

**A Thesis Submitted for the Degree of PhD at the University of Warwick**

**Permanent WRAP URL:**

<http://wrap.warwick.ac.uk/152001>

**Copyright and reuse:**

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it.

Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk)

**Autonomy, an Ethics of Freedom,  
and the Question of Answerability**

**by  
Lucy Clare Barry**

Submitted for the degree of  
Doctor of Philosophy in Philosophy

University of Warwick, Department of Philosophy

September 2019

## Table of Contents

ACKNOWLEDGEMENT.....	4
DECLARATION .....	4
ABSTRACT .....	5
INTRODUCTION .....	6
CHAPTER ONE: A SARTREAN CONCEPTION OF HUMAN AUTONOMY.....	9
Sartre's theory of Consciousness .....	11
Pre-reflective consciousness.....	13
The transcendence of the ego.....	15
The origin of nothingness .....	17
The relation of consciousness to its past.....	20
Freedom, reasons for action, and self-determination .....	21
Value.....	24
The Fundamental project .....	27
CHAPTER TWO: OBJECTIONS TO SARTRE'S ACCOUNT AND THE OBLIGATION TO TAKE UP OUR FREEDOM .....	33
Facticity.....	33
Objections to Sartre's account of Initial choice .....	38
The question of determinism.....	41
The determinist argument.....	42
An incompatibilist response .....	42
The reason view, a partial incompatibilist response .....	45
A compatibilist response .....	46
The obligation to live as if we were free .....	48
CHAPTER THREE: AN ETHICS OF FREEDOM.....	52
Kant's conception of freedom as rationality. ....	53
Acting within the moral domain .....	57
Sartre, and the absolute value of freedom.....	59
The goal of completion.....	62
Valuing the other .....	64
Perceptual theory of emotions .....	65
The need for empathy .....	68
Conclusion to chapter three. ....	71
CHAPTER FOUR: THE QUESTION OF MEANING AND ANSWERABILITY .....	72
Defining the question .....	73
Intrinsic meaning.....	74
Answerability .....	78
Meaning and answerability as experienced .....	82
CHAPTER FIVE: UNDERSTANDING METAPHOR.....	87
Creative metaphor theories.....	88
The incorporation of emotion and value.....	95
The dynamic capability of metaphor.....	100
CHAPTER SIX: RADICAL NOVELTY .....	108

The emergence of radically new objects .....	108
Accounts of creativity .....	111
The intelligibility of radically new objects .....	112
A refutation of a naturalistic account of creativity .....	118
The role of concepts .....	123
Metaphor as a model of the creative process .....	125
<b>CHAPTER SEVEN: THE CREATIVE PROCESS.....</b>	<b>129</b>
Creativity as a collaborative process .....	136
Imagination and the unreal object .....	139
Aesthetic experience and the creative process as a way of learning.....	145
An account of the creative process based on Sartre's aesthetic experience .....	150
Conclusion to chapter seven. ....	152
<b>CHAPTER EIGHT: CONCLUSION .....</b>	<b>157</b>
<b>BIBLIOGRAPHY .....</b>	<b>166</b>
<b>FIGURES .....</b>	<b>173</b>
Figure 1: Papyrus of Ani.....	173
Figure 2: The Lion Man .....	174

## ACKNOWLEDGEMENT

Throughout the duration of this project I have been overwhelmed by the support, guidance and encouragement of my two supervisors, Peter Poellner and Eileen John. Their knowledge, expertise, patience, and sense of humour have provided four of the most enjoyable and productive years imaginable. I am also extremely grateful for the help of my GPC supervisors, Fabienne Peter and Diarmuid Costello who have been a source of positive support and help along the journey.

## DECLARATION

This thesis is entirely my own work, and has not been submitted for a degree at another university.

## ABSTRACT

This thesis has two interrelated parts. Firstly, articulating an account of human autonomy and exploring the possibility of an ethics of freedom, I argue that the radical existentialism propounded by Sartre is both intellectually credible, and morally responsible.

Secondly, I consider whether this conception of autonomy is adequate to human self-understanding, or whether we need, in some sense, to be answerable to something transcendent. Historically, philosophers addressing this question have either: appealed to some conceptually graspable domain grounding values or norms, eg. mainstream religious theology; or supposed a putative external ground to be ineffable or mysterious. However, I suggest that this need can be met from within a broadly humanist position, thus making existentialism liveable.

In making this case, I consider how metaphor can allow us to surpass concepts and literal language. Additionally, some theories claim that use of metaphor can result in an instance of something unprecedented. I analyse these radical, creative theories, to establish whether metaphors can be considered as potentially *productive* of meaning, and what kind of innovation this implies.

If creative metaphors can produce something unprecedented, complex questions arise: how do new objects, events or ideas issue from human activities? Can something be radically new, and yet intelligible? I address these questions, concluding that both questions can be answered affirmatively.

Returning to answerability, I argue that in artistic processes, humans can experience a sense of creative responsibility that encounters/responds to real resistance. This resistance is experienced as originating from outside us, and it is this that we are answerable to. In our appeal to it for a sense of 'rightness', we acknowledge that it is not only other than, but somehow greater than us. In the aesthetic object we find an exemplar of freedom and self-determination, experiencing our significance in something against which to measure ourselves.

## INTRODUCTION

If we choose to address the question of how we should live, and enquire into the possibility of a meaning and purpose to human existence, many philosophies or religions offer themselves as possible routes to fulfilment. These range from the mystical to the hyper-rational. But when we consider the magnitude of adopting a philosophy of life, simply making a choice from available options doesn't seem adequate; it generates an unnerving sense of arbitrariness. There is a sense that what we are looking for should be something based on truth, something that is THE way to achieve a life of virtue or happiness for example. And this really ought to be something that forces itself on us through its identity as 'the only true path', not something impulsively selected as one possibility out of many competing possible theories.

The idea of making a choice in this context is problematic. And as the necessity of making choices is central to the theory of existentialism – the freedom of the individual to make of his life what he chooses – existentialism as a route to a 'good life', endowed with meaning, seems unpromising. Whether we are looking for the assurance of eternal happiness, eudaimonia, a virtuous life, or immortality, radical freedom, particularly as presented by Sartre, is not an obvious candidate. Sartrean freedom is associated with terms such as nihilism, despair, and angst, and the negative tenor is reinforced by Sartre's own declaration that we are 'condemned to be free'. For Sartre, everything we are and will become results from our free choice, there are no rules or guidelines telling us what we ought to do, and we can expect no externally bestowed reward or punishment for our choices. This seems to signal incompatibility with a life of contentment and moral goodness. However, I will argue that an understanding of existential freedom can indeed be life changing, to the extent that it can lead to an ethical and meaningful life.

My argument builds from the suggestion that taking up our freedom is perhaps the most honest way of life. Humans have the unique capacity to alter the course of their lives in accordance with what they think or believe. If we consider the phenomenology of human existence, it seems to us that many of our most important decisions *are* freely taken and that the choices we make are our own. And it is, perhaps, only when we start to question this capacity that we begin to speculate about conditions or limitations: we posit gods who must be obeyed; or deterministic schemes by which our autonomy is restricted or altogether denied. In order to give these standpoints the force they need so that they can reassure us, we make them into religious imperatives, or fundamental beliefs which take on the form, if not the content, of religious faith. We then uphold these positions as absolute truths which define us.

But as religions and worldviews are always only argued for, rather than conclusively proved, adopting any one of them nevertheless involves making a choice. To take the example of determinism: if we claim that our decisions are not merely *influenced* by our environment, genetic make-up, and social conditioning – a claim which is uncontroversial – but are in fact *pre-determined*, we are incorporating something which cannot be proved into our belief system. The arguments in favour of this position may seem to us to be so convincing that its validity is beyond doubt, but because we have no way of *knowing* ultimately if we are right, it is still a matter of choice that we have accepted them. If we claim with the materialist that everything that exists is made of matter, we are making an assumption which cannot be verified scientifically. Even if every object or thought was examined and found to be material, we would still not have proved the claim that *all* things are made of matter. Alternatively, if we are absolutely convinced of the existence of a God in accordance with any of the major monotheistic religions for example, the acknowledgment that due to their mutual exclusivity, at the very most only one of them can hold the status of absolute truth, reveals the precariousness of the conviction, and the inescapability of choice.

Therefore, the element of choice in the adoption of any philosophy of life seems unavoidable. However, if we accept this inevitability, and abandon the idea of a system which forces itself on us because of its irresistible claim to be *the only* true path, this need not be understood as problematic, but rather as empowering. If we are concerned with finding an explanation about how things really are, we naturally work to find what is, for us, the most convincing explanation. Ideally, having analysed the arguments for or against the position, we will be completely convinced, and we will reach a position that seems to us the most credible account of how the world is, or the most beneficial method of ensuring a good life. But if we recognise this position honestly as having been adopted by our own free choice, we are at liberty to continually interrogate it, to make changes where necessary and even to abandon it if new information diminishes its credibility. This demands intellectual honesty. But, in defending any position in the face of challengers, our identity itself is not challenged, it is only our fundamental choices that need to be argued for. And whilst it is true that the claim ‘I choose to believe that’ is weaker than ‘I believe that’, we have the advantage of a level of protection against being in serious error. I am certainly not advocating a form of relativism, according to which whatever I decide to believe is therefore true. There remains a fact of the matter whether or not all things *are* pre-determined or whether an omnipotent God *did* create the world. But we cannot *know* this from our position as finite humans. To recognise this limitation and to accept radical freedom as our default position then becomes an act of humility, with the



potential to inspire a more compassionate attitude to the various choices pursued by our fellow humans.

But is the standpoint of radical freedom intellectually credible, and is it liveable? These are the questions I will respond to. In chapters one and two, I will argue for the intellectual credibility of Sartre's account of radical freedom, setting out the ethical implications of this in chapter three. In chapter four I consider whether the unqualified humanist position implied by Sartre is liveable, or whether we need to think of ourselves as answerable in some way to something, against which we can measure our lives. If a broadly existentialist conception of autonomy is defensible against logical and empirical objections, is it adequate to the way we understand ourselves? This question will be taken up in the remaining chapters.

In chapters five, six and seven, I outline my reasons for suggesting that a sense of meaning can be achieved from within an existentialist position, as I have defined it, thus making it liveable. In the search for something that our lives might be answerable to, or find measure in, but which is beyond our powers of comprehension, metaphor is of particular interest. The use of metaphor allows us to go beyond concepts and literal language. In chapter five I carry out an analysis of the nature and function of metaphor, and suggest that it can be used to liberate us from a static, fixed conception of the world, allowing a changed ontology. In addition to stretching the boundaries of what we can know and think, metaphor offers the potential to break through these limits by creating something new. In chapter six I pursue the suggestion that this newness is not just newness of perspective but something unprecedented; and consider whether humans actually have the capacity to introduce radical novelty into the world. In the final chapter I analyse the creative process itself and suggest that something mysterious takes place in the act of creating something new. And whilst this element of mystery cannot be fully explained, I suggest that it is in this encounter that we experience something to which we might think of ourselves as answerable, whilst upholding our autonomy and radical freedom<sup>1</sup>.

---

<sup>1</sup> Throughout this project, where it has been necessary to use a personal pronoun, I have tended to use 'he' and 'himself' rather than varying between the male and female versions, or using both, as in she/he. This is purely for the sake of clarity and brevity. I take it as axiomatic that the term 'artist', for example, is a gender-neutral term and that creative ability is equally distributed amongst genders.

## CHAPTER ONE: A SARTREAN CONCEPTION OF HUMAN AUTONOMY

It is generally accepted that, to a greater or lesser extent, humans desire freedom. We want to be free to do what we want without obstruction. But we also want to be free to decide for ourselves what it is we want; we want to have free will. Karl Jaspers writes: 'To be means to decide about being ... it is still my own self that decides what it is ... for all the dependence and determinacy of my existence, I feel sure that ultimately something rests with me alone' (Jaspers, 1969, p.56).

The level of freedom we believe ourselves to have affects both our understanding of the extent to which we can hold ourselves and other people accountable, and our capacity for self-determination. We desire not merely to witness our life but to form it from the world as we find it, and we also want to be responsible for what we make of ourselves. Freedom influences our relationships as a condition of responsibility; we hold each other responsible only to the extent that we consider choices to have been freely made. Our sense of self-identity depends on freedom of choice, we distinguish ourselves from others by the choices we make, both on a superficial level in what we wear, what we eat and what we enjoy doing, and also on a deeper level in how we relate to the other inhabitants of the world, what commitments we make and what we uphold as valuable. But perhaps more importantly, free will is significant if we are to understand ourselves as capable of influencing the way things are. If we are powerless to bring about real change in our world or our environment, our sense of purpose is threatened, and the meaning of our lives is called into question. If we believe that all decisions have been made for us and the course of our lives is pre-ordained, we cannot consider ourselves to be actively participating in our own lives in any meaningful way. More importantly, we cannot accept responsibility for putting right the things in our world that need to be changed: inequality, injustice, cruelty and needless suffering. We are distinguished from non-human animals by our capacity for self-reflection, and if we were merely passive recipients of life's events, this would be a source of endless frustration.

Perhaps we cannot prove that there is not an unknown source determining us from within or beyond our world. It may be impossible to determine with any degree of accuracy the exact level of influence our genetic make-up, our environment, and our social situation has on us, on who we are and what we do. But I shall argue that if it is even possible that we are free to decide who we are and how we live, then we are morally obliged to live as if this were the case. I will return to this idea at the end of the next chapter, having considered the account of

freedom developed by Jean-Paul Sartre as a candidate for an adequate understanding of human autonomy.

There are, of course, many and varied accounts of human freedom. My reason for focussing on the work of Sartre is that his account is representative of autonomy of the strongest kind. The sub-title of his major work, *Being and Nothingness*, is '*an essay on phenomenological ontology*', which is indicative of the scope of his project. Sartre gives an account of agency and practical reason, but his freedom amounts to more than this. Sartre does not distinguish between freedom, and the being of human reality; not only can we understand ourselves as being free, we cannot escape such an understanding. We are ontologically free; our freedom is not subject to being gained or lost. Freedom is not a property among others belonging to the essence of the human. The being of man is his being free (Sartre, 2005, p.49). It is this combination that makes Sartre's account uniquely suitable for the project of considering just how autonomous we, as humans, can really consider ourselves to be.

Sartre's views on freedom are frequently described as extreme, or radical, and often rejected on these grounds. But his account is both more complex, and, on detailed examination, arguably more persuasive, than is often thought. My suggestion is that rejection of Sartre's writing on freedom because of its radical or extreme nature is not justified. Sartre comprehensively illustrates what it means, on a practical level, to acknowledge our freedom and accept responsibility for our lives in a way that reflects a potential moral obligation. For Sartre, 'freedom is a necessary feature of finite reasons-acknowledging ('human') consciousness' (Poellner, 2015, p.222). In order to fully understand what this means, it is essential to understand Sartre's account of consciousness. This is because the question of freedom only arises for a being that is conscious. However sophisticated the programming of a robot becomes, however responsive to its situation an artificial intelligence is designed to be, it does not make sense to ask of it whether it is fundamentally free. But as conscious beings we are intuitively concerned both with our own freedom, and that of others; consciousness is central to human reality, and to explore consciousness is to explore what it is to be a human being. Crucially, for Sartre consciousness is self-determining, and responsive to reasons, and the significance of both these characteristics will be explored in this chapter.

At the start, we need to establish what we normally mean when we claim freedom, in the sense of autonomy, for ourselves. Importantly, the idea is not just to ascertain negative freedom, according to which I am free if there are no obstacles preventing me achieving my goals, but rather the positive freedom to choose, or freedom of the will. And for this, three features seem

to be crucial: the first of these is that when I act, I am not simply determined to act by antecedent states of myself, or the world. If this were so, I could not reasonably claim that 'I could have done otherwise', which is usually thought to be a basic requirement of autonomy. Neither would I be able to claim responsibility for any significant change to my character, or the effect I have on the world. Secondly, it has often been held that I can only be autonomous if I act in the light of reasons that potentially justify or warrant my action. If my actions are just selected with no justifying reasons, they are no different from random or chance acts, and so cannot be described in any meaningful way as the product of a free will. Thirdly, it could be argued - and Sartre is strongly committed to this - that the reasons motivating my actions cannot consist in facts or features of the world external to me. If my motivation did come from such facts or features, my facility for self-determination would be compromised.

I will begin by outlining Sartre's account of the consciousness of the human subject. The idea of the structure of consciousness as pre-reflective, together with freedom, characterise what Sartre terms the 'human reality'. Sartre makes some fairly obscure claims to justify this; that pre-reflective consciousness is not just consciousness of something, but also unmediated and translucent self-presence; that there is nothing to be found in consciousness, it is subjectless; remarkably he claims that this consciousness is the origin of nothingness, it is what allows for nothingness to come into the world; and finally, that consciousness separates itself both from what it is conscious of, and from what it was, it negates its past, and negates itself. Ultimately, the extent to which these claims can be justified will determine how persuasive Sartre's account of freedom is.

### Sartre's theory of Consciousness

For Sartre, every human consciousness, by which he means every experience, is intentional – it is consciousness of something - and this something transcends consciousness, it is given as distinct from it: I do not just hear, or imagine, I hear *something*, or imagine *something*. Consciousness posits something as a theme of attention, it is directed towards an object, and this directedness is necessary for something to appear to a subject as an object. An object might be something directly perceived, a chair, or a dog for example, but an object can also be something imagined, or a state of affairs; perhaps the weather in Paris today. A property can be an object, the height of a tall tree can be an object for me. Essentially, an object is presented as going beyond my consciousness of it, it is more than my current experience of it. The object perceived or imagined is the synthetic unity of an infinite number of possible appearances or

profiles, present or future. In perceiving, or imagining, a tree, I see it *as* a tree only because I assume the other side of it, which currently does not appear to me: 'The object always surpasses not only our individual experiences of it but also our knowledge of it (since there are always more perspectives from which to observe it)' (Detmer, 1988, p.13). For this reason, the object is not me, it is essentially given as 'over against me'. It is given as distinct from me, and when I return to it, either in thought or experience, on a different occasion or circumstance, I can recognise it as the same object; an object essentially has a being, or objectivity, that is not completely identified with my single perception of it. An object then, whether real or imaginary, can be said to have opacity – 'we perceive the totality "cup" in and through the perception of a single aspect or perspective of the cup. Thus the cup is perceived as having an objectivity that is not identified with our perception of it' (Catalano, 1974, p.36-7).

That our consciousness is, at least for the most part, intentional seems right: most of the time we are conscious *of* something, of an object or objects in Sartre's broad sense. However, although directed outward towards objects, consciousness is, for Sartre, necessarily also awareness or consciousness of itself, and this is less intuitive. We might assume that whilst focussing on a particular project, we are unconscious of ourselves as doing the focussing. Sartre denies this, asserting that the being of consciousness, *is* its relating to a transcendent object, and *an awareness of this relating*. In this account, self-awareness is negatively characterised as a condition in which I am given to myself, but not as an object. My mode of awareness of what I am conscious of, for example my anger, or my irritation, is given in its fullness, it has no hidden profiles. What I am angry *about* does have hidden sides, I am not fully aware of all its aspects, it transcends my experience. But I am aware of myself as being angry, and unlike an object that has hidden sides, a back or an underneath, my anger is not 'over against me', but is immediate self-consciousness: 'Just as an extended object is compelled to exist according to three dimensions, so an intention a pleasure, a grief can exist only as immediate self-consciousness' (Sartre, 2005, p.10).

The claim that consciousness is *necessarily* conscious of itself, that there can be no unconscious aspects, or no opacity in consciousness is a central premise for Sartre's claims about the freedom of consciousness. If it is true, then, phenomenologically, consciousness cannot be influenced by anything outside itself, unless that influence were itself experienced at the time. To the extent that we are not directed by anything 'outside' our own present consciousness at the time of choice, our present consciousness fulfils one condition for self-determination, i.e. being autonomous. How does Sartre justify this claim for pre-reflective consciousness?

## Pre-reflective consciousness

Sartre gives three reasons for believing that un-reflected experience is not unconscious of itself; firstly, if I am, for example, counting cigarettes (Sartre, 2005, p.9), and am asked what I am doing, I know immediately that I am counting without having to ask myself. The fact that I know immediately what I am doing, without reflection, at least suggests that I have a non-observational awareness of the experience of counting. Secondly, I *can* reflect on an experience. If for example, before reflection, I was imagining a horse, but was unaware of the *experience* of imagining, reflection could only repeat a presentation of a horse, I could not reflect on my *imagining* experience, which I know I can actually do. Thirdly, Sartre's illustration of counting the cigarettes is meant to demonstrate that pre-reflective self-consciousness is a condition of the act of counting; if I am asked, I know that I am counting, but this knowing is a combination of my immediate consciousness - that for example I have got to twelve - and the previous moments, nine, ten, eleven and so on:

these apparently separate consciousnesses are internally related to one another by virtue of having counting as their "unifying theme" ... Yet this consciousness of the stages of counting as counting consciousnesses was not positional, because at the moment I was interrupted, what I was attending to was just: "Eleven" (Morris, 2008, p.68).

So while consciousness *can* be positionally conscious of itself by reflecting on its own experience, for example reflecting on imagining, seeing, hearing, and in this mode, it makes the experience into an object, most experiencing is *not* reflective. When I am seeing a cat, I am conscious of the cat, but I am not conscious of the experience of seeing the cat in the same way that I am conscious of the cat. But, as just shown, for Sartre, I do have *awareness* of my own experience of seeing the cat, and this takes the form of pre-reflective, non-positional consciousness. In this mode, consciousness does not become an object, it does not take the form of a subject-object duality of perceiver and perceived, or knowing and known. It is non-positional [self] awareness (Sartre puts the 'self' in brackets to indicate that he is not treating the self as an object. Similarly he puts 'of' in brackets when describing consciousness as being conscious [of] itself). It is this [self] consciousness that makes self-knowledge, the subject's knowledge of its own consciousness, possible. Pre-reflective consciousness as ground of self-knowledge explains the immediacy of the awareness of my hunger, as opposed to the mediated knowledge that, for example, Jupiter is a planet or even that this (object in front of me) is a chair. Knowledge of the self, in the sense that I hold conscious beliefs about myself, is reflective, but I am aware of my hunger without having to reflect on it, and this awareness is

not subject to error – in fact it is not an episodic belief, something having propositional structure, at all. I cannot have an awareness of being hungry if I am not, in fact, experiencing being hungry.

Because objects are opaque, they cannot be *in* consciousness; consciousness (the mode of experiencing), as translucent, cannot contain any unconscious components as we have just seen; whilst being dependent on the world, consciousness ‘has no “content”’ (Sartre, 2005, p.7). One significant consequence of the translucency of consciousness is that it has the potential to reveal reality as it is: a global distinction between object-as-experienced and object-as-existing is, for Sartre, not tenable: ‘...the being of an existent is exactly what it appears... [the phenomenon] does not point over its shoulder to a true being which would be, for it, absolute. What it is, it is absolutely, for it reveals itself as it is.’ (Sartre, 2005, p.2).

Whilst consciousness is always *of* something, it is not constitutive of the being of objects, as it is for an idealist. Objects have existence independently of consciousness, and are external to it. For Sartre this is demonstrated in the experienced difference between perception and imagination that can only be explained if the object perceived exists independently of me. In *The Psychology of Imagination*, Sartre details some of the features that distinguish imagined from perceived objects: ‘imagined objects are seen from several sides at the same time’, and ‘if we turn away from them they are destroyed’ (Sartre, 2001, p.141-2). He further argues that an object that is imagined can be altered in any way the subject chooses without restriction, but by contrast, a perceived object ‘presents itself with a certain stubborn obtrusiveness. It *resists* my conscious attempts to alter it, to deny it, or even to “bracket” it’ (Detmer, 1988, p.15). This coefficient of resistance illustrates that the object cannot originate from my translucent consciousness since it is given in experience as going *against* my conscious acts<sup>2</sup>.

If pre-reflective consciousness is empty and translucent, it cannot be understood as containing a substantial self, or ego: ‘the intentional character of consciousness must be absolute: it must be all and only “of” and not an ego conscious of objects.’ (Catalano, 1974, p.11). This is problematic if we understand the ego to be what unifies our extremely diverse experiences and prevents them from appearing to us as a chaotic multiplicity of disparate events. Husserl, along with many philosophers, argues that I am aware of the ego as the *subject* of any particular act or experience in time. The ego is, for him, both immanent, in my awareness, and transcendent

---

<sup>2</sup> For a detailed discussion of the relationship between perception and imagination, see Uriah Kriegel’s essay ‘Perception and imagination, A Sartrean account’ (Kriegel, 2016, pp.245-276). I will return to this point in Chapter seven where I consider in more detail Sartre’s account of the imagination.

because it goes beyond any particular experience (Zahavi, 1999, p.146-7). If we are to deny that the ego is immanent to consciousness, how can we explain that the subject of my experience today of seeing my cat, is the same subject that yesterday imagined being on holiday? And can we explain synchronic unity – the fact that in a single occurrence of consciousness at a given time there is simultaneous awareness of different sensory experiences, hearing, seeing etc. – without the presence of an ego in consciousness?

### The transcendence of the ego

Sartre locates the ego outside the structure of pre-reflective consciousness, determining it as an object *for* reflective consciousness. In *The Transcendence of the Ego*, Sartre writes: 'For most philosophers the ego is an "inhabitant" of consciousness' (Sartre, 2004, p.1). The importance, for Sartre, of refuting this, is to rule out the idea of a substance, or any object-like entity, persisting throughout our conscious states; 'it is precisely because consciousness is pure appearance, because it is total emptiness ... it is because of this identity of appearance and existence within it that it can be considered as the absolute' (Sartre, 2005, p.12). Its status as non-substantial absolute allows consciousness to be understood as self-determining. He writes: '...it is impossible to assign to a consciousness a motivation other than itself. ... Consciousness is a plenum of existence, and this determination of itself by itself is an essential characteristic.' (Sartre, 2005, p.11). The presence of a transcendental ego in consciousness would constitute a kind of essence, which would then precede and determine action. But, as the claim central to existentialism states, existence precedes essence: 'man first exists: he materializes in the world, encounters himself, and only afterwards defines himself' (Sartre, 2007, p.22).

For Sartre, the ego is the personal / which I gain awareness of by reflecting on my behaviour and intentions. By contrast, the pre-egoic awareness, which is present when we are non-positionally conscious of an object, is the more fundamental self. As a way of explaining the unity and individuality of consciousness, the existence of the transcendental ego as a transcendence that is also immanent in consciousness is, in Sartre's account, neither necessary, nor even possible. We have seen why he thinks that it is not possible, being incompatible with the character of pre-reflective [self] consciousness. But on what basis does he claim that it is not necessary, that we do not need such an ego to account for the possibility of the unity of experience, or for the fact that experience is always, and for each for-itself, *my* experience?



Sartre first argues that synchronic unity can be understood without postulating an ego. He writes: '... consciousness is defined by intentionality. Through intentionality it transcends itself, it unifies itself by going outside itself.' (Sartre, 2004, p.6). Intentionality connects consciousness to the world, and the world is already an interconnected multiplicity. For Sartre, the synchronic unity of consciousness is found in the ordered array of objects that make up the world of our experience.

In order to explain unity within duration, the fact that there is a directional, and meaningful, stream of consciousness in time – seen in the counting of cigarettes - rather than a series of unconnected happenings, Sartre refers to Husserl's account of time consciousness. Each current episode of consciousness depends on the retention of the just passed episode in order to be what it is. The elapsed episode forms part of the identity of the current episode forming one whole: 'to speak of "a consciousness" is to speak of the whole of consciousness, and this singular property belongs to consciousness itself, whatever its relations with the I may in other respects be' (Sartre, 2004, p.7). Williford writes:

If Sartre is right, one needs only internal time-consciousness to account for the diachronic unity of the stream of consciousness... There is a recurrent or invariant structure of non-positional or pre-reflective self-consciousness here, but that is not a transcendent quasi-thing-like *entity*; it is more like an ever-repeated *property*, the instances of which are temporarily linked. (Williford, 2011, p.200).

The most obvious role that we tend to assign the ego is that of providing a sense of 'mineness'. But for Sartre, consciousness' individuality derives from its nature. Non-positional self-consciousness is characterized by the first-person mode of givenness. As Zahavi writes: 'an adequate phenomenological description of lived consciousness will simply not find any ego, understood as an inhabitant in or possessor of consciousness.' (Zahavi, 1999, p.140). Consciousness *is* for Sartre, a unified and individuated stream 'for what confers personal existence on a being is not the possession of an Ego – which is only the *sign* of the personality – but it is the fact that the being exists for itself as a presence to itself.' (Sartre, 2005, p.127).

In Sartre's account then, pre-reflective consciousness is empty of content; it is fully present to itself as translucent, and even the ego is located outside it. But Sartre goes on to make the initially obscure claim that consciousness is the origin of nothingness. The argument for this claim is that if pre-reflective consciousness is empty, consciousness must be a kind of negativity, and in order for negativity to exist, there must be *experience* of lack or absence. The experience of negativity shows up in claims that: intentional consciousness *is not* its intentional

object, it is not identical to that which it is conscious of; and intentional consciousness *is not* its past, it is not determined or constituted by its past beliefs, desires or experiences.

Consciousness must be constitutively separate both from the object world, and from its past in order for Sartre's account of freedom to be viable. It is this identification of consciousness with nothingness that accounts for the claim that neither things in the world, nor my past can determine my present conscious choices, and so it is central to Sartre's account of autonomy.

### The origin of nothingness

Pronouncing the centrality of nothingness in Sartre's system, Gardner writes: 'If the philosophy of B&N were to be reconstructed in the form of a system based on a single principle, that principle would be the identification of human being with nothingness' (Gardner, 2009, p.71).

The reason why this identification is so important is that if consciousness is constitutively separate from the object world and its own past, and if it can then be understood as self-determining, the dramatic consequence ensues that all my choices or commitments must be made anew out of nothing. And it is only when we can assert this that we can consider ourselves to be fully autonomous at any one time.

How then is nothingness to be understood? Sartre describes the mode of being that is consciousness *as* nihilation; 'the Being by which Nothingness arrives in the world is a being such that in its Being, the Nothingness of its Being is in question. *The being by which Nothingness comes to the world must be its own Nothingness.*' (Sartre, 2005, p.47). The confusion generated by the claim that 'being is its own nothingness', is compounded further by the assertion that consciousness is nothing, and also that nothingness is ontologically real. Reality, the totality of existents, includes both things that exist in the mode of *being*, and 'things' that exist in the mode of *nothingness*. Nothingness depends on being, and is brought about by the nihilation of being. Reality contains negation because of the power of consciousness to negate. At this point, Sartre's account of nothingness appears to be both technical and obscure, suggesting little in the way of practical application. However, he goes on to explain that nothingness is discovered phenomenologically, as a concrete feature of experience. So, if nothingness is a concrete feature of experience, where does it come from? Nothingness cannot be produced by being-in-itself – the term Sartre uses to denote the object world - because being-in-itself is full positivity, being-in-itself lacks all relation with nothingness; a chair just *is* a chair, it lacks nothing. Nothingness cannot nihilate itself because it *is* not. The claim is that nothingness comes into the world through pre-reflective consciousness. Consciousness is not perfect equivalence either

with itself or its object. I am always pre-reflectively aware of my separation from the object of my consciousness; I am different from that object. The lack of identity with itself of the human being *qua* object, is the nothingness within it. If, for example, I am a plumber, or an Englishman, these are purely the characteristics that I am freely living; I am not defined by them. A chair is defined by having the characteristics of a chair, but no object-characteristics can define my existence: 'Concrete nothingness is the constant "elsewhereness" of consciousness.' (Catalano, 1974, p.64).

According to Sartre, the fact that we can make negative judgements presupposes the possibility of experiencing lack or negativity: 'Nothingness must be given at the heart of Being, in order for us to be able to apprehend that particular type of realities which we have called *négarités*.' (Sartre, 2005, p.46). The posing of a question provides one instance of negation in Sartre's account. Questioning, for Sartre, includes all human attitudes of doubt, or mere expectation; the human is a being for whom everything is in question, and he maintains that everything that can be thought is potentially either the subject of a question or the answer to one. As Taylor articulates, 'the human subject is such that the question arises inescapably, which kind of being he is going to realize' (Taylor, 1982, p.111-2). Unity, achieved by the in-itself – a chair *is* a chair – is an aim that consciousness never realises. By adopting the attitude of questioning, three negations are introduced: The person asking the question does not know whether the answer will be negative or affirmative – there is a conscious non-being of knowledge. Secondly, in the case of a negative response, there is a transcendent fact of non-existence. And thirdly, in the case of an affirmative reply, the existence of a truth is implied, and the situation is acknowledged as "'It is thus, and not otherwise.'" In a word the truth, as differentiated from being, introduces a third non-being as determining the question – the non-being of limitation' (Sartre, 2005, p.29).

The nihilating role consciousness plays is also observable in the form of destruction and fragility. Sartre claims that the human is the only being through whom destruction can occur. Using the example of a storm, he argues that a storm can only be said to have caused *destruction* if it is experienced as having destroyed something. There must be a relation of man to being, in which man comprehends this being as destructible. Without this relation, the storm merely *redistributes* what there is. Similarly with fragility, for something to be fragile, there must be a relation to being in which man recognises the probability of its non-being in the face of certain conditions. Man, as witness to destruction nihilates by recalling a past state of affairs, which is no longer. It is important to note that whilst for Sartre, destruction comes into being through man, for him it is also an objective fact.

Although non-being appears only within the limits of human expectation, if negative judgements presuppose non-being, negation clearly cannot be reduced to a function of judgement. This point is demonstrated in the phenomenon of absence, seen in Sartre's illustration of an anticipated meeting with a friend Pierre, in a café at 4 o'clock (Sartre, 2005, p.33). In this example, I, the subject, arrive late, and on arrival I have an intuition that Pierre is not there, I *see* that he is absent. I encounter a negativity. In itself, the café, with its tables, people, sounds, is a fullness of being. But when I enter the café in the expectation of seeing Pierre, this is given to me as a synthetic organisation of objects, in which Pierre is about to appear. Pierre's absence is given to me via the café as a double nihilation: the café is the ground on which the object of my attention will appear, and as such, everything within the café is given to me as *not* what I am looking for. This is the first nihilation. In the second nihilation, the figure of Pierre is posited as not actual, it is a perpetual disappearance. Pierre raises 'himself as nothingness on the ground of the nihilation of the café' (Sartre, 2005, p.34). Pierre's absence could not occur without my conscious expectation of seeing him, but this does not make his absence subjective. His absence is also an objective fact; he *is not* in the café. But the nothingness of Pierre has been *introduced* by my conscious act of expecting him.

Concrete nothings then, are transcendent to our consciousness. This point is emphasised in the observation that they can be distinguished from fictional nothings, for example there is a clear distinction between what is intended by the claim 'Pierre is not in the café' and the claim that 'there are no unicorns in this café'. Concrete nothing is always the absence of *something*. This account of nothingness then, can be summarised in four statements: (i) We have an immediate pre-judicative comprehension of non-being. If I am wondering whether my cat is in the shed, I am not at the same time judging that I do not know whether my cat is in the shed. I am not making a judgement on the non-being of my knowledge, but I do have a non-positional awareness of not knowing. If asked, I can reply that I do not know immediately, without asking the question of myself, or observing, or making inferences; (ii) Negations are not given as subjective. Whether or not Pierre is in the café is an objective fact; (iii) If negations were purely judgements, the category of negation would be unintelligible – if all is supposed positivity, judgement could not conceive of a negative form; (iv) 'Nothing' assumes concrete forms. In the example of Pierre and the café, the café, which is organised into a synthetic totality, 'presents' the figure of Pierre as 'not being there'. Pierre's absence is the foundation for the judgement that Pierre is not in the cafe. Negative judgement then is subsequent to non-being, and non-being is a component of the real.

Having characterised nothingness as transcendent to consciousness and as an objective reality introduced by consciousness, Sartre considers the relation of consciousness to its past in the light of this description. The past is often believed to play a major part in determining what we are and what we will become. If this is so, our sense of freedom is threatened. In Sartre's phenomenological account, he claims that the concept of nothingness demonstrates that our past does not determine us, but rather as past, it is put out of play.

### The relation of consciousness to its past

Consciousness cannot be identified with my mental states or emotions, because all these can be turned into objects for my consciousness; for example I can reflect on my happiness as a state of being which affects me, and of which I am conscious, but I am given to myself as distinct from everything that I am conscious of within the world. This also applies to my past experiences; my consciousness is distinct from all beliefs and representations about its past. In the act of reflection on my past acts I separate them from me, and so neutralise them. Sartre's point about the negation introduced by questioning becomes particularly helpful in explaining this: in questioning, the questioner breaks free from any continuity with being, creating the real distance that is necessary for the possibility of the being or non-being of what is questioned. The break with the causal series is not a break in the temporal flow, but a separation caused by nothing: 'What separates prior from subsequent is exactly nothing. This nothing is absolutely impassable, precisely because it is nothing' (Sartre, 2005, p.52). This break puts the past out of play for me, the past retains a relationship with my present consciousness but one where it is bracketed out: 'It constantly maintains a relation of interpenetration with the present consciousness, but on the basis of this existential relation it is put out of the game' (ibid). Sartre illustrates this point with the example of a gambler who has resolved to give up gambling. In the face of the gaming table, temptation makes it clear to the gambler that he *is not* his past resolution. In order for the resolution to be meaningful for him, the gambler must remake it. Anguish experienced in the realisation that there is *nothing* preventing him from gambling reveals that he is not identified with his past resolution: 'What the gambler apprehends at this instant is again the permanent rupture in determinism; it is nothingness which separates him from himself' (Sartre, 2005, p.57).

We have seen that in his theory of consciousness, Sartre makes two claims that are particularly significant for his conception of freedom: (i) Consciousness is not determined by its past. It has awareness of its past, but its past is put out of play. (ii) Consciousness *is not* the objects that it

encounters in the world. Consciousness involves a distancing of itself from everything that it *is not*, both its past and what is given to it in the present. For Sartre, freedom involves a forcible break from the world. However, the overall aim of forming an adequate conception of human autonomy demands that we go further than this. Not being determined by my past nor being the objects that I encounter are necessary but not sufficient conditions for absolute freedom. In order that our acts should not be understood as chance or random, Sartre needs to show how consciousness is self-determining: “to be free” does not mean “to obtain what one has wished” but rather “by oneself to determine oneself to wish” (in the broad sense of choosing)’ (Sartre, 2005, p.505). In the following section of this chapter I will look at how, in Sartre’s account, we decide for ourselves what we want, and show how *reasons-responsive* intentional action is constitutive of human freedom.

### Freedom, reasons for action, and self-determination

In Sartre’s account, our freedom consists in our conscious decisions to act, and our decisions being *responsive to reasons* that are ultimately *determined by ourselves* in the sense that our ends are determined by consciousness itself (Poellner, 2015, p.225). Intention, in the fundamental sense, and act are inseparable; our freedom is autonomy of choice – Sartre intends ‘choice’ not as dream or wish, but as synonymous with acting which supposes a ‘commencement of realization’ in order to distinguish it as such. If freedom *qua* autonomy does not mean obtaining what one wishes, success is not important to freedom (Sartre, 2005, p.505). What is central then is rationality; we need to have *good reasons* to believe, and to act in the light of those beliefs. We need to have the means of linking up our actions with our beliefs, of choosing the right means to achieve our ends, and of forming goals that make sense to us, in order to consider ourselves as free and responsible. Significantly, in Sartre’s account, what counts as a good reason is not found by adjusting correctly to the way the world is, but must ultimately be determined by ourselves. For an action to appear to me both as meaningful, and as ‘mine’, I need to recognise it as responsive to reasons. In order to clarify how Sartre can claim that our actions can be meaningful in this way, I will now consider how, for him, intentional action is structured.

If goals seem well chosen and worthwhile, they appear to give us reasons for action. We are phenomenally aware of the apparent reasons and this awareness enables the reasons to guide our behaviour. Sartre’s structural account of reasons-responsive actions demonstrates that in his account, our ultimate ends are not received from outside consciousness, but are produced

by it: 'human reality in and through its very upsurge decides to define its own being by its ends. It is therefore the positing of my ultimate ends which characterizes my being and which is identical with the sudden thrust of the freedom which is mine.' (Sartre, 2005, p.466). For Sartre, there are four elements that need to be involved in an understanding of intentional actions as meaningful; a) the *motif*; b) the decision, or choice; c) the end; and d) the *mobile*, or motive. To act is to organise means in view of valued ends, and to modify the world such that there is an essentially interrelated complex of *motifs*, *mobiles*, and ends, in which modifications are linked to a series to produce a desired result.

If my freedom is autonomy of choice, and my conscious decisions to act are determined only by consciousness itself, how do I decide whether to take one course of action rather than another? In Sartre's account the reason for the act is the *motif*, the 'ensemble of rational considerations which justify it' (Sartre, 2005, p.468). The objective state of affairs is grasped by the agent in evaluative terms: 'this objective appreciation can be made only in the light of a presupposed end and within the limits of a project of the for-itself toward this end' (ibid). A *motif* might be instrumental, a means to something else, perhaps a fact that can be used as an instrument for achieving specific ends, and as instrumental this requires a grasp of means/ends relations - the agent must have an understanding of the situation that is instrumentally relevant to recognising an objective feature as suitable for achieving the end that he is committed to. The instrumental suitability of the action may involve a process of deliberate, rational justification, but equally, in non-reflective action, it may involve conscious perceptual registering – in which the environment is given to the agent's consciousness as affording specific obstacles to be overcome, or opportunities to be taken up.

Secondly, an intentional, potentially reasons-responsive action requires an awareness at the time of action of a commitment to the realisation of an end, and of performing the action required in order to achieve it. This consciousness of effective desire is usually non-thematic, and it is what Sartre calls 'choice'. The choice is experienced as the subject's own, and is a response to his grasp of *motifs*. 'It is that aspect of the experience of action that grounds the agent's normal non-observational and non-inferential ability to say, with respect to many of her behaviours, under some descriptions: "I did it", in contrast to other behaviours which are truthfully reported by her as: "it happened to me"' (Poellner, 2015, p.226). This can also be contrasted with a situation in which the agent might say 'I thought about doing it but did not actually do it'.

Third, all meaningful action must have an end, or goal, in view, an end aimed at for its own sake. This intrinsic end may be a feature of the action itself; the action may be performed for its own sake. The agent pursuing an intrinsic end needs to have a conscious, but not necessarily thematic, awareness of that end, although some of the features that make the end desirable may be implicit. An intrinsic end must be characterised in evaluative terms, it is a reason for an action if it has the value of being good or desirable. Human reality chooses its ends, and by this choice, 'confers' or acknowledges value in what is chosen. For Sartre, a positive value has the nature of something that ought-to-be, or in the case of a negative value, ought-not-to-be. In *War Diaries*, Sartre writes that he initially rejected the idea that something should have the nature of an ought-to-be, considering that this would suggest a connection with the metaphysical realm, and a representation by the categorical imperative (Sartre, 1984, p.88). But later, following Scheler, he acknowledged the existence of values: '(I) understood that there existed specific natures, equipped with an existence as of right, and called values ... (that) regulated each of my acts and judgements, and that by their nature they "ought to be"' (ibid). A positive value can be described in phenomenological terms as 'a feature or property that is necessarily such as to merit, *pro tanto*, being instantiated: it ought to be' (Poellner, 2016, p.139). It is only in respect of having an end that an agent can have a *motif* – a *motif* is always in relation to an end.

Finally, a motive (*mobile*) is the conscious mental state, 'the ensemble of the desires, emotions and passions' (Sartre, 2005, p.468), in which an agent grasps a possible end as a value and is motivated to pursue it. As such, it is subjective, motivation comes from the evaluative content of the belief at the affective level, which justifies the action from the perspective of the agent. Being conscious [of] something as a conclusive *mobile* involves taking steps to pursue that end.

This fourfold structure describes how the evaluative components give us reasons for action. But can I simply 'choose' that an action is valuable for me? What would it mean to say that I 'confer' value on ends through my choice of them? Intuitively it seems right to say that it is a necessary condition of freedom that my ends are worthwhile ends, but what does it mean to consider something to be worthwhile? In Sartre's account, value plays a fundamental role, and an explanation of this is important for an understanding of what it means for consciousness to be self-determining.



## Value

Sartre denies that freedom is arbitrary or capricious (Sartre, 2005, p.475), actions can only be intelligible to me if their ends are taken as valuable in some sense. All ends are specified in terms of some apparent good, which motivates me to pursue it as my end. Phenomenologically, in some cases values do appear to us as the property of actual or possible objects, and these values place demands on us. The beauty of a piece of music appears as a property of that music, which demands our appreciation. These demands are both normative and valuational; when I am conscious of a possible course of action as 'acceptable' or 'unauthorised' for example, I am conscious of a normative demand. A valuational demand can be identified in the acknowledgement that something is good or bad, that it ought to be actual if it is good, and ought not to be actual if bad. Valuational demands can exist without the involvement of any normative demands, but normative demands depend on valuational demands.<sup>3</sup>

The understanding of consciousness as nothingness, central to his ontology as a whole, underpins Sartre's account of value; that actions are carried out as a means to achieve worthwhile ends, or are themselves such ends, introduces the acknowledgement of an objective lack. The intention is the realisation of what is a desirable, but *not yet* realised possibility. In the first consideration of the idea of the act, consciousness leaves the level of being, the full world of what is, and approaches the non-being: '... it is by a pure wrenching away from himself and the world that the worker can posit his suffering as unbearable suffering and consequently can *make of it the motive* for his revolutionary action' (Sartre, 2005, p.458). Acknowledging the situation of lack, in turn introduces the concept of value – a lack only appears as such in the face of an acknowledgement that things could be otherwise, that a desirable situation is possible but not yet actual.

Whilst acknowledging that the way the world appears to us is that objects have values, such as beautiful or disgusting, attached to them, Sartre claims that values appear at the level of being only because the for-itself, or consciousness, brings them into being; nothingness and value are both regarded as *négatités*, arising only in connection with our engagements in projects. Values can only be understood in relation to nothingness, and equally, nothingness demands the conception of value. So are values real, objective properties, or are they 'projected' by the subject?

---

<sup>3</sup> The source of normative and valuational demands is the subject of much detailed debate and controversy. For the purposes of this work I will accept Sartre's account, which appears to me to be intuitively credible.

Values do seem to have an ambiguous status; they are described alternately as subjective and objective. They are subjective in, for example, Sartre's claim that 'my freedom is the unique foundation of values ... nothing justifies me in adopting this or that particular value, this or that particular scale of values' (Sartre, 2005, p.62). In the paragraph directly following this claim, Sartre implies that values are objective: 'I am engaged in a world of values ... values are sown on my path as thousands of little real demands, like the signs which order us to keep off the grass' (ibid).

The co-existence of subjective and objective qualities is a recurring theme throughout Sartre's work. For example, in *Existentialism is a Humanism* he refers to the limitations of the human condition as having an objective dimension, because they affect all humans everywhere, and a subjective dimension because to be meaningful they must be experienced by individuals (Sartre, 2007, p.42). In Sartre's example described above, the absence of Pierre from the café does not 'exist'; his absence is not *in* the café. The café is a fullness of being, all positivity. Pierre's absence is encountered as a result of my project of looking for him. His absence has the appearance of subjectivity; it exists only in relation to my consciousness. However, Pierre's absence *is* real, if he *was* there, I would have seen him. Equally values are subjective in their origination in consciousness, in relation to projects, and because they have no independent existence in the world: 'values do *not* exist, or more precisely ... they owe such existence as they have to human consciousness which brings them into being' (Detmer, 1988, p.136-7). But again, they are also objective in that they, or many of them, are not simply invented, they can be discovered, or encountered in the world. In *Intentionality*, Sartre writes 'Being dreadful is a *property* of this Japanese mask, an inexhaustible and irreducible property which constitutes its very nature' (Quoted in Detmer, 1988, p.178).

It is this understanding of values as having both a subjective and an objective dimension that counters the accusation that, according to Sartre, I can choose anything at all, and confer a positive value on it at random, by making it an intrinsic end. This is not Sartre's position, if it were, not only would it be morally problematic, it would fail to explain choice - choice is dependent on justifying reasons, which are grounded in consciousness of values, it therefore cannot be said to create them. As Taylor writes: 'a radical choice *between* strong evaluations is quite conceivable, but not a radical choice *of* such evaluations' (Taylor, 1982, p.119). The existence of objective values discovered in the world rules out the suggestion that we randomly confer value at will. Sartre's specific understanding of choice is explained in his concept of the fundamental project, which will be discussed below. But the particular way in which subjectivity and objectivity co-exist provides a balance that avoids the difficulties arising from an

understanding of values as holding one or other status. This balance is further displayed in unreflective affective consciousness, in the behaviour of the emotions and passions.

Things that we encounter in the world, phenomenal objects, reveal themselves to us as likeable, repellent, terrifying: 'many emotions are (re)presentational acts and ... what is given in them are apparent value characteristics of phenomenal objects' (Poellner, 2015, p.229). The value characteristics merit response, and demand actions; 'my indignation has given me the negative value "baseness", my admiration has given the positive value "grandeur" (Sartre, 2005, p.62). If my indignation gives to me the negative value baseness in respect of Fred's treatment of his cat, this merits my negative emotional response, and demands that I take action, for example to persuade Fred to act differently: 'emotions such as indignation have literally, an intuitive character: they present a certain characteristic phenomenology which is representational' (Poellner, 2015, p.229). Some emotions then can be understood as grounding choices and rationalising actions. In the case of Fred's treatment of his cat, I experience what I perceive to be cruelty through indignation; my affective experience presents an evaluational demand, and also a normative demand, both of which are present in my indignation. The evaluational demand presents a disvalue – Fred's cruelty to his cat ought-not-to-be, and the normative demand is that I act to stop it.

However, does the fact that something is given to us as worthwhile demand that we act for it, rather than against it? The instance of someone affected by apparent weakness of will - where an agent's desires seem to be inconsistent with his values, calls into question whether his acts are still free. If I say that I consider it worthwhile to spend the afternoon clearing snow from my neighbour's path, but instead go to the zoo, does my apparent weakness of will not reveal a lack of freedom in the face of my own desires? For Sartre, it is how I act that reveals my real commitments. In *War Diaries*, he writes of the obligation to shoulder what happens to me: 'at the moment when I lose my grip, when my body "overcomes me", when under physical torment I confess what I wanted to keep secret – it is of my own accord, through the free consciousness of my torment, that I decide to confess' (Sartre, 1984, p.113-4). I may disapprove of my decision in retrospect, but for Sartre, it is reflective of my values (hence my ends) at that time. For Sartre, when influenced by emotion, I still act freely. In the example given, my actions reveal that at the time, I did not actually consider clearing the snow from my neighbour's path to be a worthwhile use of my afternoon, and going to the zoo instead was indeed a freely chosen act.

In Sartre's account of freedom then, our actions are responsive to reasons that apparently justify the choosing of them. All actions are intentional, and there is no distinction between an intention in the fundamental sense, and an action, or between an intentional action and a free action. This account allows for us to make free choices of means in relation to valued ends. But if these ends are not to be just a fragmented collection of things that seem to us worthwhile at the time there must be a unifying component to the structure; 'there is a unity, unique to each individual, to the end toward which each person acts: the "fundamental project"' (Morris, 2008, p.145). This ultimate, or unifying aim, or end, of consciousness is to be self-determining. What consciousness desires is not some objective good that the world can satisfy; consciousness surpasses the world and ultimately desires its own transformation. If consciousness determines itself, and objects cannot act upon consciousness to determine its actions, the *fundamental* reasons of a for-itself cannot be found in objects. Where then can these reasons be found? An explanation of what it means for consciousness to be self-determining can be found in the concept of the fundamental project.

### The Fundamental project

The fundamental project is the overall orientation towards one's life as a whole, which lays the foundation from which other purposes emerge. All our local, limited goals and ends are subject to the question 'why?' - ultimately their importance can only be understood in relation to the fundamental project, which uniquely is not intelligible in terms of a further project; 'Consciousness organizes the world in terms of "instrumental complexes" or means-end relationships which are an expression of its overall goal or "project"' (Jones, 1980, p.234). Sartre uses the term project, not only to denote acts that aim to bring about something that is not yet actual, but in a wider sense, for example he refers to the 'projects' of marriage, parenthood and even staying alive. In *Huis Clos*, Sartre gives an illustration of a soldier deserting the army and fleeing to safety, describing both as projects – the actions involved form a series which is motivated by a specific aim, and this gives the actions their meaning. Some projects are a way of pursuing larger projects, forming an integrated hierarchy. However this hierarchy is not open-ended - our projects, which are manifested in behaviour, are unified in the fundamental project. Trivial actions, such as putting on a coat for example, can be understood as manifesting the project of keeping warm, which in turn is a way of pursuing a deeper project, maintaining health for example, and so on. Reasons and motives are ultimately derived from our fundamental project 'it is this present choice which originally creates all causes and all motives which can guide us to partial actions; it is this which arranges the world with its

meaning, its instrumental-complexes, and its coefficient of adversity' (Sartre, 2005, p.487). All our actions then can be understood ultimately as ways of pursuing our fundamental project, although for Sartre, every action, however small is also constitutive of the identity we choose for ourselves; 'Our acts are related to our fundamental project, not as a conclusion to a premise, but as a partial structure to a totality.' (Catalano, 1974, p.200).

The choice of our fundamental project constitutes our 'essence' at a time. The original choice that we make of ourselves is what creates our reasons and motives. It is completely distinctive and is also changeable – we can adopt a new set of values – although Sartre acknowledges that this is not easy and is not undertaken lightly (this is a subject that I will return to in the next chapter); our values characterise who we are, and in choosing a different set of values, we choose to become a different person.

But the concept of choice, used in relation to the fundamental project, must be understood in a particular sense, specific to Sartre. Although our projects are in a sense chosen, this does not mean that we necessarily thematize that choice, or that we reflectively decide to continue upholding it. Our projects do not necessarily result from such clear and distinct decisions. Fundamental projects influence our specific aims at any given time, they direct the way the world seems to us, and how we react emotionally and rationally to it. But whilst we are aware of the structures of experiences, and of how we think and feel about them, we need not have a thematic awareness of the deeper reasons behind them, or of why we react in the way that we do. For Sartre, what reflection grasps 'is not the pure project of the for-itself as it is symbolically expressed ... by the concrete behaviour which it apprehends' but rather 'the concrete behaviour itself; ... the specific dated desire in all its characteristic network' (Sartre, 2005, p.591).

The concept of the fundamental project is illustrated in *Being and Nothingness* with the story of a hiker, walking with friends. After several hours, facing extreme fatigue, the hiker sits down, throws down his pack and gives up, while his friends go on walking. A determinist account of the hiker's reasons for giving up would focus on the gradient of the hill, and the physical condition of the hiker. But the phenomenological reason for the hiker stopping is simply the subjective assessment of the hill as being *too steep* and his body *too tired*. The actual gradient of the hill cannot be said to be the cause of his giving up, because his friends, who are equally fit, continue to walk. They also experience fatigue but choose to overcome it and continue. It is the hiker's interpretation of the steepness and his fatigue as *too much* that is significant. For Sartre, the hiker stops because he chooses to suffer his fatigue in this way, and this is a part of

his larger, or fundamental, project. The hiker could have done otherwise, he could have continued with his companions. But his fundamental project is such that he has chosen to be someone who in this instance considers his fatigue unbearable and gives up. The hiker's partial end, for example relieving his fatigue, cannot be said to compel him to give up because, as partial, it relates only to one specific part of his situation, rather than to the meaning of his life as a whole; the partial ends that he acknowledges are understood in the light of his ultimate end. The hiker is free to change his fundamental project and undergo a radical conversion, in which case he might enjoy his fatigue as a sign of his mastery over his body and nature. The difficulty involved in bringing this about, explains why he mistakenly experiences his situation as one where he could not but have stopped.

In accordance with the understanding that the for-itself cannot be affected by anything external to it in a way that determines its actions, Sartre argues that: 'events can affect you only if they're assumed by your own possibilities' (Sartre, 1984, p.96). The for-itself is affected by the valuational and normative demands characterised by the emotions: 'In being "moved" by such a demand (by acting on it), an agent is not simply passively caused to behave in some way or other, but necessarily actively envisages some possibility – an as yet absent, futural act or worldly state – as a value to be realized' (Poellner, 2015, p.231).

So, our motivation for action comes from our fundamental project, even when this is not consciously thematised. But if I am responsible for my choice of fundamental project and this choice constitutes a choice of myself, who is it that makes the choice, and what motivates this choice?

Again, it is important to understand choice as intended by Sartre. The *original* choice of fundamental project, unlike other choices we make, is not an action. Rather than regarding it as a choice, it might be better understood as a way of being-in-the-world, or as a commitment. Sartre himself refers to it as a fundamental attitude; 'If the original choice of a fundamental project is not an action then questions about when it was made are out of place ... Also out of place are questions about why – for what reasons and motives – the "choice" was made' (Morris, 2008, p.153-4). I am not responsible for my initial choice on account of being the author of it, but because I *am* it. The 'I' that creates the initial choice cannot be separated from it. Self-creation through that choice is not an act that I perform but, in de Beauvoir's words, 'the original upsurge of our existence' (Beauvoir, 1976, p.25). Because the fundamental project cannot be interpreted in terms of another project, without existential psychoanalysis, we cannot make it explicit. All our individual choices and acts are partial projects with a margin of

contingency, unpredictability and absurdity. But each of these partial projects is 'the specification of the global project on the occasion of particular elements in the situation and so is always understood in relation to the totality of my being-in-the-world' (Sartre, 2005, p.502).

The motivation for the choice of the fundamental project, also referred to by Sartre as specifying a for-itself's ultimate end, can be found in the for-itself's determination of itself as a lack of being; it 'is perpetually determining itself *not to be* the in-itself' (Sartre, 2005, p.109). The for-itself can only establish itself in terms of the in-itself, and against the in-itself, as that which it *is not*. In the face of this ontological lack of being, the for-itself points towards a totality of itself, or coincidence with itself, that would imply satisfaction - the removal of desire. Illustrating this point with the example of hunger as lack, Simont writes 'without this totalizing aim, and assuming hypothetically a purely external lack as assailing the for-itself, hunger would be blind, inexpressible panic' (Simont, 1999, p.179). Instead, hunger is organised behaviour due to the projected satisfaction of nourishment: human reality is motivated towards the projected satisfaction of the completion of the ultimate end.

The ultimate end of the for-itself is then the overcoming of lack or deficiency, a desire for 'completeness', which Sartre refers to as being an in-itself-for-itself, a fully self-conscious, necessary, self-grounding object, which although a personal being, would have a fixed nature. Sartre details both authentic and inauthentic ways in which the for-itself can pursue its ultimate end. Whilst both ways acknowledge that consciousness lacks identity with itself, in inauthentic pursuit, described by Sartre as 'bad-faith', the for-itself aims at completeness or wholeness, and implicitly takes this to be a realisable end. However, this totality is unattainable. Personal beings are necessarily conscious and self-aware, and these essentially dynamic, project-involving qualities cannot be understood as part of an unchanging being. Because the nature of the for-itself is transparent, and that of the in-itself is opaque, the for-itself and the in-itself are not compatible, and so the concept of achieving identity is incoherent – this would be 'an impossible condition in which the for-itself would be both consciousness and entirely a non-contingent, self-grounding object' (Poellner, 2015, p.233). If the freedom of the for-itself is secured by its self-determination, pursuit of the fundamental project in the belief that it is achievable, pursuit in bad faith, amounts to a denial of this freedom. In an authentic pursuit, the project of achieving completeness is not given up, but it is acknowledged as perpetually incomplete.

However, even when pursued authentically, the idea of the fundamental project presents a problem for reasons-responsive action. Sartre suggests that it is based on the ontological

structure of human reality: 'This mustn't be understood as an empty psychological desire, but as the transcendental structure of human reality' (Sartre, 1984, p.110). But Sartre's own conception of consciousness means that the desire for completion, if understood as an ontological goal, is not only impossible, but incoherent – as discussed earlier, consciousness fundamentally cannot be an object to itself. Unlike a chair, or a stone, which is just what it is, human reality is not one with itself. Consciousness, which *is*, only in relation to objects, is continually projecting towards that which *is not*, and it is in transcending itself that it is free. Man is free *because* he is structurally different from objects; 'This lack of self contained necessary being is man's radical freedom ... He is free from the causally determined world of substantial beings-in-themselves precisely insofar as he is a lack of such being and continually transcends it' (Anderson, 1979, p.25).

This presents a problem because Sartre's own conception of consciousness as free involves that consciousness is rationally motivated – when I act, there is an at least apparent justifying reason for my action. But an incoherent, contradictory reason cannot be understood as a justifying factor. The desire to be self-founding is therefore incapable of serving as an action guiding reason. Whilst an end that is contingently impossible can still motivate, and potentially justify a pursuit, an incoherent end cannot. If Sartre is committed to the desire for completeness having ontological status, perhaps we have reached an impasse.

Bearing in mind that the claim that the ultimate end of all humans is the desire to be ontologically self-sufficient and self-founding is hard to justify, and that even Sartre fails to provide reasons for us to accept this, perhaps the phenomenological status for the fundamental project, the sense of substantiality understood as experiential absence of lack, which Sartre also suggests, is enough. To this end, Poellner proposes a modification of Sartre's account that would remove the ontological status of the aim (being an in-itself-for-itself), in favour of an ultimate end merely of completeness, or wholeness; 'There is no phenomenological warrant for thinking that any adequate existential psychoanalysis of the fundamental project of an inauthentic for-itself would invariably uncover a desire for nothing less than *ontological self-sufficiency*' (Poellner, 2015, p.244 n23). This modification involves rejecting the ontological aim of becoming a God-like consciousness, whilst allowing for the phenomenological aim of overcoming lack. The 'completeness' aim can be understood as contingently impossible but not incoherent, and so still potentially a good justifying reason - thus leaving the reasons-responsiveness of action, and therefore freedom, intact. Whilst this does involve an alteration of Sartre's stated position, it satisfactorily resolves a problem that otherwise seems insurmountable.



I think that Sartre's account of humans as self-determining beings, on account of our human consciousness, is convincing: if consciousness is empty and translucent, it cannot be influenced by anything outside itself, without that influence being experienced. However, in order to secure freedom on Sartre's account, reasons-responsiveness must also be established, and it does seem that the account of the fundamental project is problematic if taken as an ontological aim. I see no reason to reject the complete theory on this single point, bearing in mind the lack of motivation for accepting it, and also that there are no apparently serious consequences for the overall theory if it is not followed. For the purposes of this project, I propose that the interpretation of the fundamental project as an ontological aim is put in brackets. After considering some further objections to Sartre's account of autonomy in the next chapter, I will return to the question of the fundamental project in chapter three, where I consider it in relation to an ethics of freedom based on the idea of freedom as possessing absolute value.

## CHAPTER TWO: OBJECTIONS TO SARTRE'S ACCOUNT AND THE OBLIGATION TO TAKE UP OUR FREEDOM

Before Sartre's account can be accepted as a suitable candidate for an adequate understanding of human autonomy, some of the major objections that have been made against it call for a response. I will look first at an objection which centres on the phenomenological aspects of Sartre's account, the apparent constraint of our facticity - we are born into a situation, and this provides things that we are powerless to change – which suggests that absolute freedom is never achievable. Having argued that Sartre is justified in the claim that our given situation does not have to be regarded as a threat to our freedom, I will consider objections that have been raised against the concept of the initial choice, or fundamental project. Looking briefly at concerns surrounding the concept of choice itself, which I suggest can be attributed to a misunderstanding of Sartre's theory, I will look at objections which centre on the modification of the initial choice. The initial choice is the original relation that the for-itself chooses with the world, and Sartre claims that it can be renewed or modified, or even abandoned in favour of another project. But, as all our reasons for action are formed in relation to our fundamental project, Sartre needs to explain how it is possible to have a reason to change this initial choice, and I will explore that next.

Sartre's account of freedom is an ontological account, and the debate over determination is clearly significant. In the next section, I will consider three examples of the determinist argument before suggesting an alternative position that is compatible with Sartre's account and accommodates the possibility of absolute freedom. In the final section of the chapter I will argue that taking up this absolute freedom is a mark of true humanity.

### Facticity

For Sartre, although we necessarily exist in a situation – 'we discover ourselves in a world peopled with demands' (Sartre, 2005, p.63) - it is only through freedom that there *is* a situation: 'Sartre's concept of 'situation' ... is accordingly not that of a *condition* of freedom – rather my situation is an *expression* and *realization* of my *unconditioned* freedom' (Gardner, 2009, p.159). Sartre's ontological freedom always arises in context: 'Granted, it is thanks to facticity that I'm thrown into war. But what war will be for me, what face it will reveal to me, what I shall myself be in war and for war – all this I shall be freely and am responsible for' (Sartre, 1984, p.113).

As illustrated by this quote, whilst there is resistance from things in our situation and we do face obstacles, these only have meaning in and through the free choice of human reality. The world resists my projects, and there are things that I cannot change, e.g. the place of my birth. But the unchangeable is only perceived as help or hindrance in the light of my freely chosen projects. A cliff face can be a challenge to the person who wishes to climb it, but a shelter to someone wanting protection from the weather. The resistance of the cliff face to the climber is an objective fact; the cliff face offers brute resistance. But for Sartre, this brute resistance is a condition of freedom: 'the coefficient of adversity of the thing and its character as an obstacle (joined to its character as an instrument) is indispensable to the existence of a freedom' (Sartre, 2005, p.506). It is only through our freedom to choose our own goals that objects or facts, which would otherwise be neutral, develop qualities which are helpful or problematic. Sartre acknowledges that certain aspects of our facticity, for example an extreme pain, create an awareness of disvalue that is independent of my specific ends – pain is experienced as bad. However, the normative force of this disvalue will relate to my ultimate ends: the pain might be experienced as to-be-removed because my ultimate end is hedonistic, or to-be-endured because my ultimate end is motivated by religious asceticism for example.

Sartre organises the concept of situation into an integrated structure, consisting of: my place, my past, my environment, my neighbour and my death, and explains why for him, these things do not determine consciousness, and further, claims that they can be understood not as a threat to human freedom but rather as a condition for it.

My place is the geographical location in which I live, my country, with its climate and its resources. My environment consists of the instrumental things around me, together with their co-efficient of adversity and utility. I am surrounded by things that are different from me, and which develop potentials for or against me. I occupy my place in connection with the place I occupied before, I moved here from there, and that place refers back until the initial place in which I was born is reached. It is here that my experience began. The connections that exist before this place have no part in explaining my place for Sartre, because they do not belong to my experience, the chain is broken at this point: 'Thus to be born *is*, among other characteristics, *to take one's place*, or rather ... *to receive it*' (Sartre, 2005, p.512). Place is introduced through human reality. My being-there is a pure given, but it is only in relation to my free choice of what I want to do that my place appears to me as a help or hindrance, and that obstacles can be raised. If, in following my particular project to reach the top of a mountain by lunchtime, I encounter a rock fall blocking the only available path, this rock fall is part of my environment. It is manifested through my project as an obstacle. But I can attempt

to overcome my obstacles, or I can choose to abandon the project as an act of free renunciation. It is only by an act of free renunciation that things can be judged to be impossible: 'at whatever point I stop, it is I who have decided I couldn't go on any longer – hence, I could have gone on a bit longer still. But if I admit – and wish – never to have any excuse, my freedom becomes *mine*, I assume for ever that terrible responsibility' (Sartre, 1984, p.114).

In order for me to act freely, it is necessary that my freedom does not extend to an ability to alter the past. If my past actions could be immediately undone, they could not have had any meaningful effect, and I could not therefore be said to have been free to effect change: 'The very notion of freedom demands that our decision should plunge into the future, that something should have been *done* by it, that the subsequent instant should benefit from its predecessor and, though not necessitated, should be at least required by it' (Merleau-Ponty, 2004, p.213). But although my past is fixed, and I must make new decisions in the light of it, my past cannot determine my actions. The relation of my past consciousness to my present is one of interpretation. The meaning that my past has for me is variable, because it is dependent on my freely chosen project. A decision that I took last year might no longer seem reasonable, and although I cannot change the fact that I did make the decision, I am free to interpret it as ridiculous and take an alternative course of action. Or conversely, I might choose to continually remake that decision on the grounds that I interpret it as honourable. In either case, the past does not restrict my freedom, because having been objectified as the past, it has been put out of play for me and must be taken up or rejected anew: 'Thus like place, the past is integrated with the situation when the for-itself by its choice of the future confers on its past facticity a value, an hierarchical order, and an urgency in terms of which this facticity *motivates* the act and conduct of the for-itself' (Sartre, 2005, p.525).

The presence of the obstacle is then, on Sartre's account, not a limit to freedom but rather demanded by the very existence of freedom. He writes 'Not to *accept* what happens to you... To *assume* it ... in other words to adopt it as one's own, exactly *as if* one had given it oneself by decree, and, accepting that responsibility, to make it an opportunity for new advances, as if that were why one had given it oneself' (Sartre, 1984, p.95). My freedom requires that there is an environment that can be changed, and obstacles that can be surmounted. Freedom, to be freedom, must be exercised on the given which is independent of it: '...freedom's very project is in general to *do* in a resisting world by means of a victory over the world's resistances' (Sartre, 2005, p.528). Because things I encounter are independent of me, the unpredictable nature of my actions is already anticipated. My actions have a margin of indeterminacy built into them, and so my projects are open projects, they carry the possibility of modification. So, whilst it is

true that I cannot be said to have chosen that which I did not foresee, there is an extent to which, for Sartre, obstacles are foreseen.

Sartre allows that the factual existence of our fellow man constitutes a possible limitation to our freedom: in my being-for-others I 'meet with a meaning which is *mine* and which I have not given to myself' (Sartre, 2005, p.531). In recognition of my fellow man, I realise that objective meanings exist, that the world is already meaningful, and the meanings are ones that I have not attributed. The co-efficient of adversity of things encountered in the world is not only experienced by me, but can be revealed as existing independently of me by things such as signs, for example, 'danger', or 'exit'. Here Sartre acknowledges 'there is no longer a *situation* as the organization of a meaningful world around the free choice of my spontaneity; there is a *state* which is imposed upon me' (Sartre, 2005, p.533). It is a matter of fact that I belong to the human species, and as a member, I have a nationality, and speak a language that is comprehensible across members of my social group. I am thrown into a world, which has a specific meaning. However, the example of language can demonstrate that these limits are not obstacles to my freedom. Words only have meaning in the context of a sentence, and the sentence only has meaning in the intentional act of designating or understanding something. When something is designated or understood, the sentence and the laws of language are transcended, or nihilated towards a posited end – once the meaning has been conveyed, the words are no longer needed. Laws of language exist but these laws are the means of making freedom realisable, they are to-be-transcended. The concrete foundation of speech is then the speaker rather than the language. For Sartre, speech is an example of a social and universal technique, and he applies this reasoning equally to nationality; I exist as the nihilation of givens that are already part of a situation created by others. The brute givens are not obstacles, but the ground from which freedom arises.

In addition to meanings being conferred onto things that I encounter, my fellow man can confer meanings onto me. Because of the other, I have objectivised qualities – for example I am a slave or a master, clever or stupid – qualities which, although referring to me, are alienated from me. The other apprehends me as the Other-as-object, my situation becomes for him an objective structure. But although I may attempt to assume my being-for-others, it is unrealisable; the other is free to ascribe certain characteristics to me, but I *am not* these characteristics. The other might describe me as stupid, and I can understand the meaning of the description, applying it to myself, but 'I can not join the meaning of this word to my person' (Sartre, 2005, p.548). So, whilst there is a factual limit to our freedom in the existence of the other, I do not *encounter* it.

Another of the brute givens, and an aspect of our being-for-others, is death. Death is part of the fundamental facticity of the for-itself, it has no justification, and its character is, for Sartre, absurd. As long as I live, I can refuse the Other's interpretation of me by projecting myself towards ends which differ from his interpretation. But, after death, the meaning of my life continues to be modified, but not by me. The Other becomes the guardian of my life, and can choose to forget or remember, to interpret it as this or that kind of life: 'Thus the very existence of death alienates us wholly in our own life to the advantage of the Other. To be dead is to be a prey for the living' (Sartre, 2005, p.564). We know of death only through the existence of the other – without the other, my death would only be the simultaneous disappearance of myself and the world, of subject and object. Death is pre-reflectively in everything that we do, but is never encountered. Death marks the end of our possibilities and the permanence of our being-for-others, but as such, it cannot be conceived. In order to realise my death, I would have to look back at it from beyond it, which I cannot do. In *War Diaries*, Sartre declares that to present what is unrealisable as realised (Sartre, 1984, p.199), amounts to being in bad-faith. So death, like my being for others, is unrealisable, it is not *my* possibility, but external to me as the inevitable reverse side of my possibilities. But it is because death is beyond my subjectivity that my freedom remains unchallenged by it.

Sartre's position can then be described as that 'no past or present reality external to consciousness can impose conclusive, all-things-considered normative demands on the for-itself' (Poellner, 2015, p.233). The situation is the things we encounter, things are as they are, and I am there among them. But Sartre makes the often-quoted claim that 'the slave in chains is as free as his master', and in order to make sense of this, it must be emphasised that he does distinguish between different kinds of freedom. Sartre is clearly concerned with freedom of the will rather than physical, political or social freedom. The prisoner in chains can be understood as having freedom in the former sense, whilst being deprived of it in the latter. So, a distinction is made between absolute, or ontological freedom, and restricted, or practical freedom. Consciousness has absolute freedom because it is not its situation, it can separate itself from all that it is not, and so, *qua* consciousness – i.e. phenomenologically – it is not causally determined. Absolute freedom is not omnipotence of course. The prisoner is free to the extent that he can resign himself to his position, or he can rebel against it and make plans to escape, so he is ontologically free. He is however not free to walk out of prison and go on holiday, and so in this particular situation, he is *practically* unfree.

But, having made a distinction between ontological and practical freedom, Sartre shows that for him, they are necessarily connected. He writes 'If man is not originally free, but determined once and for all, we cannot even conceive what his liberation might be' (quoted in Detmer, 1988, p.66). The concept of imprisonment requires some capacity for ontological freedom in order to make sense. If the prisoner lacks the power of conscious choice, it makes no sense to speak of his conscious choice as being restricted or taken away, and so in this case, he could not be described as captive. Practical freedom, the freedom to pursue freely chosen ends, is not possible for a being that is wholly determined, because such a being could not freely choose. A fully determined being could not be described as unfree, the concept of freedom would have no relevance. Practical freedom is, for Sartre, grounded in ontological freedom.

What kind of value can be placed on this freedom that Sartre assigns to the prisoner? If the prisoner is only free to dream of his release, his freedom seems to be reduced to the level of mere ideology. Again, it must be emphasised that for Sartre, freedom is autonomy of choice, and choice is identical with acting. The prisoner is not free to go out of prison, but he is free to project his escape: 'The true philosophical notion of freedom distinguishes it both from a mere internal intention and from the ability to succeed in an enterprise' (Catalano, 1974, p.203). The prisoners' projection of escape is far stronger than a dream or whim, it is a genuine intention, and this is not diminished by the fact that it has no chance of success.

### Objections to Sartre's account of Initial choice

It is in view of the initial choice that Sartre claims that we create ourselves *ex nihilo*, and therefore secure freedom and responsibility. However, in Sartre's own account, and as Heidegger argues, we are always already 'thrown' into a world. We do not exist as pure intellect looking in on a world and deciding on the value of things, but are immersed in a world that has been variously interpreted by its inhabitants. In order to choose, we must have an understanding of how the world is, which appears to challenge Sartre's argument that it is through our initial choice that we gain understanding. Among other philosophers of the existential tradition, Merleau-Ponty rejects Sartre's concept of initial choice. Sartre, he claims, confuses choice and freedom – having the freedom to refuse what I am is not synonymous with making a choice of what I am; 'It must not be said that I continually choose myself on the pretext that I *could* continually refuse what I am. But not refusing is not a choice' (Merleau-Ponty, 2014, p.479). However, in the light of the account of the fundamental project given earlier, it seems that in making this objection, Merleau-Ponty has possibly misunderstood

Sartre's concept of Initial choice. Initial choice is, for Sartre, the foundation of all deliberation, and the interpretation of deliberation takes place in terms of the Initial choice. The Initial choice is 'one with the consciousness which we have of ourselves ... And as our being is precisely our original choice, the consciousness (of) the choice is identical with the self-consciousness which we have' (Sartre, 2005, p.484). Essentially, rationality is necessary but not sufficient as an explanation of human action, a point that is manifest in the concept of bad-faith. Bad faith is a way of relating to the world; it is a project in which we avoid responsibility for our actions by deceiving ourselves that we have a fixed nature over which we have no control. In *Existentialism is a Humanism*, Sartre categorises bad faith, in which freedom seeks to negate itself, as 'an error', a 'lie' (Sartre, 2007, p.47). In bad faith, the for-itself attempts to bring about a consciousness of its human subjectivity as reified in being-in-itself. A bad faith denial of freedom is not necessarily articulated, it is rather a *lived* self-deception, in which values and qualities of life are not acknowledged to be the result of personal choice, but treated as if they were independent of the subject: 'all bad faith manifests a *choice*, made not on the reflective, voluntary level but pre-reflectively as a "spontaneous determination of our being"' (Gardner, 2009, p.173). A life lived in bad faith is, for Sartre, a life of contradiction because in bad faith, a person chooses his own values (unreflectively), but at the same time, lives as if these values were not chosen by him. As an explanation of human action, freedom then transcends rationality: 'the choice is that by which all foundations and all reasons come into being' (Sartre, 2005, p.501). Empirical desires are motivated by reasons, but these symbolise the original choice, the fundamental desire, the full meaning of which is never fully transparent to me. Desire is, for Sartre, a lack of being, but in so far as this lack can never be overcome - consciousness can never become a substance - in desiring, I am pursuing an inconsistent end. But because I only have an unthematic grasp of what I am ultimately pursuing, I can choose inconsistent ends whilst not recognising them as inconsistent. It is in this sense that Sartre can claim that the initial choice is 'prior to logic ... It is this which decides the attitude of the person when confronted with logic and principles; therefore there can be no possibility of questioning it in conformance to logic' (Sartre, 2005, p.591). The fundamental project, the goal of becoming an in-itself-for-itself, is not grasped in the abstract, but only in concrete instances which symbolise completion, and this goal exists as something not yet achieved - its full meaning is never transparent; 'Absolute concreteness, completion, existence as a totality belong then to the free and fundamental desire which is the unique person. Empirical desire is only a symbolization of this' (Sartre, 2005, p.588).

Sartre clearly states that modification of our fundamental project is always possible (Sartre, 2005, p.486). However, if the fundamental project is a choice of what constitutes a reason for



the for-itself, what kind of reasons can there be for making a change? Thomas Baldwin claims that this transformation can never be a voluntary act: 'Since through our original choice we create the only values that are our values, we will never find any reason for altering our chosen project, despite the fact that we apprehend at some implicit level of anguish that we might have chosen some other project' (Baldwin, 1980, p.34). Sartre doesn't give a clear response to this question, but his language in *Being and Nothingness* suggests that the motivation required for this extreme modification goes beyond the rationality implied in Baldwin's claim. The first thing to note is that for Sartre we are 'perpetually engaged in our choice', and this suggests a vibrant relationship that is always open to assessment, it is not made and fixed in time. Secondly, Sartre connects the capacity for revision of the fundamental project with the idea of existential psychoanalysis, and the implication of this is that I do not have full knowledge of the ontological meanings of my own concrete projects. I have to decipher my own original choices and this may only be partially successful. But rather than implying the involvement of the unconscious, for Sartre this is explained by the distinction between reflective and pre-reflective consciousness. In addition, that such a dramatic modification incorporates more than rationality is clear from a passage where Sartre considers the example of an inferiority complex. He states that the suffering and shame this causes 'can be interpreted only in and through the intention which makes me inferior' (Sartre, 2005, p.497). But he goes on to write:

Yet at each moment I apprehend this initial choice as contingent and unjustifiable; at each moment therefore I am on the site suddenly to consider it objectively and consequently to surpass it and to make-it-past by causing the liberating *instant* to arise... hence ... the frequent upsurge of "conversions" which cause me totally to metamorphose my original project' (ibid).

What Sartre means by the 'instant' becomes clear in the following sentence:

These extraordinary and marvellous instants when the prior project collapses into the past in the light of a new project which rises on its ruins and which as yet exists only in outline ... these have often appeared to furnish the clearest and most moving image of our freedom' (ibid, p.497-8).

In Sartre's account there is a hierarchy of motivations, from those relating to concrete desires to those relating to the fundamental desire to become a necessary self-caused being. But the relationship between these desires involves 'the mediation of our situations and "symbolisation" of the original choice of self of each individual for-itself, in ways that call for existential psychoanalysis' (Gardner, 2009, p.179).

So, it appears that the objections made to Sartre's concept of Initial choice depend on 'choice' being understood in a more conventional, and perhaps more simplistic, sense than he intended. If understood specifically in the way Sartre used it, his account can, arguably, withstand these concerns.

### The question of determinism

For Sartre, our choices are not determined by our past, our situation, or by events that result from laws of nature; our actions are free, and we bear total responsibility for them. But the claim that we can transcend our past, our situation etc., is, for determinists, just false. For the determinist, Sartre's phenomenological claims do not reflect the way things really are. To address the determinists' challenge to Sartrean freedom, I will consider three potential ways of responding to their claims that our freedom is either limited or illusory. In the first, Robert Kane (2003) suggests, in agreement with Sartre, that the feeling we have that our decisions are free and undetermined *is* a good guide to the way things really are. Kane proposes an incompatibilist argument in which he argues that we *do* have free will and moral responsibility, that this does preclude determinism, and that our actions are not governed by luck or chance. The second type of response I would like to consider is presented in an account by Susan Wolf (Wolf, 1990): that the phenomenology may not reflect ontology, but this doesn't matter for our understanding of autonomy. Wolf offers an argument that is incompatibilist with regard to psychological determinism, but unlike Kane, Wolf argues that there is no reason to suppose that any potential physical determinism is a threat to our having responsibility for our actions. In an example of the third type of response, Gary Watson (2004) provides a compatibilist argument with an account of responsibility as attributability.

These arguments against the possible limitations on free will have strengths but also weaknesses. In the light of this, I will suggest a fourth response. Acknowledging the fact that the debate is still open, that there are no conclusive proofs of physicalism, it is legitimate to take practical considerations as primary – the *possibility* that the phenomenology does reflect ontology, that we are ontologically free, should be enough to motivate us to live as if we possessed absolute certainty.

## The determinist argument

I will begin by defining the argument which Kane, Wolf and Watson variously refute. In an essay on the impossibility of moral responsibility, Galen Strawson lays out a 'Basic Argument' for determinism that provides a useful example of what would need to be countered for a libertarian account of free will to succeed (Strawson, 2009, p.219). This argument leads Strawson to conclude that, *whether or not determinism is true*, we cannot be morally responsible for our actions. For him, any argument that suggests that we *can* be ultimately truly responsible, must explain where his argument fails. As we are concerned with affirming freedom, and so accepting the consequence of moral responsibility, responding to this argument will be critical. Strawson's conversational form of this argument follows:

1. You do what you do, in any situation in which you find yourself, because of the way you are.
2. To be truly morally responsible for what you do you must be truly responsible for the way you are – at least in certain mental respects.
3. You cannot be truly responsible for the way you are, so you cannot be truly responsible for what you do (ibid).

Strawson proceeds to develop his argument by making two claims: 1. From the basic argument, if we are determined, we cannot be held morally responsible for what we do, because our actions are determined by something that is outside our control. 2. If we are not determined, our actions are merely the result of luck or chance, and equally we cannot be morally responsible for them.

Strawson's argument appears credible on first sight, but can be countered by arguments that fall into any of the three general types: compatibilist accounts; incompatibilist accounts; or accounts that are compatibilist about certain aspects of determinism but not others. Examples of each of these types will be considered in taking on the determinist challenge.

## An incompatibilist response

Determinist theories often highlight the determinism evident in nature, and claim that this must be reflected in human behaviour since humans are part of the natural world. However, Kane (Kane, 2005, p.132-3) draws attention to advances in the realm of physical, biological and

human sciences, which have suggested that the natural world does include a degree of indeterminism. It has been further suggested that the activity of the brain might be susceptible to quantum indeterminacy in the exact timing of the firing of neurons. Whilst acknowledging that these suggestions are speculative, and also that chaotic behaviour alone is not enough to imply genuine indeterminism in nature, Kane puts forward an argument for the possibility of a combination of quantum physics and new speculation on chaos and complexity, which would make sense of free will:

If the processing of the brain does “make chaos in order to make sense of the world” ... then the resulting chaos might magnify quantum indeterminacies in the firings of individual neurons. These chaotically magnified indeterminacies in the firings of neurons would have large-scale indeterministic effects on the activity of neural networks in the brain as a whole. The indeterminacy at the neuron level would no longer be “damped out,” it would have significant effects on cognitive processing and deliberation (Kane, 2005, p.134).

Kane applies this reasoning to specific deliberative processes - those acts about which, according to him, we make claims for ultimate responsibility, actions from which we form our will by making significant choices that affect our character. Kane calls these ‘self-forming actions’ (SFA’s). SFA’s occur in situations when we have conflicting desires and motivations. In these situations we experience tension and uncertainty, and, Kane argues, this might perhaps be reflected in a kind of stirring up of chaos in the brain – the indeterminacy of our neural processes: ‘What we experience internally as uncertainty about what to do on such occasions would correspond physically to the opening of a window of opportunity that temporarily screens off complete determination by influences of the past’ (Kane, 2005, p.135). In such a situation where we might, for example, be trying to decide between two conflicting actions, one an altruistic act and one based on personal ambition, there are strong reasons in favour of either choice. What hinders a decision in this example is the indeterminism produced internally in the subject by the conflicting desires, and it is this indeterminism that blocks the possibility of determination by the past. The subject simultaneously desires to do two mutually exclusive things at once, but at the moment of choosing, either choice will be endorsed as that which the subject most wanted to do for the reasons available to him at that moment. Whatever is chosen can then be described afterwards as having been the result of a free and rational process. Indeterminism does not feature in all acts, we do at times act from a will that has already been formed. But Kane argues that when we do so, we are acting as a result of earlier undetermined actions, or SFA’s.

We might think though that on Kane's account, actions are still subject to the Luck Principle: If an action is *undetermined* at a time *t*, then its happening rather than not happening at *t* would be a matter of *chance* or *luck*, and so could not be a *free* and *responsible* action. Kane rejects this claim however, arguing that indeterminism rules out deterministic causation, not causation *per se*. In the decision-making process, again using the example of the choice between an altruistic act and one based on personal ambition, *the cause is the effort of will* involved in reaching a decision, and it has the property of being indeterminate. The process of reaching a decision *is* the effort of will, and this continues right up to the point at which the choice is made. The choice of one course of action over another is the result of the effort: 'There is no point at which the effort stops and chance "takes over"' (Kane, 2005, p.140). Whichever choice is made, it will be the result of the success of the effort, because the choice was between two things that were considered desirable.

Kane acknowledges that despite the decision being rational, voluntary and undetermined, there is a degree of arbitrariness because there cannot be sufficient *prior* reasons for choosing one act over the other. SFA's involve an element of risk taking. This, however, is consistent with the nature of free will – the justification for a choice lies in the future rather than being explained by the past – each SFA is 'the initiation of what might be called a *value experiment* whose justification lies in the future and is not fully explained by past reasons' (Kane, 2005, p.144). To the extent that our actions can be understood as 'self forming', we can respond to (3) in the Basic Argument; we are responsible for the way we are, at least in some mental respects. If this is so, we do have power to make choices for which we are morally responsible.

Kane's argument involves some speculative claims regarding the working of the neural system in the brain. But whilst these suggestions do not prove that indeterminism *is* a feature of mental activity, or that we do have free will, if they are scientifically *possible*, they might be used to defend the libertarian position against charges of incoherence, obscurity and mystery. It would then be up to the determinist to prove that such claims are in fact not true.

Both the argument from Kane, and Sartre's ontological view of consciousness and freedom, are incompatibilist. In Sartre's ontology, consciousness cannot be determined by natural processes, because it is of a different order from any other kind of entity – consciousness is not a being. Consciousness is transparent to itself, and if anything were to physically interact with it, that force would become known to consciousness, consciousness would be aware of it. Phenomenologically this is not the case, we do not have an awareness of being determined by anything. But, the determinist might still argue that phenomenology is not reflective of

ontology, and that we *are* in fact determined by something of which we are not aware. Compatibilist arguments for freedom and responsibility assert that even if we are so determined, it makes no practical difference, and in some sections of his writing Sartre appears to be sympathetic to this view.

### The reason view, a partial incompatibilist response

In Wolf's account, to have responsibility, we must act in accordance with reason. Arguments that focus on autonomy hold as primary the ability to act in discordance with reason if we choose to do so, stated as 'the ability to do otherwise': we must not be necessitated to take one course of action rather than another. But, for Wolf, acting in discordance with reason is inconsistent with our status as morally responsible beings. Even if we have total autonomy in the sense that we are the ultimate source of our own actions, if we do not act in accordance with reason, we cannot be held morally responsible for what we do. Counter to the 'ability to do otherwise' claim, the agent must be able to do only one thing; to act in accordance with reason. If the agent does act in accordance with reason, he can clearly take responsibility for that act. If, however, he does not so act, whether or not he is responsible is a matter of whether he *could* have done. For Wolf, the ability to do only one thing does not imply a view of life as pre-ordained or inflexible, because acting in accordance with reason necessarily involves being responsive to different conditions in a changing environment. If, for example, an agent is unable to do otherwise because one course of action appears to be too rational to ignore, this does not imply a limit to that agents' freedom, but rather appears to confirm it.

Wolf's 'reason view' of freedom shifts the focus away from the determinist's concern about what metaphysical kind of persons we must be in order to be free and responsible; whether or not we are subject to the same causal networks, the same psychological and physical forces, as the rest of nature. The question is rather located in the intellectual and the meta-ethical sphere, whether we have the ability to think and so to act well, rather than badly. Wolf further defines this as 'the ability to act in accordance with, and on the basis of, the True and the Good' (Wolf, 1993, p.71). The ability to act on the basis of what is true and good clearly does not allow for being fully, exhaustively determined by one's past, rather it requires the flexibility to assess each situation rationally and decide how to act accordingly. Neither does it entail that an agent necessarily *will* always act in accordance with these values, but merely that he has the ability to do so. What is relevant then is not how many options are available, but the quality of the options, and it is the agent's own faculty of reasoning that will lead to the assessment.

For Wolf, the True and the Good are not arbitrary values attached to the agent's personality and rational capacity, however. The objective nature of these values might suggest that metaphysical concepts or Platonic values are being introduced, but this is not Wolf's understanding. That beliefs are shaped by the external world is not considered a threat to freedom and responsibility, but rather a promotion of it. Empirical evidence, which is external to the agent, is necessary in the formation of belief. An agent is less free, not more, if ignorant of external facts relevant to the situation. Wolf's claim that 'the free and responsible agent must be able to see and appreciate the True' (Wolf, 1993, p.120) appeals to the process of rational belief formation, the ability to see how things in the world actually are. Further to this, the objectivity of values suggested by reference to the True and the Good, relates to non-arbitrary standards by which an agent judges some actions, choices etc. to be better than others and to the acknowledgement that there *are* better or worse choices to be made. These values are objective, but only to the extent implied in ordinary moral discourse – an example of which is the statement 'the gratuitous torture of a child is wrong'.

In proposing this concept of freedom, Wolf claims that the 'reason view' of freedom is preferable to one of the prevalent alternative views – the 'real self' view. In the 'real self' view, an agent has responsibility for actions that can be attributed to his real self, where the agent is able to govern his actions on the basis of his system of values. For Wolf, the 'real self' view does not account for the question of the agents' responsibility *for* that real self. The desires and values of a person may be attributed, for example, to insanity, or early traumatic experience. In addition, for Wolf, the responsibility conferred by the 'real self' view is shallow, it merely attributes a pivotal role to the agent in the causal series of events leading to the act in question, and then judges whether the act is deserving of credit or discredit. Wolf wants to argue for a deeper sense of responsibility in which the non-instrumental moral quality of the agent himself is in question, and considers that the 'real self' view is inadequate in this regard. However Gary Watson (2004) defends a particular version of the 'real self' view – the 'self disclosure' view – against these charges.

### A compatibilist response

Watson distinguishes between two senses of responsibility: accountability and attributability. If we want to hold someone *accountable* for his actions, we need to ascertain the level of control an agent has, which is central to the question of determinism - if someone does not have

control over the ends they have, we cannot reasonably hold them morally responsible for their acts. In assessing accountability, something in addition to moral judgement is implied - when questions of praise and blame are introduced, there is a suggestion that a further response is appropriate, censure or sanction for example. But for Watson, *attributability* is primary: attributability focuses on the attribution of conduct to a person as its agent, responsibility is an aspect of the unity of character and conduct. The appraisal that results is 'appraisal of the individual as an adopter of ends' (Watson, 2004, p.263). The agent is responsible for an action because it is *his own*, he is the author of it, and as such he is self-governing, and his ends are necessarily self-disclosing. Attributability relates to ethical issues of practical identity, such as what it is to lead a life, and the quality or characteristics of that life, and it is these things that the agent takes responsibility for. In Watson's 'self disclosure' view of freedom, the evaluation is focussed on the ends adopted by the agent: 'If what I do flows from my values and ends, there is a stronger sense in which my activities are inescapably my own: I am committed to them. As declarations of my adopted ends, they express what I'm about, my identity as an agent' (ibid, p.270). Watson's argument, like Kane's, rejects (3) of Strawson's Basic Argument, claiming that you *can* be truly responsible for the way you are.

Like Sartre, Watson appears to be in agreement with some form of compatibilism. His phenomenological account of freedom concerns what it feels like to be an intentional agent. If I have no conscious experience of being compelled to act by deterministic causal laws, then I can be said to own my actions and to have acted autonomously. This accords with Sartre's view that we are responsible for what we freely choose to do: Sartre makes no distinction between an act, and a free act; acts are necessarily free. He counters the suggestion that we are psychologically determined, that we are just ignorant of the fact that we are actually determined to act, with the claim that even if this were the case, it would make no practical difference (Sartre, 2005, p.57). Despite this apparent concession, Sartre's ontological account of freedom, based on the self-determination of consciousness, would firmly reject a deterministic position, and must be considered as incompatibilist indeterminism.

What these counter arguments show collectively, is that we are not rationally compelled to accept determinism as an inevitable limit to our freedom, and that Strawson's basic argument is not irrefutable. Our pre-theoretical understanding of our decision-making process tells us that when our actions are intelligible to us, we choose freely and we are responsible for what we choose. This has practical connotations, as suggested by Gardner who claims that Sartre's concept of pure reflection 'refers to a non-discursive, intuitive discovery of one's *own* ontological freedom, a privileged reflexive cognition of the sort that entails a change in one's



practical orientation' (Gardner, 2009, p.171). The various arguments outlined here allow us to leave our intuitive understanding intact, whilst acknowledging that the metaphysical issue of determinism is, so far, still open.

I would like to conclude this chapter however by arguing for a stronger claim: building on the accounts of freedom offered by Kant and Sartre, my proposal is that the possibility of our being free involves an obligation to live as if this were the case.

### The obligation to live as if we were free

In *Groundwork of the Metaphysics of Morals*, Kant states that when you choose, you must do so 'under the idea of freedom' (Kant, 2012, 4:448). If a subject chooses to do something purely to satisfy a desire, he regards his act as involuntary to the extent that he considers the desire to have impelled the act. But, if he considers instead that he chose to act on the desire, the act is then regarded as a product of his will, and so as a free act. Kant does not intend that you should believe that you are free, but that you should choose as if you were. The belief in free will is theoretical, but it has a practical function. Our decisions should arise from principles that we have freely chosen, and these principles should justify what we do.

Following Kant, in the essay *The Authority of Reflection* (Korsgaard, 1996, p.90-130), Christine Korsgaard offers an account of freedom that is not threatened by determinism, in which autonomy is understood as ground for our obligations both to ourselves, and to humanity in general. Like Kant, Korsgaard argues that the question of whether or not we are determined is purely theoretical, and has no practical relevance to our status as autonomous. As beings with the capacity for self-reflection, humans are uniquely capable of calling into question their mental activity. In self-reflection, we can ask ourselves whether or not our beliefs, desires, perceptions etc. justify what action we might take. Having raised this question, we need a reason to act on our findings. If our impulses withstand the scrutiny of reflection, we have a reason to continue, if not, we have a reason to change. Korsgaard claims that it is through freedom that we make our acts our own: by scrutinising an external influence, desire for example, by way of self-reflection; we have freedom to endorse or deny it, and so our acts are freely chosen. The suggestion could still be made that, for example, whether we ultimately do endorse a desire is pre-determined. But Korsgaard denies that this threatens our freedom. In any situation the agent is faced with having to make a decision. Even if what he decides could actually be predicted by an outside observer, this makes absolutely no difference to him, he

must still decide for himself what he will do: 'It is from within the deliberative perspective that we see our desires as providing suggestions which we may take or leave' (Korsgaard, 1996, p.96).

Korsgaard is clearly right to say that the fact of being determined would make no difference to the practical side of our decision-making process. We have to go out into the world and decide how we will act from day to day. Whatever the ultimate reality, the decisions we make feel like our own free decisions. However, it is not the case that the question of whether or not we are determined is purely theoretical and has no practical relevance. As humans, we tend to take an attitude, or stance to our lives, seen as a whole. For example, some people are, as a general rule, more optimistic or more positive in their outlook than others. Some are more likely to take risks, to live dangerously, whilst others prefer to be cautious, and these stances, while not necessarily the result of conscious decisions, potentially colour everything we do. Whether or not we ultimately take ourselves to be determined, even if this thought is not articulated, or even recognised by us, will affect our general stance. To live with a conviction of absolute freedom, is to accept total responsibility for our actions. If we have absolute responsibility, we cannot, for example, justifiably say 'it's out of my hands', 'there's no more I can do' in the face of a continuing injustice. There will be no point at which we can abdicate responsibility, and say it is now up to someone else.

An example can illustrate this point: suppose a situation where X has been imprisoned for a crime that Y is convinced X did not commit. Suppose also that Y is a devout believer in a God who hears and responds to the prayers of his people. Y prays that justice will be served. Y might also campaign for X's release; he might petition the relevant politicians. But ultimately whether or not X is released, Y will consider that God has answered his prayer. He will reach the point where he says 'I have done what I can, it is in God's hands', and can move on without experiencing guilt. Contrast this with a real-life situation: Sir Michael Dummett was a philosopher whose chief philosophical interests were logic, language and mathematics. He had a major influence on analytic philosophy. He was also deeply concerned with ethical and political issues concerning refugees and immigration, and was a leading campaigner for racial tolerance and equality. As a highly educated and respected man, he was in a strong position to write about these issues with the justified expectation of being heard. In 1958, Dummett and his wife co-founded the Institute of Race Relations think-tank. It would be reasonable to consider that on the issue of race relations, Dummett had made full use of his position, and that he might reasonably consider that he had done all that he could do, leaving any necessary further work to the politicians. However, in the face of a huge rise in immigration during the

1960's, Dummett and his wife 'drove a battered van to Heathrow Airport day after day to take up the cases of Asian and West Indian immigrants threatened with deportation. On one occasion they were arrested and prosecuted after staging a protest against a market stallholder who refused to serve black customers. Police dropped charges ...' (Telegraph, 2011). The point is not that Dummett was a good man, although he might well have been. It is that by his actions he demonstrated that there was no stage at which he felt he could consider moral responsibility to be transferable to someone or something else.

The deep-seated suspicion that we cannot escape aspects of our situation, is conducive to the development of a passive attitude of resignation, and the feeling that ultimate responsibility lies somewhere beyond our reach. Similarly, the belief that all is ultimately under the control of a benevolent creator God, that all will work out for our good in the end, may result in an overall attitude of submissiveness. It is necessary to accept ultimate responsibility if we are to be motivated to take a fully active stance to life in general. Even if we are unclear as to how much freedom we actually have, perhaps judging that we are passive in some respects, and active in others, each decision that we take will contribute to our overall outlook, and be influenced by our underlying convictions. Ultimately, these two stances on life taken as a whole: passive resignation and active obligation, are in profound tension.

The question of determinism does then appear to have practical relevance. Rather than attempting to find a position that allows for the co-existence of determinism and indeterminism, in the light of the fact that we will never know, with certainty, which is the case, it is more beneficial to consider what is at stake when we favour one or other position. In microeconomic theory, the 'opportunity cost' is the value of an opportunity forgone in a situation of mutually exclusive alternatives. The opportunity cost of taking the passive stance, of allowing situations to continue on the grounds of putative inevitability, is considerable.

If, as Sartre argues, we are 'condemned to be free', then we do have full and total responsibility. And this clearly involves demands on us. If we are free, we must acknowledge our potential to effect change. The capacity for self-reflection, together with a belief in the power to try and change the things about us and our world that appear wrong, accepting full responsibility for our actions, is arguably what marks us out as truly human.

The arguments outlined above suggest that there is a clear *possibility* that we are not determined, that we *are* free and that we *can* create ourselves. The *possibility* of total freedom ought to motivate us to live under the assumption of that freedom, taking responsibility for

how things are with us, and in the world. But this real possibility does not only constitute a sufficient positive reason to act in this way. By considering the opportunity cost, we introduce the thought that if we fail to do so, we are at risk of serious moral error. By living in bad faith, as if we were not free, we risk denying our fundamental humanity.

### CHAPTER THREE: AN ETHICS OF FREEDOM

At the end of the last chapter I suggested that we had a 'responsibility' to act as if we were certain of our absolute freedom, and that by failing to do so, we run the risk of serious moral error. But the concept of responsibility introduces important questions: can an existence that aims to uphold its freedom absolutely, be morally responsible, rather than individualistic and amoral?; why should acting in the certainty of autonomy lead to morally commendable actions?; if everyone is free to choose, and to establish the value of what he chooses, how can any choice be considered wrong, inappropriate, or more critically, unethical?; and importantly, if we value our freedom absolutely can we hope to answer Schopenhauer's question, paraphrased by Joseph Campbell: 'How is it possible that suffering that is neither my own nor of my concern should immediately affect me as though it were my own, and with such force that it moves me to action?' (Campbell, 2013, p.xi).

There are views that consider autonomy to be in conflict with any familiar form of ethics, for example those of Nietzsche, and perhaps some readings of Sartre. However, there is an opposing view that autonomy is only properly realised in ethical commitment. This view is particularly associated with the Kantian tradition, and can also be found in alternative readings of Sartre. It is this opposing view that I would like to consider, beginning with an account of Kant's reasons for valuing freedom absolutely. I suggest that although, as Kant claims, we must value our freedom to determine our ends without necessitation, something central to the ethical domain, namely the emotional faculty of empathy, is left out in Kant's elevation of rationality. So I will then consider Sartre's account, which also values freedom absolutely but for very different reasons. For Sartre, ethical commitment is motivated by the necessity to value not only your own freedom but that of others: in taking up our freedom we become aware of an ultimate goal, the desire for completeness, or experiential absence of lack, and this desire extends to the whole of humanity.

Although both Kant and Sartre uphold the value of freedom absolutely, in the last section I will argue that we need to broaden the scope beyond these two accounts if we are to achieve a fully ethical account of human autonomy.

## Kant's conception of freedom as rationality.

For Kant, if you are rational, you must value freedom. In *Lectures on Ethics*, Kant claims that 'freedom is the source of all value – that it is intrinsically valuable, and that other valuable things must not merely be compatible with freedom but actually derive their value from the value of freedom' (Guyer, 2000, p.129).

Freedom, for Kant, refers to the ability of a person to determine his ends without necessitation either from his own desires and inclinations, or from external sources. Freedom then is, on this account, freedom of the will. The will does not conform to the laws of nature, but neither is it lawless, freedom of the will is autonomy: the power or right of self-government. Kant's autonomous agent is one whose will 'is not just subject to the law, but subject in such a way that it must also be viewed *as self-legislating*, and just on account of this as subject to the law (of which it can consider itself the author) in the first place' (Kant, 2012, 4:431). Self-legislation occurs in the light of reasons that, for the constructivist interpreter of Kant, are constituted by the activity of human reason, and for the realist interpreter, we antecedently recognise. All sentient beings are able to use the powers that they have in accordance with their will, but in animals, this will is bound to the impulses of the senses, and in being so determined, non-human animals are not free in a Kantian sense. Uniquely, humans are not necessitated to action by mere inclination. Humans have the capacity for practical rationality – the faculty of giving and receiving laws and principles – in accordance with which they act. So Kant maintains a clear divide between the rationality of the will, and the impulses of the senses. In order for it to be autonomous, it must be the case that 'the will is in all actions a law to itself' (Kant, 2012, 4:447). Kant prioritises the formal aspect of ethical judgements, which he equates with rationality, over the material aspect of ethical acts, which he equates with sensory content. This is in order to preserve ethical demands as having universal application, undetermined by contingent events. Actions motivated by sensory content relate purely to pleasure and pain for Kant, and are therefore contingent. One reason for freedom to be considered valuable for Kant is that it is 'ultimately by our freedom that we distinguish ourselves from the rest of nature, which merely seeks contentment or happiness' (Guyer, 2000, p.169-70). So, freedom is synonymous with rationality. But why does Kant consider freedom to be *intrinsically and unconditionally* valuable?

As rational agents, we generally act in accordance with subjective projected ends to satisfy our needs and desires. But Kant's aim was to identify an objective, universal end, common to all rational agents *because* they are rational – subjective empirical ends have only conditional

value but Kant argued that unless these conditional values are grounded in some unconditional value, 'nothing whatsoever of *absolute worth* could be found' (Kant, 2012, 4:428). This source of value, for Kant, is not found in some transcendent object such as God, or an empirical object, for instance our empirical selves, but in practical reason. How this is understood will depend on whether a realist or a constructivist interpretation of Kant is adopted. The central issue in this interpretative contention, loosely outlined, is whether values are possessed by things independently of us – a realist account, or whether things are valuable because we value them – a constructivist account. With regard to ethical judgements, are they correct (or not) as a matter of fact, independent of the choices of a rational agent? Or do they depend on such choices?

For the constructivist, the need to preserve our autonomy led Kant to seek an alternative to an understanding of moral values as objectively real. And his alternative is a conception of rationality as possessor of ultimate value and ground of all values, with the consequence that values are subjective. By making rationality the ground of all values, an infinite regress in conditional values, which would otherwise occur, is avoided; the question as to why something is valuable generally involves an answer that refers to the value of something else, for the sake of which the initial thing is valued. By making rational agency the capacity to confer values on actions and their objects, the cycle of questions of the 'but why is *that* valuable?' nature can be ended. The source of value is rationality because the rational will has the status of conferring value; what is valuable *is* just what is required by rationality. However, why should we accept that rationality can confer *ultimate* value apart from as a way of avoiding the infinite regress? Again, different interpretations of Kant offer different answers to this question, but taking Korsgaard's constitutive reading as one example, the idea would be loosely as follows: if I am an agent, I must regard an action as *my* action rather than the product of some force acting on me (Korsgaard, 2009, p.18). And for an action to qualify as such, it must be attributed to me as a unified whole rather than just to one part of me: mere behaviour *can* be attributed to one part of me, for example if I tap my foot without thinking about it, this behaviour can be attributed purely to a bodily movement, and this does not qualify as an action; by contrast, if I decide to travel the world, my decision counts as an action which can be attributable to me as a unified whole: 'For a movement to be my action, for it to be expressive of *myself* in the way that an action must be, it must result from my entire nature working as an integrated whole' (ibid, p.19). Following Kant, the parts of the agent that Korsgaard claims are in need of unification are reflective thought and desire. To pull these parts together we must make a choice, and in order to do that, we need a reason derived from a principle. If the agent is to have reason to accept this principle, it cannot be imposed on him from the outside, but must be the agent's own, it

must be a subjective principle. The relationship between the rational will and desire is determined by the principle on which the agent acts, and it is only by acting in accordance with the principle of the categorical imperative, an imperative which has universal as opposed to hypothetical mandate, that the two parts can be unified. A material principle aims at the attainment of a subjective end, it is contingent and relative, and so can only provide a hypothetical imperative – I must do X if I am to attain Y. A formal principle, by contrast, aims at objective ends which are such for all rational beings – if I am to live up to my status as a rational being, I must obey such a principle irrespective of my desires or feelings. The deliberative action organises desire and reason into a unified system, and this generates an action attributable to the agent. What rationality requires, our rational principles, are valuable because they are constitutive of what it is to be an agent, which for Kant is synonymous with being a person.

On the constructivist reading, freedom is intrinsically and unconditionally valuable for Kant because it is synonymous with rationality, and rationality is ultimately what confers personhood. The consequence of this is that values are wholly dependent on the attitudes of those who value, there are no antecedent facts about what is valuable apart from those conferred by a particular point of view. What is considered good is that which is chosen by a creature exercising rationality, nothing is ontologically good. If we confer value merely by the act of choosing, this would suggest that we can declare anything to be good, just because we regard it so. However Korsgaard denies this, claiming that as agents, we have value-conferring status by virtue of having the power of rational choice, and therefore all humans with this power also have value conferring status. The implication of this is that whatever one agent calls 'good' must be in harmony with whatever all the other rational agents claim; 'for the good is a consistent, harmonious object shared by all rational beings' (Korsgaard, 2000, p.123). Because rational nature is an end in itself, it must be treated as such in all instances – in our own person as well as that of another. To act otherwise is contradictory: 'If you overturn the *source* of the goodness of your end, neither your end nor the action which aims at it can possibly be good, and your action will not be fully rational' (ibid).

Alternatively, for the realist, our worth as rational beings is not constructed but is objectively valid. The categorical imperative is grounded in the absolute value possessed by all rational beings who exist as objective ends in themselves; 'The content of the law is not a creation of my will, or the outcome of any constructive procedure on my part. The law of autonomy is objectively valid for rational volition because it is based on an objective end – the dignity of rational nature as an end in itself' (Wood, 2008, p.108). Our commitment to the categorical imperative must be explained by an underlying value which we take it to have; 'A good will is



good not because of what it effects, or accomplishes, not because of its fitness to attain some intended end, but good just by its willing, i.e. in itself' (Kant, 2012, 4:394). To act in accordance with the noumenal level of our moral selves is to act at a higher level than that of appearance in which our desiring selves are located; 'In this way, Kant suggests, he can explain why even 'the most hardened scoundrel' will admit that he would be a 'better person' (i.e. a person of greater worth) by bringing himself to act morally, where he would then be acting in the 'higher' world of the understanding and not the 'lower' world of sense' (Stern, 2012, p.31). On the realist reading, a prior order of values provides the measure against which our actions, guided by practical reason, are judged. The source of the value of rational agency is then independent of any act of willing or valuing by us. Objective ends, which alone can bind unconditionally, must ground principles that apply to and are necessary for all rational beings, and for the Kantian realist, these objective ends are persons. Whilst the realist does not have to seek agreement over what to *determine* valuable, he is obliged to treat all rational agents as equal co-authorities over what *counts* as a reason; 'Persons thus have a normative status in virtue of their rational capacities' (Formosa, 2011, p.183-4). What counts as a good act is limited to what there is a reason to do, but we do not determine this independently of other rational agents. Our choice of what counts as a reason are influenced by the requirement for agreement by other rational agents. Rationality is unconditionally valuable for the realist because it is in virtue of possessing it that persons are regarded as having absolute value as ends in themselves.

The constructivist/realist debate is on-going and extensive, involving many varying interpretations of Kant. There are well-rehearsed arguments in favour of, and against, both realist and constructivist positions. But the aspect I am interested in here is shared by both: Kant is commonly held as being committed to the absolute value of rational agency. So what are the implications of valuing rational nature absolutely?

Kant holds that a universal objective *principle* is required to ground moral action. Emotions, for Kant, fall outside the scope of personal agency, they cannot be produced by the power of the will, and therefore there can be no obligation to feel an emotion. As Kant links moral agency with following rational obligations, emotions must be excluded from the moral domain. Feelings and desires, being idiosyncratic rather than valid for all rational beings, cannot give rise to a practical law. A rational being's maxims – subjective principles of intention – are practical laws not in virtue of their content, but of their form; the moral law becomes the mere form of a universal legislation. For Kant, the categorical imperative - acting in accordance with a maxim is permissible only if that maxim could be universally accepted as a principle of action (Kant, 2012, 4:447) – entails that a free will is the same as a will under moral law.

For Kant, if you act in accordance with morally right thinking, you are beyond reproach. Reason is the power guiding moral conduct and facilitating respect for the dignity of the other '*never merely as a means, but always at the same time as an end in itself*' (Kant, 2012, 4:433). It is by virtue of having the gift of reason that we are commanded by the moral law, laid down in us as rational beings, to respect the other. The input of emotions and feelings must be subordinate to rationality in order to preserve our autonomy, and as Iris Murdoch writes; 'Act, choice, decision, responsibility, independence are emphasized in this philosophy of puritanical origin and apparent austerity' (Murdoch, 1970, p.81). Kant's vision of the ideal, is a vision of human rational perfection and moral virtue, which ultimately downgrades empirical characteristics. Richard Dean writes; 'There is, then, a sustained line of thinking in Kant's texts that the power of reason produces an idea of humanity, and that this idea is meant to play an action-guiding role in practical deliberation' (Dean, 2013. p.174). But is it reasonable to expect rational ideas alone to perform an action-guiding role in practical deliberation?

### Acting within the moral domain

In the Universal Law formulation of the Categorical Imperative, Kant claims to provide a procedure for guiding deliberation both on how to act, and on the deontic status of actions (Kant, 2013, 6:225). An action will be morally right if its associated maxim can be willed as a universal law. This then is supposed to show us how to choose what is right and to act accordingly. However, there are many different maxims that might be associated with an action, and these will vary according to how the relevant circumstances are conceptualised. The maxim which we associate with an action will depend on how that action and the circumstances surrounding it are described. So, the moral status of any action will be determined by its description, and about this, as Anscombe claims, Kant gives no guidance; 'His rule about universalizable maxims is useless without stipulations as to what shall count as a relevant description of an action with a view to constructing a maxim about it' (Anscombe, 1958, p.2).

O'Neill responds to this charge arguing that an answer to the problem of relevant descriptions is built into Kant's formula: in Kant's account a maxim is not just any principle that an act exemplifies, but is the principle of a particular rational agent at a particular time – a subjective principle; it 'expresses a determination of the power of choice. To say that an agent's power of choice is determined is simply to say that he intends to do a specific sort of act or pursue a

specific end in some situation' (O'Neill, 1975, p.40). A maxim is then understood as an intention, which may be formulated by the agent either on the occasion of acting, or beforehand. There may also be future based general intentions to pursue certain sorts of ends. For O'Neill, Kant addresses the problem of relevant descriptions in his claim that any voluntary act has an intention, and it is this to which the categorical imperative should be applied. Applying the principle to the intention focusses on the first-person perspective of the agent, so offering only a subjective representation of the situation. But whilst remaining silent on the objective rightness of an action, it does offer guidance on how to act. As O'Neill admits however, any act has many parts, and these correspond to many intentions. How do we decide which intention is primary? This response does not seem to answer Anscombe's point, we still have no way of knowing that any particular act intention is appropriately described. There have been many other attempts to respond to these challenges. For example, Schumski argues for an interpretation of Kant's universality requirement which switches the focus from the particular case, to a call for agreement among rational agents: 'a principle satisfies the universality requirement if and only if all rational agents can agree that it would hold for them if they were to find themselves in the circumstances in question' (Schumski, 2017, p.19). But a fundamental problem facing Kant's ethical system still stands: approaching a moral situation either by choosing an appropriate maxim from an already formulated stock, or creating a new one in accordance with the circumstances of action, or indeed considering whether the action would be appropriate from the perspective of any and all potential rational agents if willed as a universal law, all involve *rational* assessment of the morally relevant features. Each situation demanding a moral response is different in subtle ways from all other situations. And whilst offering little guidance on how to determine moral relevance, Kant assumes that this is a task that must be accomplished purely by rationality. This is in spite of the fact that situations are presented to us not as theoretical concepts, but as involving people possessing both intellectual and emotional faculties, who have physical feelings and sensations, and are sensitive to their specific environment.

When we think of what it means to live ethically, I propose that the ability to act in accordance with rational principles is not enough. And further to this, that being a unified agent cannot simply be a matter of the alignment of the will and desire. Contrary to Kant's claim, in a moral situation, we must primarily be able to 'see' the situation as being in the moral domain, and to engage with the moral implications. Unless we can really perceive the moral implications in a wider sense, we will be unable to rationalise about how to act morally. In a discussion of the primacy of the act of perception, Vetlesen writes; 'moral perception is not conceived as something to be naively taken for granted, to be presupposed as a mere given, but is conceived

as an accomplishment, a complex feat, and an utterly precarious one at that' (Vetlesen, 1994, p.6-7). To perceive the situation as being in the domain of morality, involves being able to visualise and engage with not-yet-realised possibilities, and to see that things could be better, a uniquely human capacity. Vetlesen suggests that moral performance runs through a three-level sequence that is necessary in order for a subject to recognise the other as a moral addressee: perception, judgement and action. For action not to be devoid of intentionality requires judgement, and for judgement not to be blind or uninformed, requires perception. The order of the sequence is fixed – before we can get to the point of judgement, we need perception that 'gives' judgement its object, and, for Vetlesen, we 'perceive' not through rationality but through the basic emotional faculty of empathy:

To "see" the circumstance and to see oneself as addressed by it, and thus to be susceptible to the way a situation affects the weal and woe of others, in short, to identify a situation as carrying *moral significance* in the first place – all of this is required in order to enter the domain of the moral, and none of it would come about without the basic emotional faculty of empathy (Vetlesen, 1994, p.4).

Moral perception sets the scene in which moral judgement – which provides the norms and principles required for moral action – can take place. If I do not first 'see' the situation as morally relevant, calling for judgemental powers, the cognitive knowledge is impotent. Sartre refers to this requirement as intuition, the direct 'seeing' or apprehension of a point or a principle – we can sometimes know something to be true by direct, unmediated inspection of it rather than by use of inferential reason: 'intuition is the presence of consciousness to the thing' (Sartre, 2005, p.196). By prioritising the rational absolutely, and regarding moral feeling merely as the effect the law has, Kant, in a quest to preserve the absolute value of freedom, disrupts the sequence that allows entry to the moral domain. Before considering how this sequence might be reinstated, I will consider Sartre's account, in which the human subject, by virtue of its humanity, has a common goal: it seeks to become its own foundation and to achieve coincidence with itself, or completion. And, on some readings of Sartre, valuing not only your own freedom but that of others, motivates ethical commitment.

### Sartre, and the absolute value of freedom.

Sartre claims that valuing freedom absolutely, as the highest value, is a pre-requisite of an authentic life. Freedom is both the source of all value, and is itself a value because it is the constitutive condition of possibility of anything showing up as a value at all. In reference to

Sartre's ontology, Simone de Beauvoir writes: 'The man who seeks to justify his life must want freedom itself absolutely and above everything else' (Beauvoir, 1976, p.24). The accounts of valuing freedom in both Kant and Sartre connect it to man's fundamental desire for a meaningful and justified existence. But what constitutes such an existence is markedly different. For Kant, as discussed above, a meaningful and justified existence is a unified one entailing principles: because we must act, we need reasons to act and therefore we need some principles. The principles we choose relate in certain ways to the conception of practical identity that we adopt, for example we define ourselves as members of a religious or ethnic group, or by gender or career choice – the argument being that our way of identifying ourselves provides the source of our reasons. A conception of practical identity is, as Korsgaard writes, 'a description under which you value yourself and find your life worth living and your actions to be worth undertaking' (Korsgaard, 2009, p.20). On this account, it is only if we are unified by our principles that actions can be attributable to us as agents, without this unification, our lives are merely a random collection of impulses. For Sartre however, the route to meaning and justification is far more complicated.

Historically, man has looked to satisfy this desire for meaning by recourse, for example, to faith in a divine creator, perhaps suggesting recognition by man of his own contingency. But Sartre is clear that morality, and therefore the concept of value, relates strictly to the *human* realm: 'if morality is the law that regulates through the world the relationship between human reality and itself, the first consequence is that human reality is obliged to account only to itself for its morality' (Sartre, 1984. p108). This thought is reinforced by Sartre's typically sardonic statement 'Whether God exists or does not exist, morality is an affair "between men" and God has no right to poke his nose in' (ibid). Clearly though, Sartre does have a concern for morality within the human realm, his ethical commitment is evident when he writes: 'We always choose the good, and nothing can be good for any of us unless it is good for all' (Sartre, 2007, p.24). Within the human realm, valuing my own freedom necessitates valuing that of others.

Whilst upholding the absolute value of freedom, Sartre asserts that the authentic person is responsible for the free creation of value, and having accepted this, 'he can will but one thing: freedom as the foundation of all values' (Sartre, 2007, p.48). So, according to Sartre, values are freely created by authentic consciousness, and the implication of this is that they would seem to be subjective and particular to the individual. However, he also maintains that freedom is absolutely and *universally* valuable. These two ideas appear to be in conflict: if all values are relative to an individual, how can one specific value be considered universally valuable, and therefore higher than the others?

Values are disclosed in the light of our chosen ends or projects, in the recognition that something is lacked. A lack appears as such in the face of an acknowledgement that things could be otherwise, that a desirable situation is possible but not yet actual, for example. So values, for Sartre, show up only in the light of a transformative project, and they demand to be acted on; if a value is experienced as positive, it makes a demand that it 'ought to be' (Sartre, 1984, p.88). To reiterate a point made in Chapter one, the reference to values 'showing up' or 'being disclosed', implies that value is a response-independent concept, it applies to the nature of objects in themselves, rather than on the cognitive impression that the object makes on the subject. But unlike Kant, who seems to inspire an either/or response to the value realism/anti-realism interpretation, Sartre can be read as upholding both that values are created and that they are discovered. Values are not facts, which might be recognised by an indifferent observer, but importantly for Sartre's ethical position, neither are they absolutely agent relative: 'Nothing can be valuable for me or any other particular agent unless it is in this sense valuable *simpliciter* and *therefore* valuable, although not necessarily with the same determinate valuational force, 'for all' who adequately grasp the relevant features' (Poellner, 2015, p.237). Values must appear to a consciousness, in relation to a project, but this should not be taken to imply that all values are wholly subjective. The concept of smallness in relation to a mouse sitting next to an elephant requires a consciousness to apply it. But this does not mean that it therefore depends on this consciousness, the mouse really *is* small in relation to the elephant. So whilst some values are clearly subjective, the value of bread to a starving person for example, the disvalue of cruelty when applied to the torture of a child is not. Value is then dependent on consciousness without being internal to it, but is not dependent on an individual consciousness 'just as there can be no actual objective value without some consciousness to which it manifests itself, so there can be no free consciousness without value-manifesting objects' (Poellner, 2015, p.240). Some values are not agent relative but shareable by all subjects because they are essentially universal – the beauty of a sunset is not indexed to an individual viewer but can potentially be appreciated universally. Sartre's inference of a hierarchy of values in his claim for the absolute status of the value of freedom underscores this.

Sartre takes two different approaches to explain why he thinks that freedom has the status of absolute value, one based on the ontological structure of human reality, the other on phenomenological experience. According to Sartre, the ontological structure of human consciousness is such that it is necessarily directed towards ends or projects, has an awareness of its current situation as incomplete, or lacking, and has a desire to overcome this lack and escape the gratuitousness of existing without having a foundation. Absolute value then is

located in the freedom which pertains to a necessary and self-founded being: 'The source of all value, and the supreme value, is the substantiality or nature of the being which is its own foundation' (Sartre, 1984, p.111). The desire of human consciousness involves becoming a necessary not contingent being, whose non-existence would not be possible, which Sartre likens to the desire of 'becoming God'. As I argued in Chapter one, this approach, based on ontological structure, cannot overcome the charge of contradiction. However, if we pursue the alternative phenomenological justification, can freedom be understood as absolutely valuable?

### The goal of completion.

According to Sartre, the desire for completion is, in effect, the desire of consciousness to *deny* its freedom by becoming object-like. In the inauthentic life, in bad faith, this desire to 'become God' is taken as realisable, and in this state, Sartre describes man as a 'useless passion' (Sartre, 2005, p.636).

If man is ontologically structured to believe he can realise unrealisable goals, he does indeed appear to be a 'useless passion' and questions of a moral nature become irrelevant – it matters little *how* a man lives if he is doomed to failure from the start. However, in the final section of *Being and Nothingness*, Sartre suggests that it is not inevitable that man remains in the state of bad faith. Man chooses values in the light of his freely chosen goals, and he does this in view of an ultimate goal, the value or goal of becoming God, which Sartre claims is a pre-reflective aim, shared by all intentional consciousnesses. But by emphasising the subjective aspect of value – as conferred by humans – Sartre reasons that the value of becoming God, like all other values, has no objective or intrinsic value. It only appears to have this kind of value to those who believe both that values are written in things, *and* that they are independent of human subjectivity. So if man achieves self-transparency through pure reflection, and gains awareness of the incompatibility of this fundamental project with the structure of consciousness, he might see that instead of taking up a project which is doomed to failure, he might choose freedom itself, which he would then recognise as the source of all value, as his ultimate value: '...will freedom by the very fact that it apprehends itself as a freedom in relation to itself, be able to put an end to the reign of this value? [the value of becoming God] In particular is it possible for freedom to take itself for a value as the source of all value...' (Sartre, 2005, p.647). By raising this question, Sartre at least opens up the possibility that man might not be doomed to failure after all. If nothing has totally objective value, all that is left to be a candidate for *ultimate* value is the freedom to *confer* value: 'Only by valuing freedom first and foremost will any of the

values it subsequently creates be meaningful or justified, including the value it places on my own existence' (Anderson, 1979, p.47). The idea is that if, as humans, we necessarily must value *something*, we must value that which makes valuing possible at all, and this something, for Sartre, is freedom itself.

The phenomenological approach referred to earlier, involves the notion of the desire for completeness as unrealisable but coherent. That man desires to surpass himself in order to be the foundation of his own future is plausible – consciousness 'projects a certain future of itself beyond the world, on the horizon, in the illusion that when it becomes that future, it will be so in the guise of its own foundation' (Sartre, 1984, p.110). The undesirable lack that is evident in man's end-directed character (Sartre, 2005, p.111-5) suggests that man is continually attempting to achieve completeness. The aim of the for-itself of achieving coincidence with itself amounts to the phenomenological aim of eliminating experiential lack. Man is then aware of a lack for which he is somehow accountable, and he is oriented towards overcoming it in completion, which he knows is not possible. This coincidence of human reality with itself is not possible to the extent that it cannot be *realised* – but this does not mean that it could not exist or that it cannot motivate.

What we can see from this discussion, is that Sartre's concept of the unity of the person extends beyond the rational unity of principles demanded by Kant in the quest for a meaningful life. An authentic subject is orientated towards experiential completeness both of itself and of others – this goal of completeness must encompass the whole of humanity, with its environment. On a phenomenological level, Sartre offers only hints at what he means by the desire for 'completion', and 'coincidence' with oneself. But does Sartre's broad commitment to valuing freedom equally in others as well as ourselves constitute an ethical approach to existence? Whilst signalling an intention to do so, Sartre did not produce a complete theory of ethics. There have been attempts to construct such a theory by extracting from the works we do have, but from Sartre's own writing, particularly in *Being and Nothingness*, some serious doubts arise as to whether Sartre's concept of the human subject and intersubjective relationships preclude any kind of empathetic involvement. The 'other' in Sartrean terms appears to be a negative factor in being a locus of conflict, and a positive factor only as a potentially useful object in my own quest for self-justification:

For if in one sense my being-as-object is an unbearable contingency and the pure "possession" of myself by another, still in another sense this being stands as the indication of what I should be obliged to recover and found in order to be the



foundation of myself. But this is conceivable only if I assimilate the Other's freedom (Sartre 2005, p.386).

This is hardly compatible with anything we might understand as empathy, although it should be acknowledged that Sartre's position on this subject was somewhat revised in his later works. What seems to motivate Sartre's discussion here on our being-for-others is his very specific ontology of man's fundamental desire to become God. This is an ontological goal, the coherence of which we have already called into question. My interest then is not so much in ascertaining what a Sartrean ethics would look like, but rather whether an ethics based on the absolute value of freedom is available, or whether we must give up any part of that freedom if we are to live ethically. I suggest that there is a way of seeing the other and respecting his values which does not compromise our freedom.

### Valuing the other

On Husserl's account, in empathy the other is given to me not as an object, but as a centre of orientation – the other is given as a subject directed at the same world, understood from a different perspective. The consequence of this is that 'As soon as the other appears on the scene, my relation to the world will change, since the other will always be given to me in a situation or meaningful context that points back to the other as a new centre of reference' (Zahavi, 2014, p.137). I perceive the other's anger, but it is essentially given to me in a situation revealing a different perspective. Stein writes 'Here emerges the possibility of enriching our own world image through another's, the significance of empathy for experiencing the real outer world' (Stein, 1989, p.62-3). Empathy then becomes a source of self-knowledge, the perspective of the other and the meaning it has for him affects the meaning the world has for me. The 'other' then is, if empathetically regarded, not a threat but an opportunity for mutual fulfilment.

A closer look at emotions in general offers an explanation as to why this should be so. Understood phenomenologically, some emotions seem to reveal to us things about our environment – we come to understand a work of art as beautiful, for example, through the experience of surprise, or joy, and trying to say anything about what it is for the painting to be beautiful necessarily involves referring to feelings of a relevant kind. In these instances there is no clear divide between rationality and emotion. Recognition of this has led to the idea that basic emotions either are, or are analogous to, perceptions, and that the representational element of emotions is a perception of value. Values then are understood as essentially

phenomenal. If this is so, emotions have the potential to reveal a world of values to us in an experientially significant way. The theory is not without its critics, but if it can answer the objections that have been raised, emotions become ethically relevant as the way in which our most basic access to values is given to us. If the theory holds, rationality is not accepted as our primary means of accessing the moral domain as Kant suggests. Rather than formally reasoning towards an understanding of a situation, or subsuming facts under principles to reach a moral viewpoint, the idea is that we register the complexity of a situation by means of emotional perception. If we understand emotions as revealing evaluative properties of the object as it is given to us, we gain an insight into some of our most deeply held convictions which are otherwise hard to explain – the fact that we *know*, for example, that the Holocaust is abhorrent in a way that is different from merely holding a rational belief that it is so. Pelser writes: ‘while I might believe that slavery is unjust and I might be averse to slavery because of my belief that it is unjust, until I have been angered by or felt indignation toward an instance of slavery, I have not directly “seen” the injustice of it’ (Pelser, 2014, Ch7.2). My rational and general belief that racism is hateful, although strongly held, has an effect on me that is markedly different from an actual instance of discrimination, whether experienced directly, or witnessed. Further to this, if the content of emotional experience is evaluative, the normative content is built into the experience itself, there is no necessity to posit any additional component to indicate what ‘I ought’ to do and why. So are we justified in holding the idea that emotions are, or are analogous to, perceptions, and that they reveal values to us? Recalling Vetlesen’s three level sequence which facilitates entry into the moral domain: perception; judgement; action, I would like to consider the suggestion of a close link between emotion and perception.

## Perceptual theory of emotions

For the purposes of this discussion, a brief outline of this broad and complex argument must suffice. However it is worth noting that it is a view which, as well as attracting contemporary attention in the form of a revival, was held by classical phenomenologists, such as Scheler and Sartre, who claimed that ‘emotions are construed as, potentially, perceptual disclosures of values’ (Poellner, 2016a, p.261). Broadly speaking, perceptual theories of emotion identify several significant parallels between sense perception and some emotions. Both are initially receptive experiences, caused by aspects of the environment of the subject, and both are intentional mental states with phenomenal properties. Access to the phenomenal characteristics of the intended objects seems to be direct in both cases; in sense perception the colour of a red object, for example, appears as directly perceived, as do evaluative properties

such as beauty, or injustice in the emotional case. Whilst neither sense perceptions nor emotions can convincingly be explained merely as judgements, both can give rise to beliefs that are subject to change in accordance with the beliefs about their object. I can believe that my friend has arrived but then realise I mistook the sight of another person for him, and so my belief changes. An appropriate emotion such as indignation at an unkind remark might subside on the realisation that it was intended as a joke. The analogy can be extended to cases of irrational emotions, which are understood as perceptual illusions; genuine horror experienced by a person seeing a mouse whilst being fully aware that the mouse is harmless, is understood as a case of emotional illusion rather than self-contradiction. We can experience emotions in cases where the object is not *directly* perceived, in imagination, or memory for example. In these cases the perceptual account would suggest that representations of the evaluative experience would also be indirect. In imagination, emotion would proceed from an imagination of direct experience; ‘the evaluative properties – determinates of fearsomeness and beauty, respectively – would need to be *intuitively* represented via images or analogia taken as resembling evaluative features of the intentional targets’ (Poellner 2016a, p.262).

However, if sense perception and basic emotions provide direct access to phenomenal properties of objects, emotions are typically particular, evaluative perceptions, and in this respect they seem to differ from sense perception. Perceptual experience is generally understood as transparent – a visual experience for example can be described purely by properties of the object involved, including the particular point of view of the subject, and the environment in which the object is viewed. To describe my visual experience of a car, I describe aspects of the car’s appearance, it is red, it has four wheels, and so on. By contrast, emotions are typically particular, evaluative perceptions, involving valenced attitudes such as approval or disapproval, and in this respect they differ from sense perception. So does the inclusion of valenced attitudes in emotions mean that whilst analogous, they cannot, after all, be considered literally as perceptions? Not necessarily; whilst objects of sense perception are perceived as possessing qualities such as colour or dimension, objects of emotion are perceived as possessing specific evaluative properties such as injustice or sublimity, direct awareness of which includes a feeling of approval or disapproval; the perceptual acquaintance with the injustice *is* my experience of anger, perceptual acquaintance with the sublimity of the painting *is* my experience of admiration. In both direct and indirect emotional experiences of value properties, values are experienced as motivating; a value ‘presents itself as an *uptake* of the value’s *pro tanto* justified “demand” to be or remain actual (if it is a positive value). And it is that property of meriting-to-be-actual that partly constitutes a positive evaluative property as such’ (Poellner, 2016a, p.266). To experience that property is to register that it merits to be, or

remain, realised. According to Scheler, 'this is just what the valenced attitudinal component of the emotional experience – in, for example, admiration – presents itself as' (Poellner, 2016a, p.275). The value properties are not given *in addition* to the emotional experience, but are part of the emotional content, this is just how the contents of emotional experience are given.

To respond to concerns about the claim that evaluative properties are part of the content of emotional experience, consider an example of the experience of awe: if I am awed by a painting, and am asked to explain this experience by reference to non-evaluative properties, I immediately run into difficulty. I might start by pointing out that it is beautiful because of the symmetry, perspective or colour balance or I might declare that I like it because it complements the colour scheme in the room. But if I have to explain my actual experience of *awe*, I cannot do this without referring to its evaluative qualities. Equally, I might well experience the painting as beautiful without consciously registering its non-evaluative properties, the arrangement of its specific elements and colours for example. However, Dokic suggests that: 'the apparent evaluative content of emotion is in fact the product of an informational enrichment initiated outside emotion' (Dokic, Lemaire, 2013, p.228). Whilst it is true that experiential phenomenology is permeated by previous experiences or judgements, it is not the case that it is this non-evaluative aspect that constitutes the content of the emotional *experience*. If Dokic and Lemaire are right, and emotional experience is a response to non-evaluative information, informational enrichment should provide knowledge that, for example, the mouse I observe is harmless, and therefore prevent genuine anxiety. However, in that example, even when I do possess the information that the mouse is not dangerous, my emotional experience is still real. Emotional experience is not a response to non-evaluative information, but to that which must be characterised in valenced terms, for example the elegance of a piece of music or the unkindness of a remark.

The example of emotional illusion involving phobia might lead to a further concern about the acceptability of the theory. In the emotional experience described above, the emotion cannot be wholly explained by reference to properties of the mouse, but involves the reaction of the subject, the *felt* quality of anxiety, and in this, it differs from perception. To reiterate, this theory does not claim that all emotion types fit exactly into the perceptual model. Some emotions, of which fear is an example, have self-referential components; to be afraid, necessarily involves awareness of oneself as threatened, and relevant physiological responses. But even discarding, for the moment, emotions of this type, if we take an emotion such as awe, which is not obviously self-referential and does not necessarily involve any bodily response, it still involves some kind of feeling, approval or disapproval for example. Can we say that this

value component is a *property* of the object rather than merely caused by it? The conviction of the (in)appropriateness of an emotion suggests that we can. The disgust felt at the cruel treatment of an animal is experienced as an appropriate response. If, by contrast, I am in a bad mood due to receiving a tax bill and fly into a rage at some mild carelessness witnessed, the carelessness has perhaps motivated my fury, but I am aware of an experience of it as an inappropriate response. The limiting of examples to certain types of emotion does little to diminish the importance of the theory; 'If at least some emotions are indeed originary intuitions – perceptions – of values, it follows that their condition of success is veridicality, and that there can be literal (i.e. not "quasi-realist") knowledge of genuine moral, aesthetic, and other evaluative facts' (Poellner, 2014, p.300).

### The need for empathy

If emotions can be understood as giving us direct access to values, how can the specific emotional faculty of empathy, which I suggested above is necessary for moral perception, be understood? As Zahavi points out, we do not have a clear consensus as to what empathy actually is (Zahavi, 2014, p.129). Empathy can be understood variously as sharing the feelings of the other either by experiencing in some way the same experience, or by imagining being the other, or by making inferences about their mental states based on our own experiences. There are many different models which aim to explain what empathy is, but rather than contributing to that enquiry, my aim is to concentrate on a phenomenological understanding which focuses on the intentional structure of empathy, in order to highlight some particular insights which it provides. So for this purpose, I will accept a basic definition of empathy provided by Stein, who claims that what we are trying to understand is how we experience others: 'All these data of foreign experience point back to the basic nature of acts in which foreign experience is grasped. We now want to designate these acts as empathy' (Stein, 1989, p.6). Zahavi writes 'none of the phenomenologists would accept the claim that one can only empathize with affective states, rather they would take empathy to refer to our general ability to access the life of the mind of others in their expressions, expressive behaviour and meaningful actions' (Zahavi, 2014, p.138). In the phenomenological account, empathy is not just a question of mirroring or imitating another's mental state, or an instance of imaginative projection or reasoning by inference, rather it is a *direct* experience of the other, which takes account of his values. But whilst I can directly perceive your mental state, your joy or grief for example, this isn't given to me originally. If it was, there would be a fusion of consciousness; your joy would be my joy. And the empathised experience *necessarily* belongs to the other. It is unmediated and non-inferential,

and in this respect it does resemble sense perception. But the direct access to the experience of the mind of the other in empathy is of a kind different from the access in self-awareness.

Husserl explains 'Just as what is past can be originally given as past only through memory, and what is to come in the future can as such only be originally given through expectation, the foreign can only be originally given as foreign through empathy. Original givenness in this sense is the same as experience' (quoted in Zahavi, 2014, p.135-6). It is the difference between my experiential access to my own mind and my experiential access to the minds of others that constitutes empathy, and it is this difference that allows me to experience the other *as* other.

On a perceptual account, in experiencing the emotion of sympathy for the pain of another for example, the pain is given to me as a disvalue: it should not be. I perceive the other's pain and empathise with the other at the same time. Using the explanation Husserl gives for perceptual experience, we can understand the pain as-experienced-by-the-other as being co-presented with the facial expression and bodily behaviour that we directly observe. In a visual experience, when I look at a house from the front, I see only one aspect – the back and the underneath of it are hidden from me. However, I assume that they are there, and that the house is a three-dimensional structure. The hidden sides are co-presented with the parts of the house that I do see. In a similar way, when I see the bodily expression, which is the outer side of a mental state, the inner experience, which is one with that mental state, is co-presented. With the direct perception of the physical expression of the other, we perceive his inner mental state as present here and now. The recognition of what is co-presented in empathy is possible because of our shared humanity; 'the call for empathy can be met because we are all human beings, principally sharing the same access to the experience of pain' (Vetlesen, 1994, p.119). The pain is necessarily that of the other, but the access to it is shared in empathetic experience. Sartre writes 'For my consciousness, one thing alone exists at that moment: Peter-having-to-be-aided. This quality of "having-to-be-aided" is to be found in Peter' (Sartre, 2004, p.18). Empathy is the basic mode of access by which we understand Peter's situation, it precedes, and is necessary for, the moral judgement that follows. Importantly, if I do not perceive the situation as one essentially involving the 'weal and woe' of a co-subject, I will fail to see myself as obligated by it. Someone who is indifferent to the suffering of another can be said to be just not 'seeing clearly'.

But empathy goes beyond sympathy or shared joy. Above I referred to the sequence necessary for entry into the moral domain; perception, judgement and action, and claimed that Kant disrupts this sequence by discounting emotion from the ethical domain. In order to see how things are with the other, we have to understand not only his situation, but his values. If each

individual is free to determine his own being, his actions will reflect his choices, he will choose to make real in the world those things that his distinctive personal form of being perceives to be good, whilst diminishing those that he considers bad. The values good and bad are, in the light of the individual's view of existence, undoubtedly real aspects of the world. However, in existing alongside other individual consciousnesses, an empathetic outlook will necessitate seeing the other's ways of valuing as equally real for him, differences must occur between individual subjects as to the degree to which things are (dis)valuable; 'So long as we identify the world presented in our own consciousness with *the* world, we fail to realize that our values are just one limited set of the world's values. We cannot adequately take in the real character of what is going on in another consciousness without seeing the aspirations and fulfilments present there as having some validity of their own' (Sprigge, 1988, p.255). And, as John Berger notes, 'To try to understand the experience of another it is necessary to dismantle the world as seen from one's own place within it and to reassemble it as seen from his' (Berger, Mohr *et al.*, 2010, p.96-7).

So what kind of ontology is presupposed by these ethical requirements? An ethics based on values understood as emotional perceptions demands some kind of realism; as we have said values, as found, are really in the world. What is outlined here is an ethics where our moral motivation comes not from a rational principle, but from the perception of objective values which are given in intentional feeling, and from appreciating the equally real values of the other. Values are given in conscious intentional emotions 'as both objective and as inherently such as to "demand" or *merit* certain emotional responses ... which accordingly can be said to be "appropriate" to them (Poellner, 2014, p.300). But if values appear as qualifications of emotions, they require a consciousness in order to be recognised; they are given in the feelings they evoke in conscious, sentient beings. And these feelings are influenced by our own projects, by how we have chosen to direct our lives. Emotions themselves are not fixed entities, they vary from person to person and with different circumstances. So our realism must be of the kind that also acknowledges the contribution of individual consciousness, their subjective feelings and experiences. This demand is, I think, captured in the claim of Maurizio Ferraris:

On the one hand, we should maintain that there is an unamendable kernel in being and experience that gives itself in complete independence from conceptual schemes and knowledge. On the other hand, we must leave open the possibility of constructing, starting from this unamendable layer, knowledge as a conceptual, linguistic, deliberate, and especially emancipatory activity (Ferraris, 2015, p.43).

## Conclusion to chapter three.

Having considered Kant's claim that any end suitable to play a role in morality must be 'given by mere reason' so it can 'hold equally for all rational beings' (Kant, 2012, 4:427), I have argued that this offers an incomplete account of humanity. There are feelings, desires, hopes and aspirations that we share by virtue of our humanity, and humanity must be recognised as more than our capacity for rationality. Our moral obligations must extend to whole persons, and real-life situations rather than rational principles and laws, and for that we need empathy. If we reject Kant's ideal of human rational perfection and moral virtue as an action guiding vision, we should suggest an alternative. By taking up our freedom as the source of value, and the ultimate value, we are free to pursue our own projects and to live in accordance with what is valuable to us, taking responsibility for what we choose. But we also recognise that our values are grounded by how the world really is, and this is a world which we share with others. Our freedom allows us to see not only how things are and how they might be for us, but to extend that to appreciate the freedom of the other, respecting their values and perspectives. The life of a free rational being then, does not necessarily lead to an isolated, amoral, individual existence, as indicated in a quote from Ellen Dissanayake:

Once everyone recognises that customs are made by humans themselves rather than divinely ordained, perhaps the world – derinded of superstition and fatalism – can be taken into human hands and molded to human needs. If all can accept that there is no justification for suffering, that pain and misery are arbitrary and not supernal punishments for wrongdoing or natural inferiority, we will perhaps be better able to reduce their ubiquity. If we agree that neither love nor light nor certitude nor peace nor help for pain occur in the nature of things, but exist only by virtue of our willing them into being, we might then ourselves create and maintain these qualities rather than bemoan a lost paradise. (Dissanayake, 1991, p.193-4).



## CHAPTER FOUR: THE QUESTION OF MEANING AND ANSWERABILITY

In the first two chapters, I set out to establish whether we could credibly consider ourselves as absolutely free and autonomous beings, and argued that we could. Whilst we cannot prove beyond doubt that our lives are *not* ultimately determined, attempts to show that they are have certainly not been conclusive. Furthermore, I have argued that we have a moral responsibility to embrace our freedom and live as ontologically free, autonomous beings. Acknowledging and utilising our freedom to own the decisions we make about who we are and what we achieve, whilst attempting to change the things that are wrong in our world, is a necessary step in the process of fulfilling our human potential. In fact, according to Sartre, the ability to see things both as they are, and as they are not is a manifestation of our freedom: 'If man could not, first, describe a present given situation both as it is *and* as it is not; and if he could not, secondly and consequentially, envisage a given situation as possibly being otherwise than how it is, then he would have no power to intervene in the world to change it' (Sartre, 2001, p.xvii). In chapter three, I considered whether living a free and autonomous existence necessarily gives rise to an individualistic and amoral existence. Given both the process of value formation, and the rational requirement to value not only our own freedom but that of others, argued for in the accounts of both Kant and Sartre and outlined above, that clearly does not follow. This leads us to an understanding of humans as free, autonomous beings who can justifiably take full moral responsibility for their actions and the consequences of them. It also allows us to think about what we value, what kind of life we want to achieve, and how we might act, in a way that offers the best chance of realising our ideals.

In this chapter, I would like to consider whether this is enough. In spite of his conviction that we are radically free, Sartre acknowledges an experiential sense of lack and an inescapable orientation towards an unrealisable goal; to present our ultimate ends as realisable is, for him, merely to pretend, and to be in bad faith (Sartre, 1984, p.199). This human condition suggests that living a free and morally responsible life will not alone ensure a contented or fulfilled existence. But whether or not we are ultimately condemned to dissatisfaction as Sartre suggests, we are often drawn to further questions about whether we can understand life as having significance, or being meaningful in some way. Is a sense of meaning achievable if we are answerable only to ourselves and other members of our society, however this is defined, with regard to the actions we take and the people we become? And what do the terms 'meaning' or 'significance' actually imply in this context? With no apparent divine principle to appeal to and no pre-determined path to follow, we are left with some serious decisions about the point and purpose of the life which we have somehow acquired. Perhaps it is our

experienced sense of lack that drives us to explore these wider questions about whether our lives do, or might, have ultimate meaning, understood as a profound significance that goes beyond human purposes and interests, one that could adequately answer to this unfulfilled desire. Sartre makes the claim that 'the for itself is the being by which "there is" a world' (Sartre 2005, p.470), but can everything really depend on us for its significance? These are questions that are of utmost importance for many of us.

## Defining the question

Whilst being important, such questions are rarely clearly defined or concisely articulated, the issues raised by them are not fixed or static areas of enquiry. The idea of meaningfulness in relation to life itself has many components, leading to very different ways of attempting to answer the question. One component might relate to whether our lives have an overall purpose, whether there is an answer to questions such as 'what are humans here for?', or 'is there a role that we are supposed to fulfil?'. Another component might be the call for a wider context; everything we encounter is experienced against a background, and this provides context which enables intelligibility. For example, a question such as 'are you ready?' means nothing unless the context is understood, we need first to know 'ready for what?'. So, by meaning, we might intend context, in the sense of a wider narrative illuminating our desires and actions. We also find meaning in the connections we make between things, a cup handle in complete isolation from a cup cannot be thought of as making real sense. If meaning is a matter of making connections, the meaning of our lives might be indicated in the way we connect with a wider horizon beyond our limited perspectives, how we connect with 'reality' if this is understood perhaps as 'the totality of all there is'. And in looking for connections with what is beyond our subjective point of view, the investigation will involve, to some extent, testing and attempting to surpass the limits of what we can know and do. We will want to push beyond the boundaries to expand our reach. This might lead to an enquiry into our potential as humans: can we not only expand our knowledge of what there is, but contribute to the total of what there is? And there is an additional component to the broad question with which we began, which is whether there is anything to which we are answerable. We are familiar with the idea of being answerable to other humans in various ways, but is there anything beyond what is immediately accessible to us, against which we might measure or assess our lives? Whilst I think that these different areas converge in many ways, I will deal with them initially by dividing them into two separate areas of interest: meaning; and answerability. This will entail considering what kind of questions can usefully be asked in relation to these two concepts. The

enquiry necessarily involves conceptual understanding in order to provide content, but this can only be partial, there is much we cannot know for certain and an element of ineffability will certainly be encountered. But at the point where concepts fail, by turning to experience, and in addition by introducing the mechanism of metaphor, I suggest that we can make some progress in this investigation.

### Intrinsic meaning

Relatively casual questions about whether life has ultimate meaning have a kind of familiarity within everyday conversation, but this reference to meaning needs qualification if it is to be adequately addressed. I take it that the question relates primarily to the concepts of relative importance, value, and objective significance in regard to human life. But we will need to consider which of the many components we need to focus on, in order to arrive at a satisfactory response. At its most extreme, what appears to be demanded is an ultimate overall purpose to our existence, perhaps a blueprint or set of reasons that make sense of what seems incomprehensible to us. Unless we can argue for an external source which confers this kind of meaning on human life, then life itself, if it is to be understood as meaningful in this sense, would need to be self-sufficiently so. To consider what seems an improbable proposition, I will outline the proposal made by Robert Nozick, who suggests that there *is* a way of understanding life as meaningful, and that the totality of everything there is, can be thought of as uniquely, *intrinsically* meaningful (Nozick, 1981, p.604-5).

According to Nozick, attempts to find meaning involve a process of transcending the limits of an individual life, and asking how it is connected to other things, to values beyond itself: 'The narrower the limits of a life, the less meaningful it is' (Nozick, 1981, p.594). A solitary life, motivated purely to satisfy the basic requirements for survival, with no aims or goals, strikes us as a fairly meaningless existence. The meaning of something then seems to relate to how it connects with what is outside it, and also greater in stature. By deciding how you transcend your life and which external connections you will form, you give your life a specific character. A fundamental limit is the fact of our finitude, and so we look beyond our own limited existence to something for the sake of which we act. In the face of our own mortality, we might look for meaning in something that will outlast us, maybe dedicating ourselves to the improvement of some aspect of society, or the creation of a work of art, or producing descendants to continue a family line. In this way we appear to transcend our limits. But the problem is that unless our created works, or our descendants, are intrinsically meaningful, it is not clear how meaning

transfers to us in this way, the question just seems to be postponed and the responsibility passed down to others. Nozick observes that for anything that we hold to be meaningful, there will always be a wider standpoint from which its significance could be diminished: 'However widely we connect and link, however far our web of meaningfulness extends, we can imagine drawing a boundary around all that, standing outside looking at the totality of it, and asking "but what is the meaning of that...?"' (ibid., p.596). If we hold, for example, that our life has meaning because it contributes to the overall wellbeing of the members of our particular society, a viewpoint that takes in the whole nation will reveal our society to be just one small, relatively insignificant group of individuals amongst billions of others. From there, it is conceivable that we might reevaluate our life's work and judge it, after all, to be meaningless. If our life seems to have significance for us because it makes a real positive contribution to the lives of all humans, when we consider that humans are mortal and the whole human race is dependent on the diminishing lifespan of the sun, even this project is impoverished. And as Nozick points out, even if we were to posit an omnipotent God whose plan involves a specific role that we are created to fulfil, we can still question whether that plan ultimately has any meaning, and why, leaving aside the question of reward or punishment, we should be motivated to conform to it. We might desire to fulfil God's plan on the grounds that God is good, but why should doing good, in itself, relate to meaning?

The problem of meaning then shows up in the question of limits. To recognise that something is *limited*, to describe it as 'merely that', is to question its meaning. So, Nozick argues, we need to look to what is unlimited. The questioning about meaning only stops with the *totality* of all there is, in the unlimited or all-encompassing, from which there are no limits that can be observed from outside, because there is nowhere outside to stand.

Before considering Nozick's idea, we have to ask whether the idea of the existence of the all-encompassing is coherent, and if it is, why should it be a meaningful rather than meaningless entity? If existing implies limitation on the grounds that to be, is to be something rather than another thing, the all-encompassing cannot be said to exist, it transcends the category of existence and non-existence. Similarly, the pair of terms meaningful-meaningless presupposes that what is so described is either meaningful or meaningless in relation to something external. How then can the all-encompassing be understood as either meaningful or meaningless, when it must equally transcend those terms? Although understood in this way the unlimited provides a stopping point for questioning about meaning, it appears to do so by transcending the category rather than by being meaningful itself. So, Nozick's question still remains: 'How in the

world (or out of it) can there be something whose nature contains meaning, something which just glows meaning?’ (Nozick, 1981, p.593).

Nozick argues that the unlimited, if it can be said to exist in any sense, does provide meaning, and it does this by becoming its own source of meaning. It can do this only *because* it is unlimited. If something has meaning, it is found by looking at its wider context, which for a limited being, will be something external to it. But for the all-encompassing, which includes all actuality and all possibility, there is nothing external. Therefore, its wider context must be contained within itself: ‘only an unlimited being can map onto and so connect with something apparently larger and external which turns out to be itself. Only an unlimited being can have its “wider” context be itself, and so be its own meaning’ (Nozick, 1981, p.603). Nozick bases his reasoning on an analogy with the mathematical concept of the infinite set: with a finite set of positive integers, there cannot be one-to-one mapping between a set and a sub-set of itself – something will always be left over. This is not so however with an infinite set: the infinite set can expand without limit to allow for one-to-one mapping with a subset of itself. Nozick clarifies this with a further explanation: ‘Only an infinite being could fully specify all its own details, including the details of that very act of specifying. To the extent that the meaning of something fully specifies it, only something unlimited could be its own meaning’ (ibid., p.747 n16). So, if there is meaning to be found, it must be found within the unlimited, the all-encompassing, or totality of all there is. In this way, the all-encompassing can be understood as being its own meaning, as intrinsically meaningful.

If we accept Nozick’s reasoning for the moment, it clearly isn’t enough that we vaguely believe that the all-encompassing can be understood as intrinsically meaningful, for our lives to suddenly gain significance. If a difference is to be made, we need to form some kind of *experiential* relation with that source of meaning. How might it be possible to *encounter* meaning in the totality of all there is? Even if the all-encompassing exists, and is intrinsically meaningful, how might meaning transfer to us?

One possibility is that we reach a realisation that fundamentally we also are unlimited, that we are by nature, identical with the totality. There are many religions and worldviews in which this belief is held. Spinoza for example held that the universe is a necessarily existing, perfect whole, within which our lives take place as an essential part of the necessary cosmic process. For Bernard Bosanquet, we are part of the totality, he writes ‘I am working with the idea which I have maintained throughout, that the universe is one, and each finite mind is a factor in the effort which sustains its unity’ (quoted in Sprigge, 2006, p.331). In Hindu Philosophy of Advaita

Vedanta, the true self, or fundamental identity (Atman) is identical with the highest metaphysical reality (Brahman), from which all things emanate. Sprigge, commenting on the work of F H Bradley, writes 'to understand anything is to see it as a necessary element in a systematic whole possessing a certain completeness to it, or as seeking to become or to join such a whole itself' (Sprigge, 2006, p.272). For some absolute idealists, consciousness as such is a common essence, present in all people, animals and nature – Sprigge writes, 'it should be that essence in all its myriad finite forms, and its actualization in the great cosmic whole of all things, to which one's final devotion should be given' (ibid., p.543). There are many ways of understanding the possibility of identifying with the totality and this is an interesting idea, which being beyond the scope of this study, I will leave open.

Briefly returning to Nozick, his account introduces some key components which are central to the concept of meaning, but by arguing for them collectively under the broad headings of meaning and significance, it is hard to see anything particularly valuable emerging. Perhaps the all-encompassing totality might offer a context which is sufficient unto itself, but the tenor of the search for meaning suggests rather a call for a context of value. Nozick's reasoning appears at best abstract and theoretical. This is perhaps an inevitable result of the nature of the starting point, and maybe this is an instance of the assertion made by Deleuze that we get the answers we deserve: 'We always have as much truth as we deserve in accordance with the sense of what we say' (Deleuze, 2014, p.201). For our lives to be genuinely understood as having ultimate significance this must be expected to make a real difference to the choices we make. As James writes 'For what seriousness can possibly remain in debating philosophic propositions that will never make an appreciable difference to us in action? And what could it matter, if all propositions were practically indifferent, which of them we should agree to call true or which false?' (James, 1985, p.444).

If the mere possibility of an inherent meaning in the totality of all there is falls short of the vital answer we were hoping for, we could of course abandon the question. Equally, we could conclude that the entire human species is insignificant and that there is no ultimate meaning. What are the consequences if we do? An existential nihilistic existence is available to humans, as perhaps suggested by Nietzsche. This might lead to a life of despair and anguish, one of extreme boredom, but alternatively it might lead to one of hedonism, lived purely for the moment, driven by the immediate avoidance of pain, and pursuit of momentary pleasure, aiming for a happy nihilism. Some kind of morality or mutuality towards other humans is still conceivable in the face of existential scepticism. But although these examples are theoretically possible, a definition of them as lived existences seems to necessitate the use of negative

terms; even a description of a hedonistic existence lived for the pursuit of immediate gratification strongly implies that it is *only* that. What they hold in common is a sense that something is lacking.

Instead of abandoning the question, perhaps we can reframe it in a more pragmatic vein. The idea of an ultimate meaning or significance to everything that exists, appears to be too abstract to offer an adequate response, and in addition, it seems unlikely to be realised. But, as stated above, meaning can be understood to relate to purpose, intention and agency, involving questions about what humans can or should do, whether we are serving a greater purpose beyond our routine existence, what are humans aiming for and what can we meaningfully achieve? Having established that we are autonomous beings, our purposes as individuals must also be freely chosen. We can choose to make the fulfilment of our own potential our purpose, taking up the challenge of becoming the best we can be. But we live in a world necessarily involving other individuals, and this introduces the possibility of a collective human purpose. Our enquiry then perhaps concerns the potential of humans as a species.

If fulfilment of shared human potential is our collective purpose, we will need to continually expand the realm of what is knowable. But can we also add to the total of what can be known? I would like to suggest that we find a meaningful existence in relation to purposeful activity towards this shared human endeavour of advancing the boundaries of what we can know and achieve. In the next two chapters I will develop this idea: in chapter five I propose the use of metaphor as a means of expanding our knowledge and as a tool for creativity; in chapter six I will discuss the potential for humans to create things that are radically new, and the suggestion that a sense of purpose results from our capacity to contribute to the sum of what there is.

Before addressing these issues, I will return to the second of the two main components of meaningfulness in relation to human life: answerability.

## Answerability

Like meaning, the concept of answerability can be understood in many different ways. Within the framework of human autonomy, many of the available definitions are clearly not relevant. If we are radically free, it follows that we are not answerable to a higher being in the sense of being under an obligation to respond in a certain way, to gain favour or avoid retribution. I do not think that there are a set of pre-ordained tasks which it is our mission to complete

correctly. Equally, answerability should not be confused with responsibility, we are, as I have argued, ultimately responsible for our own decisions, for the choices we make and the people we become.

The answerability intended here is linked to the need for purpose just described. If our purpose, as humans, is somehow to fulfil our human potential, against what do we measure our actions? We want our purposes to be worthwhile, and so we need some kind of external validation from a source beyond our own limited perspectives. David Cooper writes: 'To want to be answerable to ... is to want the elements of one's life – one's actions, commitments, and so on – and perhaps the whole life they constitute to be subject to measure ... a "qualitative criterion", that enables assessment' (Cooper, 2012, p.260).

Cooper's claim stems from the thought that the call for answerability is motivated by the desire that the actions and commitments that structure our lives are subject to some kind of external qualitative standard allowing for appraisal. Humans need not only to be, but to see themselves as engaged in projects that matter, to have aims and goals that provide a reason for living. And if projects are considered to matter, they must be subject to some kind of measure. Whilst we can be answerable to the distinction between truth and falsity for example, and we can learn by experience whether what we do is judged by our fellow humans to be better or worse than some given standard, the necessary element of subjectivity involved means that this does not go far enough to satisfy fully the demand for measure. What we need is connection with a source of significance beyond our human perspective: 'If significance is to be conducted back along the chain, its source must be "beyond the human", for whatever remains within the precincts of the human always inspires the question of its own significance' (Cooper, 2012, p.271).

In addressing the question of whether there is indeed anything beyond the human world, Cooper argues that the idea of the existence of a conceptualisable world, which can be conceptualised irrespective of human perspectives, is unbelievable. And to hold this absolutist view is, for Cooper, hubristic. It is to pretend to an exalted cognitive capacity by claiming that there is a world that we *could* conceptualise and understand even though it is totally remote from any context of human practice and so inaccessible to us as things stand. However, Cooper considers that to adopt the humanist assumption that there is *not* such a world, and that all there is, is the world as conceptualised by us, is unliveable. The adoption of what he describes as a stance of dis-incumbence; the 'posture of those who, accepting the thesis of "the human world", aspire to live in the recognition that, ultimately, their beliefs and values "lean on



nothing” beyond themselves ...’ (Cooper, 2012, p.210) is not something that humans are capable of enduring. A state of mind reflecting this is identified in the existentialist tradition as a feeling of *Angst*, or *unheimlichkeit*, loosely understood as an uncanny feeling, suggesting estrangement or not-being-at-home-ness. Sartre’s notion of absurdity refers to this tension created between the realisation of the absence of transcendent (rather than merely internal) reason or foundation for the choices we make which shape our lives, and the gravity and dedication we give to pursuing our projects and participating in life.

On Cooper’s account, humanism, uncompensated by a further doctrine that provides this answerability, is not liveable. Absolutism, as used to refer to an account of the world as it anyway is, independent of human perspective, is unbelievable, and uncompensated humanism is unliveable. So, Cooper concludes that there is a world beyond the human, but that it is not discursable, rather it is the realm of the unconceptualisable, or ineffable; it is mystery. Discourse can only capture a human world, and when the human contribution has been abstracted, that which remains cannot be explained.

One concern with this line of thought is the worry that it is underpinned by an argument from an identification of a real sense of dissatisfaction, lack or incompleteness to the existence of something to satisfy it, something beyond the human to which we are, after all, in some way answerable. But nothing suggests that there is any link between what we feel able to endure, and what there *is* ontologically. It is possible that dissatisfaction and lack is just part of the human condition. The point is surely that we want to know not whether we *need* to believe in something, but whether there *is* anything, beyond the human, a desire inspired by a search for knowledge of what there is. But a more serious concern is that Cooper’s conclusion returns us to the question raised above: why should the existence of something beyond the human provide meaning for our lives? Cooper claims that we need to be answerable to something beyond ourselves, and that the existence of a realm of mystery provides a measure by which we can assess our lives in meaning terms. But why should the mere fact of being beyond-the-human make an entity necessarily qualified to provide a measure for us? Why should we feel compelled to answer to that which is mysterious, just on the grounds that it *is* mysterious?

I would like to consider in more detail two statements made by Cooper: if projects are considered to matter, they must be subject to some kind of measure; and what we need is connection with a source of significance beyond our human perspective. Whilst agreeing with both these points, my aim is to respond to the challenge Cooper poses in his claim: humanism, uncompensated by a further doctrine that provides this answerability, is not liveable.

Ultimately, I will argue that uncompensated humanism *can* answer the two stated requirements, and therefore is not un-liveable on these grounds. Although a source of significance beyond our human perspective will not be established by rational argument alone, I will show that if we look closely at experience, something to which we are answerable can be detected. And whilst there are experiences which suggest interpretation as intimation of the existence of an order beyond the human, they do not demand it. I leave open the question of whether there actually *is* anything beyond the human realm, because my argument is just that we can find answerability and purpose without this being a necessary condition.

A traditional response to the problem of what, if anything, we are answerable to, is to invoke a God who has a plan and purpose for everything, including mankind. Although we don't know what this purpose is, we are required to trust that it is worthy to be followed because God is perfectly good. This then provides meaning; acting in accordance with God's plan is the point and purpose of our existence. But the meaning here provided is theoretical. The traditional theological account depends on an acceptance of the stated premise that God is necessarily a perfect being, and then stipulates what our response should be. The premise is something that we can only 'know' by inference, and the initial consent it requires is intellectual. We are required to accept that there is God, that he is necessarily perfect, and that we are able to know enough about him to adjust our orientation to life accordingly. But whilst this could potentially provide an acceptable answer to an intellectual enquiry, the concern is that if I am purely persuaded intellectually, what I might expect to *feel* would be more a sense of obligation, than a real sense of meaning that could answer an experiential lack. It must be acknowledged that there are people for whom an intellectual approach is absolutely appropriate: Spinoza for example believed that the highest human good is the intellectual love of God. For Spinoza, the universe is necessarily what it is, and consists in one substance: God, or nature, the existence of which Spinoza considered that he had proved. Hegel's route to a Christian model of God was by way of conceptual thought, which he considered to be the highest form of Spirit or Mind, and to be capable of revealing how things really are. The major monotheistic religions all offer intellectual arguments for the existence of God. However the classic arguments often fail to convince. One reason for this might be that as the arguments all purport to refer to the 'Highest One', and as the differences between the details of what they claim make them mutually exclusive in many aspects, at most only one of them can be right. As William James writes: 'If you have a God already whom you believe in, these arguments confirm you. If you are atheistic, they fail to set you right' (James, 1985, p.437).

An alternative approach is suggested by recalling that the search for meaning and answerability was partially set in motion by what Sartre describes as an *experiential* sense of lack. At the end of Nozick's account of life as having intrinsic meaning, there was a sense that the answer was not adequate, and this is perhaps because the 'totality of all there is', even if it exists, does not relate to anything that can be experienced. Nozick's account then is, at best, a partial response to a problem which, being about life as lived by us, requires the inclusion of experience. If Aristotle is right, our fundamental desire is for knowledge that we can believe in, and this implies knowledge that is *real* to us. The difference between things we merely know and things we believe, and which therefore become real to us, is to some extent an experiential one. There is a feeling attached to a strongly held belief. According to James, 'as long as we deal with the cosmic and the general, we deal only with the symbols of reality, but *as soon as we deal with private and personal phenomena as such, we deal with realities in the completest sense of the term*' (James, 1985, p.498).

### Meaning and answerability as experienced

So, I suggest that a search for answerability can be satisfactorily answered only by something offering the chance of *experienced* fulfilment. Timothy Sprigge points to the fundamental status of experience by making the radical claim that everything that exists is either an experience or part of an experience (Sprigge, 2006, p.478), asserting that the attempt to imagine something where its being unexperienced is part of what you are imagining – to imagine it as positively unexperienced – invariably fails. (Although it might be possible to *conceive* of something lacking features that would mark it out as being within consciousness, Sprigge argues that this is merely a form of indirect imagining). For Sprigge, we can understand ourselves as momentary centres of experience, or states of consciousness. We experience perceiving and acting on apparently external things, having particular thoughts and emotions, and we experience ourselves as the centre of our somatico-perceptual field. The character of what we are as conscious beings changes from moment to moment to make up a stream of total moments of experience. A detailed analysis of Sprigge's main claim would take us too far off course, but his account does highlight that what is undeniably real to us is experience, and this seems to be right; even at a really basic level, it is my experience of hunger that gives me the knowledge that I am hungry. So, in the search for answerability, as I suggested earlier, it is to experience that we should turn. And when we do this, we find that although our search for answerability by means of rational argument is unpromising, in experience, something shows up to suggest that the pursuit is still worthwhile. I would like to consider briefly the kind of experiences which

give us this encouragement, before discussing, in chapter seven, an account of answerability experienced within the paradigmatically human endeavour that is aesthetic creativity. I suggest that although a positive response to the desire for answerability looks unlikely within a philosophy of radical freedom, something mysterious that cannot be explained can be identified within aesthetic experience, and that the best explanation of this phenomena is that the artist is in touch with something beyond his immediate reach.

Clearly not all experience holds this kind of significance. There are common experiences which might indicate a slightly more meaningful dimension to life and intimations that our responsibilities might not just be those of humans to each other. We are used to the idea of a responsibility to care for the planet, not just as a human resource but because of its intrinsic value. But in terms of a call for a source of significance beyond our human perspective, these don't give us what we need. There are, however, experiences, the nature and depth of which it is less easy to understand. This kind of experience might usefully be described as mystical, where this is understood as different from ordinary conscious awareness, and as giving a strong impression of an encounter with a reality that is different from, and, in some fundamental sense, higher than that of everyday experience. As a concrete example of the kind of experience I am suggesting, imagine that whilst walking up a mountain, we vaguely picture the awe-inspiring scene awaiting us when we reach the top. When we do indeed reach the top, we fleetingly but completely *realise* that experience. Before we have time to think about it, or name it, we have it. The experience is total and un-fragmented, the experience itself lacks nothing. The relevant 'content' is fully experienced, but it cannot be adequately described. What is it about this phenomenon which might be, for convenience, termed as a moment of awe, that is different from the everyday experience of realising an imagined scene? The moment of awe is something that is, from our point of view, passive; it happens to us. Essentially, its nature is to arrest us, to bring us to a standstill. Our contribution to the experience is merely to have placed ourselves in a position to receive it. And, unlike for example, something that we might have previously or subsequently conceptualised as 'an adventure', with a composite nature, it appears to us as if it were outside the progression of time. If a composite adventure is realised as meaningful, it is so in a different way, the meaning is gradually grasped by making narrative connections in a temporal order. The mystical experience is unrealisable to the extent that it cannot be perpetuated; if we try to recreate it, to re-enter the experience, we find it can't be done at will. We can still experience the mountain view from a great height, but the actual moment of awe which seized us can, as far as volition is concerned, only be remembered. This moment of awe then, in its apparent

timelessness, seems to indicate an encounter with something at least beyond complete conceptual specification.

William James describes experiences with these characteristics as mystical states of consciousness, and suggests four marks by which we might recognise them as such (James, 1985, p.380-2):

1. Ineffability. They defy expression, no words can be given to describe their contents, from which it follows that their quality must be directly experienced.
2. Noetic quality. Mystical states seem also to be states of knowledge, of insight into depths of truth, to those who experience them. They are illuminations, or revelations, full of significance and importance, and carrying a curious sense of authority.
3. Transiency. They are of short temporal duration, whilst being outside our normal perception of space and time.
4. Passivity. Although there are activities that can enhance the possibility of mystical experience, the experience itself cannot be generated at will, it happens largely without conscious control and is sometimes experienced as emanating from a superior power (I think the aspect of passivity requires slight qualification, and I will address this in the final chapter).

To these four marks I would suggest a further two;

5. Surprise. Because mystical experience appears to come from outside us, it takes us by surprise. This is perhaps one reason why we find it fascinating. As Campbell writes 'One way to deprive yourself of an experience is indeed to expect it. Another is to have a name for it before you have the experience' (Campbell, 2013, p.13).
6. Scepticism. The very nature of mystical experience is such that the experiencer is susceptible both to incredulity from anyone with whom they confide, and to self-doubt. RK Elliott describes this well when he refers to Rodin who 'has just described enthusiastically how he had heard the bells of Rheims cathedral as the soul of the life of the city, and as a prophet turning this way and that to proclaim the festival' and then Rodin goes on to write "'Suddenly I hear", 'What a heap of rubbish'" (Elliott, 2006, p.43). It is suggestive of the ambitious nature of the experience that, whilst it is hard to understand and to defend, the experiencer feels a disinclination to discount it.

Again, this kind of experience is often associated with traditional religion; Christian mystics for example, might interpret such experience as revelation of our inclusion in the being of God. However, this list of defining qualities resonates with the simple illustration given above; the experience of the moment of awe on the top of the mountain. This offers a glimpse of something apparently timeless that *seems* to point beyond the human realm. Whilst this kind of experience can clearly not be taken as proof that there *is* something beyond the human realm, the experience is undeniably real, and on James's account, mystical. Having established our radical freedom to create our own essence and determine our own lives, the intention is not to establish new ways of conceptualising our existence as, for example, part of God, or to establish foundational experiences which will determine our choices, but rather to show that we can *experience* for ourselves what seems non-conceptually real, and is not of our own making. By losing sight of the ineffability of the experience, and attempting to conceptualise it, we necessarily diminish it, we bring it under powers that are limited by our finitude and standing as creatures of lack, and so cut it down to our size. Concepts are the product of human attempts to bring things under our control, by conceptualising something I can manipulate it to my own advantage. But in the type of experience under investigation here, such manipulation results in the disintegration of the very thing under scrutiny.

Sartre himself makes use of mystical language to describe experience, particularly in his fictional work, and it is an open question whether or not his atheism had a mystical foundation<sup>4</sup>. In support of this theory, Gellman cites Sartre's claim to have had a mystical experience of the non-existence of God, and proposes that Sartre was a mystical-atheist, whilst Kirkpatrick, who denies the theory, points out that Sartre later dismissed his own interpretation of that particular experience, saying 'that was obviously untrue ... but it was how I always saw it – a thought that came suddenly, an intuition that rose up and that determined my life' (Kirkpatrick, 2013, p.162). What seems apparent from these two conflicting remarks, both written by Sartre, is that what he dismissed was the *interpretation* of the experience rather than the experience itself. For Sartre, *nothing* determines your life, and his interpretation of the experience as revealing a truth with that potentiality must clearly be untrue for him. As Kirkpatrick argues, Sartre certainly would have objected to *theistic* mysticism on the grounds that 'it imposes categories on experience which are not present in the experience itself (phenomenologically)' (ibid., p.166). This imposition is problematic both for the content and the methodology of mysticism in general when it is presumed to be revealing a truth to be interpreted in conceptual form. Kirkpatrick goes on to write 'for Sartre ... phenomenological

---

<sup>4</sup> For a discussion on this point, see Gellman (2009), and Kirkpatrick (2013).

awareness should be pure consciousness, prior to epistemology' (ibid.). I would argue that the kind of experience which strengthens the claim for a source of significance apparently beyond the human is better understood as phenomenal experience itself, in all its ineffability, recognised as illuminating, or revealing insight into depths of truth, full of significance and importance but without the inevitable distortion arising from an attempt to formulate any kind of understanding of how what appears ineffable might be conceptualised.

Speaking about the concepts that for the early Greeks make up virtue – courage, justice, piety and wisdom – Paul Feyerabend writes:

According to Protagoras ..., one learns these virtues as one learns Greek, without an appeal to specialists, without teachers, simply by growing up in a community that practices them... Concepts of this kind are on the same level as color, swiftness, beauty of motion, expertise in the handling of weapons and words. They are affected by the circumstances in which they arise, by dreams, emotions, wishes; they are not subjected to rigid rules (Feyerabend, 1999, p.270-1).

It is by experience of the practices in which they arise that these concepts are explained. I suggest that in a search for meaning and answerability, many of the concepts used must be learned in the way Feyerabend illustrates here. Conceptual thought will be required in order to give content to the investigation, but this will only be partial and indeterminate, making room for an element of ineffability, and non-conceptual experience.

With this in mind, in the next chapter I will consider metaphor. Investigation into what is unknown can, perhaps, only be attempted by use of metaphorical language – metaphor is a tool uniquely suited to investigating the unconceptualisable. But rather than being solely a way of discovering what there is, I argue that by metaphor we can create something altogether new, expanding the scope of what it makes sense to think:

when a poet carries the mind into a context of meanings and then pitches it past those, one knows that marvellous rapture that comes from going past all categories of definition. Here we sense the function of metaphor that allows us to make a journey we could not otherwise make, past all categories of definition' (Campbell, 2013, p.9).

This will then lead to a discussion of what it means to create a radically new thing and how this might be achieved. By radical creativity we push beyond the boundaries of what exists and what is known. Ultimately, I will argue that in the creative process we can come to *experience* something apparently beyond our human world, and find ourselves answerable to it.

## CHAPTER FIVE: UNDERSTANDING METAPHOR

Metaphors permeate many aspects of human interaction and have many uses. But there is something particularly useful about metaphor when used as a tool for confronting things that are mysterious. In an effort to understand the human condition, we find ourselves asking substantially the same questions that have been asked by humans as far back as we can imagine, and yet in many aspects our lives are unrecognisable from those of our ancestors. This phenomenon implies that whatever meaning life might be thought to have for us, it is unlikely to be communicated in one absolutely fixed and timeless message. In which case we need a means of communication that is essentially fluid. Metaphor seems to be uniquely suitable for the task of bridging the gap between timeless truths and contemporary concerns. Via metaphor we can address the profound mysteries of life in a language that makes sense to a community of people with a shared realm of life experience. Because this endeavour is central to religious traditions, the historical use of metaphor in this context is widespread. However, as Joseph Campbell points out: 'The continuing confusion about the nature and function of metaphor is one of the major obstacles – often placed in our path by organised religions that focus short-sightedly on concrete times and places – to our capacity to experience mystery' (Campbell, 2013, p.8). Campbell's claim is that in religious teaching we have taken literally the central myths that were intended to be understood metaphorically, and in so doing we have fixed what we perceive to be the message in a particular time, geographical place and culture, and have depleted their power to convey spiritual truths: 'It is vital to recall that if you mistake the denotation of the metaphor for its connotation, you completely lose the message that is contained in the symbol' (Campbell, 2013, p.48). What is crucial here is not only the danger of misinterpreting whatever message the metaphor was intended to convey, but of restricting the creative power metaphor has to go beyond what is being asserted. If metaphor is to be significant in the enquiry into what is mysterious, particularly in relation to the meaning and purpose of human existence, the potential for creating something new needs to be taken seriously.

Metaphors can be recognised uncontroversially as offering an indirect way to achieve understanding of phenomena we find hard to comprehend or cannot express directly. However, claims about what is involved in the process of creating and interpreting metaphor, and the nature of what is produced, vary widely between different metaphor theories and linguistic traditions. According to traditional models, which can be traced back to Aristotle, metaphorical meaning is conceived simply as a transformation of antecedent literal meaning in relation to an already present object: language gives pre-existing objects a proper name which



presents the essence of the object; and in the use of a metaphor, an improper name is assigned which has an aesthetic or rhetorical function but adds no real knowledge. So, within the context of religious tradition for example, metaphors have often been thought of as the result of the substitution of figurative language for some literal or 'real' meaning in order to ornament a text or to draw out a doctrinal or ethical idea. In this instance, interpretation might be understood as a kind of de-coding – the figurative meanings that emerge are subject to certain criteria of hermeneutical adequacy and appropriateness, in relation to the substantive and normative claims particular to a specific tradition. This way of understanding metaphor takes it to be an embellishment of discourse, but one that is unable to uncover truth. But there are a number of theories that challenge this conception and attribute a creative capacity to metaphor, claiming that it is possible that the use of metaphor can result in an instance of something that did not previously exist. The two extremes of interpretation of the relationship between metaphorical expressions and their contexts are described by Hausman as reductionism and originativism: 'the reductionist interprets metaphors as cognitive expressions translatable into analogies, similes, or, in general ... language that meets antecedently accepted standards of significance ... the originativist believes that some metaphors ... are creations that constitute the things, qualities, or relationships they signify' (Hausman, 1989, p.24).

### Creative metaphor theories

Creative, or originativist, metaphor theories can be loosely sub-divided into two categories, those that maintain that what metaphors produce is world-constituting, and those that claim it is world-disclosing. If it is the case that metaphor is able to exceed the boundaries of what *is* in an objective sense, reaching beyond facts and objects to redefine the world as a whole, whether it does this by re-constituting the world, in the sense of introducing something wholly new, or by shedding new light on antecedently existing, but unavailable elements, is a matter of importance. Is metaphor a tool for perfecting the principles of expression by means of which we attain the true path to reality? Or does experience become purer to the extent that we strip away language altogether? In many descriptions of metaphor, this is not entirely transparent, for example:

to the extent that metaphor is a movement in language that exceeds and escapes conceptualization, it can have no properly definable object. It signifies ... an indeterminate or indeterminable referentiality, a reference to something infinite, what for Wittgenstein was ... "the mystical ... To the extent that it exceeds proper concepts,

metaphor (at least live metaphor) is about what is, strictly speaking, inconceivable” (Franke, 2000, p.148).

It is the more radical, creative theories that I wish to investigate, in order to establish whether certain kinds of metaphors can in fact be considered as, in some way, genuinely *productive* of meaning, and if so, what kind of innovation this implies. And it will be seen that in the theories under consideration, the dichotomy between world constitution and world disclosure does not maintain; if a metaphor is to be truly innovative it will, in an important sense, both disclose and constitute reality.

Claims have certainly been made which suggest that ambitious expectations might be justified: ‘the rediscovery of metaphor in recent decades by hermeneutic thought has emphasized its dynamic capability of bringing absolutely new possibilities of significance – and even the possibility of sense itself – into existence’ (Franke, 2000, p.144). If mere substitution of words with the aim of achieving a more decorative phrase is at one end of the scale of metaphor theory, the claims made by Hans Blumenberg about what we might expect from metaphor are at the other extreme: ‘Absolute metaphors “answer” the supposedly naive, in principle unanswerable questions whose relevance lies quite simply in the fact that they cannot be brushed aside, since we do not *pose* them ourselves but find them already *posed* in the ground of our existence’ (Blumenberg, 2010, p.14). Blumenberg’s account, which is both unconventional and unsystematic, cautions against naive acceptance of revealed truth, and suggests that through metaphor we might be better equipped to attempt an apprehension of human reality. Whilst remaining sympathetic to the all-encompassing nature of Blumenberg’s account, it will be helpful to consider first some of the more accessible, systematic accounts of what metaphor must be like if it is to be understood as genuinely productive of significance. The work of Max Black gives some guidance, as does the radical approach offered by Paul Ricoeur. An examination of these accounts might prepare the ground for an appreciation of some of Blumenberg’s more elaborate and innovative claims.

At a time when the prevailing view, influenced by the work of Hobbes and Locke, was that metaphor was a frivolous rhetorical device having no direct connection with facts and no capacity to transmit knowledge, Max Black argued that, on the contrary, metaphors have cognitive status. Building on the work of I.A. Richards, Black articulated the concept of interaction metaphor ‘arguing that the similarities into which a metaphor offers insight are created through the novel conjunction of terms, and *they* do not merely describe a set of similarities already resident in the perceptual world’ (Hagberg, 2001, p.373). In the context of

Biblical interpretation for example, David Tracy gives as an example 'the charged and tensive series of metaphors employed by Paul in Romans to disclose the scandal and paradox of Christian understanding' (Tracy, 1978, p.96), and quotes Ricoeur: 'The modern interaction theory of metaphor makes the functioning of the metaphors in the parables far more understandable by noting the inner tensions which ... by the twist of a semantic impertinence become a genuinely informative semantic innovation' (ibid, p.99).

In Black's theory, a semantic innovation is triggered by a metaphor with the transfer of a concept from its established territory to an unfamiliar one. By connection and contrast, a meaning is expressed which could not have emerged by use of literal wording. In the interaction view Black identifies the two distinct subjects of the metaphor as the primary and secondary subjects – in the familiar example of 'man is a wolf', man is the primary subject, wolf is the secondary – with the secondary subject being regarded as a system rather than an individual entity. This system Black terms an implicative complex or a system of commonplaces; described as 'current opinions shared by members of a certain speech-community' (Black, 1993, p.28).

The initiator of the metaphor 'selects, emphasizes, suppresses, and organizes features of the primary subject by applying to it statements isomorphic with the members of the secondary subject's implicative complex' (ibid). The two subjects interact within the context of the particular metaphor, and within the mind of the speaker/hearer in cumulative response: 'the presence of the primary subject incites the hearer to select some of the secondary subject's properties; and invites him to construct a parallel implication-complex that can fit the primary subject; and reciprocally induces parallel changes in the secondary subject' (ibid). Importantly, the interaction is multi-dimensional. In response to the charge that if these opinions and commonplaces already exist, nothing new is generated, Black emphasises that 'a metaphor producer may introduce a novel and nonplatitudinous "implication-complex" (ibid). Black's reference to opinions and commonplaces in connection with both subjects, allows for a wide range of influences to be encompassed, not only word meanings but associations and oral traditions which have built up within specific cultures over many generations. Metaphor then imports much more than is immediately apparent, a long history of associations plays a part in the strength of the conjunction. Consequently, elasticity is apparent in both use and interpretation; 'Because metaphors juxtapose disjoint and complex domains of connotation, their meaning varies with the interpretive finesse of their audience, and so cannot be specified by any interpretive rule' (Pillar, 2001, p.198).

Black assigns a strong cognitive function to some metaphors, making the much-disputed claim that by *changing* the relationship between the primary and secondary subjects, new knowledge and insight can sometimes be generated. A fairly straightforward way metaphors might be thought to achieve this is by creating new perspectives: as we necessarily view the world from one perspective or another rather than a-perspectively, by creating new perspectives metaphors could have the capacity to generate insight. But an investigation into an example provided by Black to illustrate his point suggests that what he considers to be at stake is more than merely the chance of a new perspective.

To demonstrate the idea of the generation of new knowledge, Black introduces the question of whether the slow-motion appearance of a galloping horse existed before the invention of cinematography (Black, 1993, p.37). Black's point is that what is seen in the slow-motion film is an antecedently unavailable, and so wholly new, phenomena. Although necessarily mediated by a man-made instrument, it becomes a part of the world, and in appearing, really comes into existence.

In this example, the *ground*, or *basis* for what is now available to be seen, the slow-motion appearance, was there before, but the appearance itself was not. To enable perception, more was demanded than merely a change of orientation in the prospective perceiver. The slow-motion *appearance* did not exist, it was not *available* to be perceived. It was only made available with the introduction of a new instrument, which removed a limitation in our capacities for perception, and gave us access to something that reality offered, but we were unable to take up. The new appearance, the slow-motion image, is objective in the sense that it will be similarly experienced by anyone who views it under similar conditions, the same information will be conveyed.

And the thought is that metaphors can function as cognitive instruments in a similar way; metaphor can generate new knowledge to the extent that when the composer and his audience engage creatively with it, it is possible that they will attain knowledge that they did not previously have, they gain something new; 'some metaphors enable us to see aspects of reality that the metaphor's production helps to constitute' (Black, 1993, p.38). This statement, together with the slow-motion film illustration, makes it apparent that in Black's account, the dichotomy between the accomplishment of metaphor in constituting reality and its facility for disclosing reality is not distinct. What the world can disclose to us depends on our capacity to see or to think, there are some things that are just not able to be disclosed. A metaphor institutes a paradigm according to which we see what was not visible before, and so our

position as knowers is altered as we overcome a restriction in our view or capacity. A similarity is created which is understood within the context of the implicative complex; the world is given intelligibility by meanings which have not previously been recognised because the metaphor shows up a new system or ordered relation of meanings.

Language that can be expressed literally doesn't exhaust what it makes sense to think, so by using metaphor as an instrument of enquiry, the scope of the questions we are enabled to ask can be expanded. The active shaping of a metaphor creates a new thing that it now makes sense to think. The point that the new knowledge 'makes sense' is critical. For Black, what we build on initially when producing a metaphor does not necessarily have truth-value in the way that statements generally do; in the example 'man is a wolf' it makes no sense to ask whether he *really* is a wolf. The implication of this is a broadening of possibility, we are not restricted to working with what is known to be true about the species wolf, but can draw on what people have associated with both men and wolves, incorporating perspectives, cultural correlations, and wider concerns about motivation, vulnerability and character for example. However, metaphors are not purely fiction, and Black claims that they do tell us something about the way things are. The truth-value they possess is symbolic truth in that the similarity between the primary and secondary subjects is similarity in the way of seeing, or intuition. Metaphors are not subject to the charge of being true or false, but might be described as being accurate or inaccurate, and can be assessed for appropriateness in this regard. So, the suggested correspondences between the primary and secondary subjects are rendered explicit by the metaphor and are then determined appropriate or inappropriate, faithful, partial, and so on: 'Metaphors that survive such critical examination can properly be held to convey, in indispensable fashion insight into the systems to which they refer. In this way, they can, and sometimes do, generate insight about "how things are" in reality' (Black, 1993, p.39). While metaphors themselves are not subject to being true or false, the situations they lead us to have the potential for being so described, the insights created by metaphor cannot be reduced to subjective, unconstrained reflections. Returning to Black's illustration again, the appearance of the slow-motion galloping horse offers us real insight into how horses run, by use of an image that is unrealistic with regard to actual lived experience. The analogies of structure grounding the correspondences are, for Black, partly created, and partly discovered.

According to the interaction theory then, metaphors yield genuine semantic innovation. Metaphors can be said to deepen and enrich our cognitive grasp of the world, they can communicate some kind of understanding or insight and, although without propositional form, can be said to mean something.

An example taken from contemporary literature reinforces the claims made by Black about the potential for metaphor. In Chapter One of John Berger's novel *G.*, the narrator is describing a crisis faced by the English upper class towards the end of the 18<sup>th</sup> century when it became apparent to them that their chosen way of life was becoming incompatible with the modern world: 'They had long since accommodated themselves to industrial capitalism and trade, but they had chosen to continue the way of life of an hereditary, landed elite' (Berger, 1996, p.31). The narrator uses metaphor to convey the idea that when threatened by new requirements with regard to modern finance, industry and imperialist investment, together with demands from the masses for democracy, the upper classes chose to transform their way of life into a theatrical performance: 'If their way of life had to disappear, they would first apotheosize it by openly and shamelessly transforming it into a spectacle' (ibid, p.32). The narrator describes the social life of the upper classes, the Court Balls, the Race Meetings, the Shoots, as the performance of a play on a stage with its own laws and conventions. The main character, Jocelyn, is described as a peripheral member of the class, which only 'increases his need to believe that the play is life and that the rest of life is a suspended empty interval' (ibid). The clothes Jocelyn wears are described as his costume for the part, and the hounds, the fences to be jumped, the winter weather etc. as the scene and the props.

Portraying the way of life of these people as a play enacted on a stage prompts a wide range of novel associations and a distortion in the customary way of thinking. Members of the upper classes during the 1800's are generally depicted as being powerful, and in control of society, with the lower classes being subservient and degraded. Berger's metaphor retains these characteristics, all those not on the stage are suspended in an interval rather than taking part in real life, so by setting the non-performers apart from the performers, the metaphor evokes deeply unequal social roles. But equally, a deep insecurity among the upper class is suggested. If the social life of Balls and Race Meetings are the scenes in the play, the question of how to live in-between scenes, how to live as a non-performer is simultaneously evoked, and aligned with the plight of those living while suspended in an interval. In addition, the metaphor sets up a situation where the members of the upper class are, to an extent, under the control of the masses: 'The general public welcomed the apotheosis. Like most audiences they felt that, to some degree, they owned the performing players' (Berger, 1972, p.32). The cast of the play, the members of the upper class, arouse contempt and ridicule for their bravado but now also sympathy for their identity crisis over what is real life and what is purely a spectacle for the amusement of others. The character of Jocelyn, whose attempt at self-delusion is only partially successful, and who is inadequately qualified even for full membership of his own class,

personifies the quality of pathos. There is a suggestion of justice in the empowerment of the masses, and an intimation of ultimate equality with regard to feelings of insecurity, self-doubt, and so on, shared by all humans regardless of class. And this shift in perception within a given society can arguably engender real and radical change. The vast range of associations and traditions invoked by the inclusion of the concept of theatrical performance greatly enhance what is given by literal description of the characters in the novel and bring us to a wholly new way of seeing.

So, as this example illustrates, Black is justified in his claim that metaphor has the potential to yield genuine semantic innovation and to communicate insight. However, if we understand Black as assigning a purely cognitive role to metaphor, an assumption provoked by his declared determination to counter the claims of Hobbes and Locke who deny it such a role, the implication is that metaphor is still wholly a matter of working with existing linguistic meanings, cultural associations, oral traditions and so on, that pre-date it. Framing one thing in terms of another depends on selecting particular conceptual aspects of the subjects that are relevant, and interpreting a metaphor involves an understanding of what concepts the speaker has in mind and what he is trying to communicate linguistically. A concern with this view is that it seems as if what is produced must be wholly determined by the intellectual input. The suspicion remains that what is really created is a new perspective or perhaps just a new way of carving the world, and so the potential for innovation and creativity is thus limited. Can we demand even more from metaphor?

Two points should be considered before proceeding: Firstly, accounts of creative metaphor such as that of Carl Hausman which seem to offer greater creative potential, highlight a phenomenon which is not clear in Black's account but might have been intended. Hausman insists that what metaphor does cannot be reduced to the shifting of attention from properties conventionally associated with the individual terms, to properties that have been unrecognised or unnoticed previously. If this were the case, there would be no advance in language, but merely a shift within the body of linguistic terms already given and complete. Instead, the individual terms must be transformed into constituents of the whole, and function together internally. As constituents, they are no longer amenable to translation or abstractable from the whole creation. It is the transformation of individual terms into constituents, and their functions as consequent elements that facilitates new meaning.

Secondly, in fairness to Black, it should be acknowledged that dependence on the initiating beliefs of the composer for the structuring and framing-effect does not necessarily restrict the

metaphor's power to go beyond those beliefs. For example Richard Moran claims that 'The composer adopts the perspective he does partly because he expects that it will lead the mind in unanticipated directions. It is possible to get more out of it than one has explicitly put into it' (Moran, 1989, p.109). Moran's claim implies that the mind has resources that it is unaware of, and so doesn't exploit, and that, metaphorically speaking, we don't already know all the places that the mind can go. Perhaps we have the capacity to see more of reality than we recognise. When working with metaphor, the speaker and the audience pay attention to facts about the world, not merely about the language used:

Effective metaphors capture, that is, represent relations of real interest between entities in the world .... metaphors, in a sense, *transcend* language by opening themselves up to aspects of the world that may not initially be embodied in the words themselves, and may perhaps even be unknown to the person who first deploys the metaphor in question (Jollimore, 2009, p.147).

In accordance with this view, Merleau-Ponty describes the power of metaphor as a capacity to go beyond the free initiative of the philosopher:

It would be a language of which he [i.e.: the philosopher] would not be the organizer, words he would not assemble, that would combine through him by virtue of a natural intertwining of their meaning [...] where what counts is no longer the manifest meaning of each word and of each image, but the lateral relations, the kinships that are implicated in their transfers and their exchanges (Merleau-Ponty, 1968, p.125).

## The incorporation of emotion and value

In Merleau-Ponty's account, metaphor is not limited to the intellectual content fed in by the composer and the interpreter. If language is recognised as having the ability to let things speak for themselves, a challenge is posed to the traditional view of truth as a one-to-one relationship of coincidence, in which univocal names are assigned to pre-established objects, and in which proper and figurative senses are sharply distinguished. Merleau-Ponty names the mode of language that functions autonomously as the 'occult trade of the metaphor [which] is able to institute relations that the thinker can grasp only afterwards ... and in a certain sense despite himself' (Vanzago, 2005, p.431). If metaphor is said to transcend language and the powers of the thinker, how then can we understand what is discovered or created, and who, or what, is responsible for the lateral relations, and the implicated kinships mentioned above?



Whilst acknowledging the cognitive role metaphor plays in the influence of perception and interpretation, many theorists recognise, in addition, an emotional and valuational capacity; metaphor incorporates not just thought, but feeling. This allows metaphor to take into account what matters to me, my interests, goals and commitments, and how I relate to the world. Whereas in thought, I survey the world as that-which-is-not-me, my feelings anchor me to reality, when I feel something I have some part in it because it is my projects that are at stake. In addition, with regard to forming appropriate responses to morally charged situations, it is possible that by helping us imagine what the emotional realities might be, and in inciting emotional responses, some uses of metaphor, for example in poetry or drama, can take us further toward a full understanding of complex cases. In the example of poetry, metaphor can be seen to express something of the poet's experience and to evoke an appropriate response in the reader, whilst making suggestions about the nature of reality. The inclusion of feelings and attitudes adds to, rather than replaces, claims that are made about the world.

Perhaps this might lead us to a conception of metaphor as a kind of fiction built out of reality. Kendall Walton suggests understanding metaphor as a kind of make-believe (Walton, 1993). In Walton's account, the metaphoric statement suggests a potential game of make-believe in which the focus is on the prop. In a simple game of make-believe, a child might use the prop of a wooden stick to make it fictional – 'true-in-the-world-of-make-believe' – that he is riding a horse. In this instance, what is interesting is not the wooden stick – indeed this must be ignored for the game to succeed – but rather the interest is in what the child 'discovers' about the world by means of the make-believe. But, Walton claims, there are instances in which the reverse is true, where the make-believe provides insight into the prop. Walton illustrates this idea with the metaphor 'argument is war' in which the argument is the prop, and the state of war, the fiction. What takes place is 'a game in which what people say in the course of an argument generates fictional truths about acts of war' (Walton, 1993, p.45). The verbal elements are war-related, of winning and losing the argument, of arguments being on-target, or shot down. The make-believe is orientated towards the prop, the argument, rather than the fiction, but playing this game can create new understanding of the situation. The participants in the argument might actually imagine they are participating in a war, but all that is required is an awareness of the potential for such a make-believe situation.

Walton's claim that metaphors are instances of 'prop-oriented' make-believe does accommodate the inclusion of imagination and feeling, but it also creates a division within metaphor that, I would argue, diminishes its true potential. In the 'argument is war' example, Walton rightly argues that our real interest is not on war – the make-believe, but on a better

understanding of the argument – the prop. But if this is our intention, the metaphor is merely an instrument into which we feed information in order to get new insight or a new perspective. The view of the argument as a prop in the fiction of war might indeed highlight previously unnoticed aspects of our behaviour or that of others. But this new understanding is well within our powers of imagination, and there are other routes that we might follow to reach the same end. Depending on how we define metaphor, there are many quite mundane examples that regularly occur in everyday conversation, metaphors are often used merely to show things in a new light or offer a new perspective. But if we are looking for something unique from metaphor, something that cannot be achieved by any other means, I suggest that we must focus not on either prop or content, but on the way all the elements combine and work together as a whole, and for this holistic approach, the metaphor theory of Paul Ricoeur is illuminating.

Ricoeur's understanding of metaphor as having the capacity to 'provide untranslatable information and, accordingly ... to yield some true insight about reality' (Ricoeur, 1978, p.143) involves the inclusion of 'a psychological moment of the kind usually described as "image" or "feeling"' (ibid). Ricoeur claims that the informative core of a metaphor is necessarily accompanied by factors extrinsic to it, arguing towards a semantic role for imagination and feeling. For Ricoeur, these elements were lacking in the accounts of Richards and Black who appeared to focus essentially on the cognitive aspect. Whether or not this charge is justified, by arguing for the inclusion of imagination and feeling as essential components of metaphor, Ricoeur made explicit what I would suggest was implicit in the account of the implication-complex given by Black.

Believing that humans are essentially creative in their use of language, Ricoeur considered that metaphors 'create new meaning and thus play an important epistemological role... metaphors create new conceptual systems that enrich our understanding of the world' (paraphrased in Snvarr, 2010, p.57-8). In what Ricoeur calls a living metaphor, what is rearranged is not the aspects or properties that are ascribed to a certain object, but instead the conditions of visibility of the object itself. Something is salient for us not just because of its features, but with regard to how we feel about it, how it affects us. A central claim in Ricoeur's argument is that the functioning of metaphorical sense challenges the dichotomy between sense and representation – sense understood here as the objective content of an expression, and representation as the images and feeling that constitute its mental actualisation. Rather than ascribing the potential for a semantic role to the objective content alone, Ricoeur suggests that in metaphor, imagination and feeling equally have such a role. This idea seems to resonate with

the suggestion in Chapter Three that our capacity to assess moral significance is heightened by the inclusion of affect and emotion in conjunction with rationality. In both instances, feelings can be seen to be both significant, and generative.

The pictorial dimension associated with representation is suggested in the expression 'figure of speech', which is highlighted by Aristotle who claimed that 'the vividness of such good metaphors consists in their ability to "set before the eyes" the sense that they display' (paraphrased in Ricoeur, 1978, p.144). The picturing function of metaphoric meaning, the capacity to turn the message into a figure, makes discourse *appear* in a particular way. Things that resist conceptualisation are given sense by being *shown* within the working of the whole metaphor.

The semantic role ascribed to imagination is underpinned by Ricoeur's understanding of imagination as productive; we need imagination in order to both see and produce likeness. The role of the productive imagination is understood as facilitating the insight necessary for the reconciliation of divergent elements, the shift in logical distance, which Ricoeur terms predicative assimilation. The imagination is productive to the extent that the assimilation *makes* the similarity between the subjects of the metaphoric utterance. Predicative assimilation is paradoxical in character in that the incongruence is preserved;

In order that a metaphor obtains, one must continue to identify the previous incompatibility through the new compatibility ... Imagination, accordingly, is this ability to produce new kinds by assimilation and to produce them not above the differences, as in the concept, but in spite of and through the differences (Ricoeur, 1978, p.148).

Crucially, in imagination, Ricoeur does not intend that what is imagined is a mental picture of something absent, or a representation of what has already been perceived, but rather the image is understood as 'a being pertaining to language ... The image is an emerging meaning' (Ricoeur, 1978, p.149). Reflecting the way in which a new semantic pertinence emerges from the suspension of literal sense due to semantic absurdity, the suspension of ordinary descriptive reference in favour of imaginative interpretation invites a more radical way of looking at things. However, the total mental chaos, or incoherence, this might suggest is avoided by what Ricoeur refers to as split-reference, metaphorical reference maintains ordinary vision *and* the new way of looking at things at the same time; 'the perspective prior to and subsequent to the transformation of the metaphor's principle and subsidiary subjects must both be conjointly maintained' – imagination 'contributes concretely to the *epoché* of ordinary reference and to the *projection* of new possibilities of re-describing the world' (Ricoeur, 1978,

p.154). With the split-reference view, metaphor forces us to do more than one thing at a time, and so enhances our capacities. The metaphor then is not a phase during the process of reaching a new set of circumstances, the metaphor and the interpretative work must stay together, not as component parts but bound by internal tension, as somehow complete, but never resolved or finished.

Referring back to the traditional understanding of metaphor as deviant, Ricoeur reintroduces the idea of deviance in relation to the structure of the whole metaphoric sentence, speaking of the semantic impertinence caused by a violation of the code of relevance governing the ascription of predicates in common use. In a standard sentence structure, 'x has the property y' for example, the subject and predicate can be assessed separately and the assimilation leads you to what you should think, the meaning is fixed. But in metaphor, the production of sense is carried by the complete utterance as a whole, not as a harmonious organic form, but as a structure characterised by internal tension, incorporating subjectivity and consciousness, thought and experience into linguistic play. In the transition from 'literal incongruence to metaphorical congruence between two semantic fields ... the new pertinence or congruence proper to a meaningful metaphoric utterance proceeds from the kind of semantic proximity which suddenly obtains between terms in spite of their distance' (Ricoeur, 1978, p.147). Ricoeur's term 'semantic proximity', itself a metaphor, demonstrates the incorporation of an experiential aspect, as the proximity is experienced but recognised as not literally true. The transfer of meaning that occurs with the metaphor is purely this shift in the cognitive distance from far to near revealing a 'generic kinship between heterogeneous ideas' (ibid). Metaphors are then no longer determined by any antecedent theory of meaning: 'Meaning is essentially a determination by shifts in application of words in ever new contexts' (Franke, 2000, p.150).

Meaning, in Ricoeur's account, is not a fixed and separate entity, but rather an event; metaphorical meaning is centred on the context and the tension created within it. Stefan Snvarr describes Ricoeur's theory of metaphor as: 'an integral part of a panoramic vision of the human world in which metaphors, narratives, religious language, imaginative literature, imagination, and even actions form a great chain of meaningful being' (Snvarr, 2010, p.55). From this can be seen an intention to relocate metaphor away from linguistic theory, and to focus not only on what metaphors mean, but on what kind of striving they are part of, indicating, as suggested above, that Ricoeur's approach is indeed a holistic endeavour. Metaphor is not only the semantic clash and resultant collapse of literal meaning, but necessarily includes the new interpretation which renders the metaphor acceptable: 'The metaphor is not the enigma, but the solution of the enigma' (Ricoeur, 1978, p.146).

As stated above, Ricoeur ascribes a semantic role not only to imagination, but also to feeling. Feelings are not purely inner states but interiorized thoughts accompanying the work of imagination in a 'picturing' relationship: 'To *feel*, in the emotional sense of the word, is to make *ours* what has been put at a distance by thought in its objectifying phase' (Ricoeur, 1978, p.156). The feelings inspired by poetry, for example, are not literally felt as bodily emotions, they follow the split-reference form assigned to imagination – tragic terror and tragic pity both deny and transfigure the literal feelings of fear and compassion. In poetic feeling we claim ownership of thought by abolishing the distance that thinking generates between knower and known, whilst maintaining the cognitive structure of thought and the intentional distance implied. Metaphor teaches us to see things in this split-reference form, it enables our capacity to be close, to be *in* the tragic story for example, whilst also seeing from a distance. Our perspective on human reality is limited, but metaphor alters this by using what we *can* see to take us to somewhere new; we realise that our normal mode of reacting to the stuff of reality isn't the only way there is. Again this complicates the dichotomy between disclosing and constituting reality, metaphors can give us access to a bigger context for feeling and enrichment than literal descriptions can offer, I experience and feel a new thing. Ricoeur argues for a '*structural analogy* between the cognitive, the imaginative, and the emotional components of the complete metaphorical act ... the metaphorical process draws its concreteness and its completeness from this structural analogy and this complementary functioning' (Ricoeur, 1978, p.159). Ricoeur has argued that 'the employment of metaphor can liberate us from the referentially static or fixed conception of the world that literal language is thought to enforce, allowing not only a changed metaphor-induced perception of the world, but more extensively, a changed ontology, indeed a changed world' (Hagberg, 2001, p.376).

### The dynamic capability of metaphor

Metaphor is then no longer necessarily a derivation from some meaning that pre-dates it, but can be original. The dynamic capability of metaphor to bring about new possibilities of significance brings into question the inflexibility of previously established literal meanings, and invites recognition of language as intrinsically dynamic, language as event. Metaphor is not a reaction to some 'proper' meaning, but original and absolute, asserting itself 'as an autonomous and irreducible dimension of language and thought' (Franke, 2000, p.146). The actuality of proper, or real, meaning is not denied in the use of metaphor, but it is acknowledged as unable to be independently identified and set apart to act as a foundation in

terms of an absolute frame of reference. What might be considered absolute in metaphor is the whole meaning, which can only be achieved by the creation and interpretation of the metaphor, and to which the literal is secondary. If we want to make an attempt to express and make sense of that which actually resists conceptualisation, experiences that have a mystical quality for example, those that occur and occupy the mind as real but which are not wholly translatable by literal truth claims, then the suggestion of an absolute nature to metaphor is interesting. Which leads us back to the metaphorology of Hans Blumenberg;

A metaphor is absolute if and only if it cannot be dissolved in concepts, and it must also have a content that gives direction for modes of relating to the world. Such a metaphor structures the world and represents the totality of reality, which cannot be experienced and of which we cannot ever get a full view (Snavarr, 2010, p.94).

For Blumenberg, absolute metaphors reveal 'the substance of thought, assuming thereby a kind of resistance to conceptual explanation, which can never exhaust them' (Franke, 2000, p.146), and they occur when what is referred to metaphorically resists thorough intelligibility, and therefore cannot be identified by any other proper terms. Absolute metaphors reflect reality in their capacity to resist and exceed our grasp.

I suggested earlier that the dichotomy between world constitution and world disclosure does not maintain in creative metaphor use, and that if a metaphor is to be truly innovative it will, in an important sense, both disclose and constitute reality. I will return to this point now. How we deal with the question of referential function in discussions of metaphor theory, whether we see the value of metaphor as resting in the capacity to reveal new truths, or to create out of nothing, will almost certainly be influenced by how we respond to the much bigger question alluded to in chapter four; whether or not we consider that there is 'a way the world anyway is', independent of human perspectives – broadly, whether an idealist or realist attitude is adopted. What does it say about our conception of reality if, by our action, we are able, in some way, to alter it?

The direction that we have been following in relation to the overall aim of this project, with its concern for uncovering truths about reality, and the potential for finding meaning and purpose, certainly seems to demand a realist interpretation of what there is. If what we seek ultimately is connection to that which seems to reach beyond the human realm, we are not looking to 'create' reality merely by perfecting principles of expression, but to *approach*, as nearly as we can, that which exceeds our conceptual grasp. On this understanding, metaphors are produced in the mind by the creative imagination, but they reach out to, and interact with, what is there

already. A realist interpretation is called for, but this needs to be qualified. As metaphors are clearly not literal descriptions of reality, and as they possess the quality of being continually eligible for revision and reinterpretation, a naïve realist approach will not be adequate. The very nature of, and need for, the metaphoric process rules out this kind of realism. If pure experience existed, as available to be accessed by a process of stripping away language, metaphor would have no role. Rather, the kind of metaphor theory suited to the project of conveying what might be thought of as mystical truth, namely the theories of creative or absolute metaphor, point towards an acceptance of a qualified, or critical, realism: metaphor does tell us something about the world, but this is a reality which *includes* the necessary subjective element of perspective and bias, the work of our imagination and our emotional responses: ‘the critical realist recognizes the role of the imagination and the place of symbols – the subjective component – in knowledge ... For our knowledge is not acquired apart from the social, economic and political conditions in which we work at it’ (Avis, 1999, p.131). Metaphors tell us something, not only about the world, but about ourselves, and our way of being part of reality. Our fundamental thoughts and our conceptual ability cannot adequately describe how things really are because they too form a part of that reality. In their inclusive and open-ended nature, metaphors can offer us a creative and constructive way to contribute to, and try to make sense of, the situation in which we find ourselves.

The capacity of metaphor to contribute to this end is demonstrated in an existing work of art,



See figure 1, page 173

which can be seen in the British Museum: *The Papyrus of Ani*. This papyrus manuscript was created c. 1250 BCE, in the 19th dynasty of the New Kingdom of ancient Egypt. The work depicts a hall of judgment, dominated by huge black scales in the centre. The scale pan on the left holds the heart of the scribe Ani, which is to be weighed against the pan on the right, which holds a feather representing *Maat*, designating the Egyptian

conception of the law of cosmic order. According to S. G. F. Brandon, the concept of *Maat* ‘connotes not only cosmic order but social order also, so that the term can be legitimately translated as “truth,” or “justice,” or even “righteousness”’. In its social extension, accordingly, *Maat* represented a criterion or standard for personal character and conduct’ (Brandon, 1969, p.92). Ani and his wife are watching anxiously. The mortuary god Anubis observes how the scales fall, and the scribe-god Thoth will record the verdict. The crocodile headed monster waits to consume the heart of Ani if he fails to pass the test. The picture also shows Ani to have been vindicated, and so conducted into the presence of Osiris, the god of the dead. So within this

one static image, a whole journey unfolds, and the experience of suspense and uncertainty is presented, somehow portraying what-is-about-to-happen rather than depicting what-is.

This picture can be read as an example of a complex metaphor, rich with symbolism and cultural references, which, by the standards set by Blumenberg, could justifiably be termed an 'absolute' metaphor. But, to what extent does the newly generated thought disclose or contribute to reality?

*The Papyrus of Ani* presents a version of a motif that reappears in ancient Greek and Hebrew religion, and also in medieval Christianity, often referred to as the weighing of the soul. This version can be found in the *Book of the Dead*, in which it is one of two conceptions of the judgment after death. In many cases, including this one, what is weighed is not actually the soul but the heart, but the significance is in the representation of the moral worth of an individual human being. If we pare this particular instance of the metaphor down to one core message, we are faced with the existence of an impartial, objective standard that all human lives will be measured against. This standard is one that we can have no way of comprehending as it is of the cosmic order, and so ultimately we have no way of knowing, or ensuring, that we will ultimately find favour. There are no enticements we might make to the authorities, no indulgences to be bought, no strings we might pull to influence the way things will turn out for us after death. The challenge to the standards of judgment that are normally accepted is underlined by the measuring of the heart against a feather, which would usually represent weightlessness. The suggestion is that judgment after death is immune from external interference, and commonly accepted standards are not the only ones that pertain.

The metaphor appears to be the result of the combination of a multitude of pre-existing ideas and practices, drawn together and interwoven. Brandon writes 'The concept of weighing first appears in the *Coffin Texts*, which document Egyptian mortuary belief during the Middle Kingdom (2160-1580 BC). References to the idea are imprecise and suggest that the notion was still at a formative stage' (Brandon, 1969, p.97). With the introduction of the scales as an instrument of measurement, this metaphor provides a new way of looking at the idea of weighing in relation to human worth. The scales represent an impartial inviolable process of evaluation. Someone facing a pictorial representation of this idea for the first time, and taking up the message conveyed, might well have an experience that could be considered life changing.

The open-ended nature of metaphor must be emphasized here. What is conveyed by this image



is not, for example, the fixed 'truth' that there *is* an objective standard imposed by a cosmic order. The context in which a metaphor is received must be taken into account. At its inception, the truth conveyed might be the idea that assessment of individual worth is ultimately impartial, leading to the questioning of the need to justify one's life to the authorities, or to denial of the possibility of buying eternal life. Perhaps in a modern context, the truth to be conveyed would be along the lines of the importance of taking ultimate responsibility for one's own life and ethical choices.

So it seems that use of metaphor has the potential to innovate, to the extent that metaphors constitute something not antecedently available for interpretation, but also to facilitate discovery, in that they articulate meanings that are not random, but constrained by something that is independent of them. How human creativity can perform this dual function will be considered in the next chapter.

With its inimitable power to impact on our imagination and consciousness metaphor transcends the power of language and points beyond human understanding, suggesting that there is more than the total of what we are and what we can know: 'An absolute metaphor is one that imposes itself with the force of a revelation' (Vanzago, 2005, p.437). In so doing, metaphor signifies that as humans we are creatures of lack, we are not self-sufficient. Blumenberg saw absolute metaphor as response to the human condition, as a 'need to give a face to what is faceless, the exigency to cope with man's inability to be fully integrated in the world' (Vanzago, 2005, p.436). Having no 'proper' meaning, metaphors act as a supplement for what is absent. But, for Blumenberg, the absence is not the determinate absence of something that might have been present, but the making visible of that which never was, or could have been present, the infinite and inexhaustible aspect of that which is concealed beneath the metaphor.

Blumenberg's recognition of man's inability to be fully integrated in the world reflects the sense of lack as part of the human condition referred to by Sartre. If Blumenberg considered absolute metaphor as an appropriate response to this condition, could metaphor equally respond to the Sartrean experiential lack and unrealisable goal, allowing us to reach beyond our human limitations and our desire for self-sufficient necessity? The problem of the call for self-sufficiency is central to Karsten Harries' consideration of metaphor in poetry. In this case Harries critiques self-sufficiency in the aesthetic object, but there are, I think, parallels to be drawn between man's demand for his own self-sufficiency and the demand for self-sufficiency in the aesthetic object. Harries claims that an ontological understanding of poetic metaphor

challenges the frequent demand that the work of art be recognised as a self-sufficient unity. This demand for unity in a successful aesthetic object requires an acceptance that nothing inessential is included and that to add anything would weaken the aesthetic whole. The result of this autonomy in the artwork is the separation of art from life. A poet meeting this demand would offer 'a self-sufficient presence strong enough to lead us to accept it instead of referring us to something beyond the poem' (Harries, 1979, p.74). But Harries claims that this attempt at aesthetic purity is bound to fail – poetry is necessarily immersed in common language and in temporality. By contrast, taking a more ontological approach to poetry suggests a very different role 'a poem should reveal what matters and thus help the individual to determine what his place in the world is to be' (Harries, 1979, p.86). What poetic metaphor can do is to reveal the inadequacy of the ways of speaking that we take for granted: 'what metaphor names may transcend human understanding so that our language cannot capture it ... in such a view metaphor has to open the work of art to a dimension that transcends it, thus it destroys our experience of the work of art as a self-sufficient whole' (ibid, p72). Harries goes on to suggest that the 'disintegration of language does not lead to silence. Against the background of silence the presence of things manifests itself. As language falls apart, contact with being is re-established' (ibid, p.88). As with the aesthetic object, humans are necessarily immersed in a world. If we limit our horizon purely to what we can think and know as humans, without acknowledging even the potential of there being anything outside that realm, we might only ever see a limited part. If there was a human totality that, even potentially, was fully comprehensible, we ought to be able to articulate it literally. Equally, if metaphor was just the random ascription of subjective meaning, resulting in incomprehensible chatter, perhaps we might conclude that reality was too distant for us to attain any awareness of. But in metaphor something distinctive happens. Joseph Campbell describes metaphor as that which gives vitality to a mythology 'delivering, not simply the idea, but a sense of actual participation in a realisation of transcendence, infinity and abundance' (Campbell, 2002, p.xx). Metaphor demands agency, it is not something to be passively received, but a process to be actively engaged in.

So, why might metaphor be uniquely suited to addressing the profound mysteries of life? The answer is suggested in the word itself, *Meta*: from Greek – a passing over, or going from one place to another, and *phorein*: to move or to carry. 'Metaphors carry us from one place to another, they enable us to cross boundaries that would otherwise be closed to us' (Campbell, 2013, p.xvi). Questions arise within metaphor that would not appear if the boundaries, for example between the known and unknown, were not approached. One important feature that can be observed from this discussion is the manifold of ideas, the sheer quantity of information

in myriad forms, which can be conveyed by metaphor in a single utterance. The enormity of possibility is seen clearly in Blumenberg's essay *Light as a Metaphor for Truth*, in which he looks at the vast range of ways in which light can be metaphorically understood and experienced:

In their expressive power and subtle capacity to change, metaphors of light are incomparable ... Light can be a directed beam, a guiding beacon in the dark, an advancing dethronement of darkness ... but also a dazzling superabundance, as well as an indefinite, omnipresent brightness containing all (Blumenberg, 1993, p.31).

Later in the essay Blumenberg describes how light becomes the metaphor for salvation, for immortality (ibid, p.34). Essentially open-ended in its capacity for being interpreted, and indefinable to the extent that it transcends conceptualisation by incorporating the realms of imagination and feeling, metaphor reaches beyond linguistic theory into the arena of experience:

the coincidence of seeing and not-seeing found in the dazzling effect of pure light is the fundamental confirming experience of all mysticism, in which the presence of the absolute attests itself, in which all thinking and speaking is surpassed, and which represents the uniquely adequate way of encountering transcendence (Blumenberg, 1993, p.45).

If metaphor encompasses not only new conceptions and perspectives, but ways of imagining and feeling, and ultimately new experience itself, it has a capacity which uniquely belongs to it. This capacity is by nature open-ended. Metaphoric meaning can be made available to an individual, but equally it can be passed from one human to another, it has the power to make sense on a collective level. The limits of metaphor are not just my limits, it is not a private and subjective enquiry but a collective adventure, greater than any one of us, as illustrated in Black's conception of associated commonplaces, and demonstrated in the chapter of John Berger's novel *G.* discussed above. And the effects of metaphor can be transmitted with great efficiency, sometimes compressed into a single phrase or sentence. For these reasons it looks like the ideal vehicle for conveying the necessary magnitude of material about what is ineffable, which we would need in order to do justice to the enquiry into mystical experience.

Not all metaphors are creative or particularly interesting, and we don't always use metaphor in the ways explored here. Some metaphor use sets a particular agenda with the intention of shutting down further debate. Metaphors can cause offence, either intentionally or due to carelessness in their use or creation. But this account gives an indication of the potential for metaphor to meet the need to describe and explore experience that is individual and relevant

to us personally, but by using terms in common language. Metaphor can be interpreted subjectively, but it is grounded in objectivity, possessing symbolic truth-value and thus avoiding descent into relativism, enabling it to play a part in the joint human endeavour. By facilitating this exploration into our own experiences and imaginings, metaphor has the capacity both to disclose and to constitute the way the world is, making accessible things that were previously unavailable to us. The ambiguity involved in metaphor allows us to test what the world is, or might be like – one interpretation does not negate another being held at the same time. In the imaginative mode I can relate to something in an experimental way that is not exhausted by content and this feature can play a unique role in providing experiences that cannot be given by things that behave in an understandable way: metaphor ‘suggests, reveals, unconceals ... the deep structures of reality to which we are related as mortals who are born into this world and who *dwell* in it for a while’ (Ricoeur, 1978, p.153).

## CHAPTER SIX: RADICAL NOVELTY

The discussion on metaphor so far has suggested that the products of some creative metaphors can be recognised as new entities, the referent of a creative metaphor can be understood as being, in some sense, both created and discovered. The suggestion is that this newness is not just newness of perspective but something unprecedented. But a close look at this idea suggests some complex questions about human potential: if created, how do new objects, events or ideas issue from human activities? A further question arises from Carl Hausman's claim in *Metaphor and Art*: 'A metaphorical expression functions so that it creates its significance, thus providing new insight, through designating a unique, extralinguistic and extraconceptual referent that had no place in the intelligible world before the metaphor was articulated' (Hausman, 1989, p.94). If metaphors designate a referent that is unique and extra-conceptual, the referent must be new and individual, but at the same time, if it is extra-conceptual, it must relate in some sense to the way the world is, or at least appears to be. How could we have a genuinely new experience, and yet that experience be anchored by the possibilities that the world really does afford, thus making it intelligible?

In this chapter I will develop a view of radical newness that can capture the kind of unprecedented novelty which I claim can result from creativity. It will be seen that these questions, which arise from discussions of metaphor, are not specific to metaphor, but directly relate to the creation of new objects in general. So, I will first consider these points in relation to the creation of novelty generally, before returning to look at metaphor as a model of the creative process.

### The emergence of radically new objects

My starting point is the simple acknowledgement that there are at least appearances of radical novelty in humanly created objects, and in this phenomenological approach I echo Carl Hausman, whose account I will discuss below. Some instances of creativity appear to exemplify radical difference that interrupts the regularity expected in an intelligible world, and it is this radical, spontaneous difference which marks these instances out as interesting, and demanding investigation. Spontaneity, understood in this sense, will then be a central component in my account of radical novelty. But rather than just *appearing* to be radically new, I will argue that these radically new objects must be considered as potentially making an ontological change by instantiating the introduction of new actualities into existence. The basis for this claim is that if a

genuine appearance cannot be dismissed as illusory, and cannot be explained rationally as proceeding from the existing order of things, the possibility of new ontological significance remains open. My intention is not to attempt to offer compelling proof that instances of radical novelty do indeed arise. But rather I argue that negative accounts offered to explain away the appearance of such instances of spontaneity by casting them as consistent with natural processes and rationalistic explanation, have not been successful. These arguments fail to do justice to the phenomena of creativity, and its role in the way we make sense of, and contribute to, the world. In the absence of a compelling negative argument, I suggest that there is a need for a positive account that does justice to our experience of radical discontinuity and allows that it can make a legitimate claim on our ontology. If appearances of radical novelty in humanly created objects are to be considered as having the potential for ontological agency, we need to establish what it means for something to be radically new, how we recognise it as such, and whether a radically new created object can achieve intelligibility in spite of its disconnection from that which precedes it.



See figure 2, page 174

The claim that apparently unprecedented objects come into existence is, I think, uncontroversial, and I assume general agreement on the fact that instances of creativity contribute to advances in tradition, whether in art, science, evolutionary theory etc., leading to new developments in the world, and new things that it makes sense to think. But in this general agreement, the philosophical difficulty of what this commits us to, as indicated in the questions raised above, is often overlooked, and it is acknowledgement of this that motivates my conviction that there is a need for an account of radical novelty.

To offer an example of what might constitute an instance of radical novelty, I suggest the production of the twelve-tone system, the new formal composing technique introduced by Schoenberg. It is claimed that Schoenberg created this technique in order to express historical violence through aesthetic form (Kaushall, 2018, p.15). Another example can be seen in the *Lion Man*, a sculpture that is 40,000 years old, found in a cave in Germany in 1939. This sculpture is the first known representation of a being that does not exist in physical form, but symbolises ideas about the supernatural. The suggestion is that the figure, with the head of a cave lion and a partly human body, carved from mammoth ivory, was intended to help people to come to terms with their place in nature on a deeper, religious level or in some way to transcend or reshape it, and so exemplifies the first known instance of humans thinking beyond the subjectivity of their own existence. The wear on his body caused by handling suggests that *Lion Man* was passed around, and formed part of a narrative or ritual that would explain his

appearance and meaning. Significantly, *Lion Man* is the oldest known representation of something that cannot actually exist. As a new kind of object with an unfamiliar identity which displayed intelligibility even though disconnected from its past, a past assumed to include only that which could be observed, *Lion Man* presents just one example of an occurrence of radical novelty with the potential to transform the world as previously understood.

Creativity has a major part to play in the way we make sense of the world, and the way we expand the horizon of things we can think and are capable of. The world is different after an appearance of a radically new thing, and creativity forces us, by causing surprise and perhaps a sense of wonder, to acknowledge this difference. Within creativity there are clearly different kinds, which display different degrees of novelty. For example, Margaret Boden highlights a distinction between 'psychological' creativity and 'historical' creativity (Boden, 2004, p.43). Psychological creativity results in a surprising, valuable idea which is new to the person who thinks of it, regardless of whether or not another person has already had the same idea. In 'historical' creativity, the idea, as far as we know, has never been held before, it has arisen for the first time in human history. Clearly the degree of creativity which needs to be analysed depends on the area of interest; in order to understand the psychology of creativity for example, the two forms outlined here will be of equal relevance. Indeed, much of what can be said about creativity will be common to both forms, bearing in mind that the intention of the artist is unlikely to be directed towards achieving a particular scale of novelty. However, whilst recognising the value of all forms of creativity, my interest at this point relates to the more contentious question of the potential for ontological change; the appearance of something unprecedented and unpredictable, the actuality of radical novelty. My focus then relates most closely to Boden's 'historical' creativity, but my central interest differs from hers to the extent that I will directly address an aspect which she does not; the ontological potency of creativity.

In attempts to explain creativity, two traditional and opposing perspectives particularly stand out, one is a rationalistic account, and one is non-rationalistic. Whilst obviously not exhaustive, a brief overview of these general positions might be useful to illustrate two aspects that are often either neglected or denied in accounts of creativity. The first aspect is the responsibility of the agent as primary source of his creation and the difficulty presented for rational explanation if this attribution is accepted. Hausman frames this as a paradox inherent in the process of creativity; 'the creator must act as an agent which is a cause without a prior cause, a cause which causes itself' (Hausman, 1975, p.11). The second is the provision of a coherent account of the phenomena of spontaneity which does not undermine the possibility of it actually

occurring; an account of what the world would have to be like in order to meet the conditions necessary for an instance of radical novelty.

### Accounts of creativity

In the first account, a rationalistic perspective, an attempt is made to explain creativity in much the same way as any other thing or event might be explained: by finding a source, and identifying antecedent conditions by which the thing or event might be predicted. Rationalistic explanations may take many forms, perhaps the most prevalent being naturalism which broadly conceives creative acts as natural events, explainable as components of natural processes. This interpretation derives most directly from Aristotle, according to whom creativity consists in making an object that exhibits a new form that has been imposed on pre-existing material (based on Aristotle 1034a, p.195). The form is like a pre-conceived plan that sets out the rules for making the product, and, in accordance with which, the material is shaped. This view takes the essential features of creativity to be both rational, and potentially predictable, being brought about in accordance with a predetermined logic.

In the second, non-rationalistic account, creativity is understood to elude ordinary rational principles, and so is viewed as a mystery for finite human knowledge. This interpretation derives mainly from Plato, who proposed the view, as stated in the *Ion*, that the artist works by inspiration; ‘all our good epic poets speak all their beautiful poems, not through art, but because they’re inspired and possessed’ (Plato, 1996, 533e, p13). The poet is a spokesman for a source that transcends his own conscious effort. Whilst finite humans would be unable to predict creative outcomes on this view, prediction is conceivable from the perspective of a transcendent being, and a complete explanation would be possible given access to this source.

Both these approaches neglect the human agent as primary source of the creative act. And although both accounts allow that new and intelligible kinds of things *appear* to arise, neither of them allows for the actual occurrence of spontaneity which distinguishes radical novelty from creativity more loosely understood. The rationalist claims that with adequate time and knowledge, the new appearance would be shown to be merely surprising, and ultimately could be exhaustively subjected to conceptual understanding. Alternatively, for the non-rationalist, the apparent discontinuities are by nature unintelligible and mysterious to the finite existent, but potentially explicable from the viewpoint of a transcendent being. If a complete, rational



explanation of the creative process is available even potentially, nothing is in theory unpredictable and nothing is unprecedented.

We can now set out the central problem to be addressed in this chapter concerning radical novelty and the potential for intelligibility in newly created objects: in accounting for appearances of radical novelty we necessarily attempt to offer a rational explanation. The moment we do this, we are at risk of reducing appearances of spontaneity to the status of illusion; the very act of describing how something might come about involves making links with what went before and knowledge we already have. And it is those links which cannot maintain in an instance of something that is unprecedented and unpredictable. However, the alternative seems to be to declare spontaneity to be wholly mysterious and inaccessible to human enquiry, requiring the existence of a transcendent source. Can anything meaningful be established about the occurrence of radical novelty if we don't accept the inevitability of one or other of these two alternatives?

### **The intelligibility of radically new objects**

Before considering whether this can be achieved, it will be helpful to outline what I intend by the term object. An object in this sense can clearly mean a physical object like a sculpture, a chair or a fork, but equally a mental object expressed as an image or an idea, a novel, or a symphony, or a firework display for example. Objects are understood as coming into being in time, and can be opposed to abstract, eternal entities. Essentially, an object in this sense is an entity that has a level of endurance, it can reappear throughout separate occasions of experience: 'a realm of entities which transcend that immediate occasion in that they have analogous or different connections with other occasions of experience' (Whitehead, 1967, p.158). Whilst objects can be understood as bearers of properties, and so things we can think about and re-identify, they are not necessarily well-mapped by concepts, as should be acknowledged in an instance of something which appears as a radically new kind of thing.

What then can we say about radically new objects? I will elucidate some further terms in order to answer this question. As I intend the terms, a radically new object must at least instance irreducible and unprecedented difference. However, while this condition is necessary, it is not sufficient, because it could allow for all things equally to be described as novel: with respect to their singularity all discriminable things, every event or object can be considered new. For example, in an account of possibility based on Whitehead, who also argues for the actuality of

radical novelty, Alex Haitos writes: 'Every new moment of existence is radically novel – an extension of the past, mixed with new and unforeseen flavors' (Haitos 2010, p.12). In Haitos's account, novelty arises out of the infinite realm of ideal entities or possibilities, consisting of new combinations or associations. It is important to distinguish the sense of radical novelty I intend from such accounts. If every moment is described as radically novel, there is a sense in which the term is unsatisfying. For example, all new things are not new in the same way; a new coat does not have the significance of a new scientific discovery. Recognising that the term difference can be ambiguous, in this account I demand the highest degree of difference, a difference which comes about by a radical leap and which resists traceability. If the new thing cannot be totally explained by what went before, I also maintain that an element of spontaneity is necessarily present, and that this should be understood as a condition of radical novelty. In order to be radically new in an interesting sense, a created thing, whether a physical object or mental product such as an image or an idea, must instance difference in kind. Something that is radically new exhibits newness of intelligible character: neither the object, nor the kind of thing it is, existed previously. An additional condition then for radical novelty is unpredictability: a radically new object must appear in a reality that was not prepared for it. Spontaneity is by nature unpredictable, and it is this unpredictability that alerts us, as enquirers, to what is radically new. So, to summarise, a radically new object must instance irreducible and unprecedented difference which is brought about by a spontaneous and radical break from a traceable past. As a consequence of being unprecedented, its appearance could not have been predicted. Spontaneity is at the core of what the determinist denies but, whilst contentious, it seems phenomenally true and, I argue, must be taken seriously.

We identify an object as a new kind of thing by contrasting it with what we already know. Radical novelty in painting might occur in an object whose definite identity, or kind, initiates a new style or tradition: for example, divisionism, founded by George Seurat around 1884, introduced the idea of combining colours optically as opposed to blending pigments physically, in order to create high levels of luminosity. In music, as referred to above, Schoenberg's twelve-tone system instanced a new composing technique which affected the course of musical composition as a whole, following on from its introduction. Importantly then, the kind of newness under consideration is that which is attributed to instances of creativity that contribute to new advances within a tradition, or even new traditions. These advances will be seen, perhaps only retrospectively, to have been assumed by subsequent practitioners and have impacted on the tradition in a positive way.

As mentioned above, there are degrees of creativity, and the idea can be used casually to refer to all kinds of things. Creativity, as used here, refers to the kind of activity carried out by creative artists who initiate new conventions, creating work displaying originality and spontaneity, and for this, some conditions for acceptance should be established, only some of which are unique to radical creativity: to be original, work cannot have been carried out in accordance with existing conventions, for example craftsmanship; to be radical, creativity must advance beyond what was previously known. Instances of accidental processes can be ruled out; if the agent is to be responsible for what is created, the creative activity must result from some voluntary and conscious effort to control and direct the process that will be relevant to the new structure. But at the same time, in order to display spontaneity, there must be some independence from the intentional control and the routine practices of the artist. To be original and spontaneous, a work must not be a copy of an original form, or the result of something taught. If agents were deemed creative purely in relation to their own knowledge, all learning would result in new creation. So the radical newness resulting from the creative process is relative not only to the individual agent's consciousness, but also to the orderly development of processes leading to the agent's activity; in Hausman's terminology:

given the "horizontal dimension" that provides background data available to the agent before creation, and given assumptions about the normal capabilities of agents, there is introduced a vertical thrust, a leap, seen in the novelty for which he is responsible. This novelty is unpredictable and not traceable to past data and the agent's ordinary capabilities (Hausman, 1975, p.45).

In addition, the work itself must be valuable. As Kant argues in a discussion on genius, originality can be exhibited by nonsense and so be worthless; 'since there can also be original nonsense, its products must at the same time be models, i.e., exemplary...' (Kant, 2000, 5:308, p.186). If something newly created is not nonsense, it has value in representing a new thing that makes sense, it is an instance of new intelligibility. In the creative process the artist strives to generate a specific product the exact nature of which he does not envisage. In spite of this, he does apparently know when the process is complete, and whether the result is acceptable to him. There must then be evaluative scrutiny on the part of the artist, a process integral to the process of creating. (This process itself bears analysis, and I will return to it in chapter seven.) The work is expected to display artistic merit, but also to be a completion of what was being expressed in some sense. The production of a new and unfamiliar work is deemed by the artist to be valuable, in the sense of being 'right' or 'finished'. If it is successful, this judgement will be confirmed by those members of its wider audience capable of understanding it; when it is made public it will be subject to a collective consensual testing. One basis for this consensual value

ascription is suggested by Coleridge's assertion; 'we know a man for a poet by the fact that he makes us poets' (paraphrased in Collingwood, 1958, p.118), suggesting a collaborative effort between artist, work and audience and a widening of the beneficial effect of the work's value. A creative object then has value as a new being, as a coherent structure that exemplifies new intelligibility. This might manifest as the freshness or originality that we find intriguing. But it is also valuable in its specific inherent intelligibility, in being not only determinate but being the kind of structure that 'ought to be' sustained. And whilst it must be valuable for its own sake, independently of tradition, it is this inherent value that may lead to it becoming instrumentally valuable in advancing or initiating a tradition.

Among the many accounts of creativity, that of Carl Hausman is unusual in the close attention he pays to the two aspects which I have identified as often neglected or denied: the primacy of the agent as responsible source of creativity, and the question of how to allow for the actuality of spontaneity in radically new created objects. Hausman bases his account on the structure of intelligible objects in general, claiming that this offers us a way both of identifying radical novelty and making it intelligible. This is in spite of what he considers to be an inherent paradox at the centre of creativity:

Creativity exhibits a tension between two poles: on the one hand, the presence of spontaneity – of radical, irreducible difference – and on the other hand, the presence of an identity that invites a search for a continuity between creations and their contexts. Newness exhibits the first pole; intelligibility seems to demand the second (Hausman, 1975, p.53).

According to Hausman, a characterisation of radical novelty must be based on an understanding of the structure of intelligible objects in general, which are encountered as structures and Forms, and this is because 'Radical novelty appears on those occasions in which there is difference not only of singularity, but also of structure and Form' (Hausman, 1975, p.21). This then is what it means for something to be radically new; recognition of it as such, as opposed to being just a new instance of something which already exists, comes from noting a new structure that exemplifies a new Form. Of course, many philosophers make use of the terms structure and Form, and the meaning of these terms can vary widely from one account to another, so it will be important to determine exactly what Hausman intends. Particular emphasis should be given to the stark difference between his account and theories in which art is *equated* with forms or structures, theories proposing that art consists just in form, understood as systems of relations between colours or sounds for example. In some traditional theories of form and matter, matter refers to what is identical in the raw material and in the

finished product, and the form is imposed on this matter. But, as Collingwood points out, while this distinction might be appropriate to artefacts where the form is a preconceived plan, in art this distinction cannot be made; 'when the poem is written, there is nothing in it of which we can say, "this is a matter which might have taken on a different form", or "this is a form which might have been realised in a different matter' (Collingwood, 1958, p.24). What Hausman is suggesting, is not an alternative methodology, rather his characterisation of newness is based solely on how created things *appear* to us. When we look at things that we think of as creations, we find them to be identifiable and partly describable in terms of their assembled characteristics, and we find that what we identify persists throughout the act of identification. His starting point then is phenomenological; 'I begin with the hope of describing things adequately without initially imposing on them a preconstructed total theory' (Hausman, 1975, p.30). His attention to the appearance of spontaneity, generally neglected in theories of creativity, makes Hausman's theory particularly interesting to me, and I will follow his approach in an attempt to do justice to the evidence it represents.

It is the assembled characteristics of an object that constitute what Hausman calls structure. Singular objects are rendered intelligible (though it must be acknowledged, not exhausted) by the cluster of identified and interconnected characteristics that they exhibit. The identity thus formed and exhibited in the object is the structure. In new structures, the way the connections provide the identity and so the intelligibility of each structure is unique, and so new structures are uniquely dependent on the special role of their constituents. An unprecedented and unique structure appears as unfamiliar because it is different from all other structures. At the point when it came into existence as created, it was new, and thereafter it is new in relation to what went before it. It displays newness of kind as newness of what is singular, but this singularity is singularity of intelligibility; 'intelligibility of a discernibly coherent complex' (Hausman, 1975, p.26).

The structure of radically new objects is intimately bound to their instantiation, structures appear to intelligence as having all their specificity in determinate objects and in that sense they are concrete. As we have said, things are identifiable and partially describable in terms of collected characteristics, and this provides intelligibility. But, to be identified, the structure must endure, it shows up as having an identifiable character before we have identified it, and that character persists throughout the identification process. In enduring, the structure manifests something a temporal, to be sustained as an identity it displays something unchangeable, and this something is what Hausman calls Form. The radically new object then '*exhibits* a new structure which newly *exemplifies* a Form' (Hausman, 1975, p.28).

Form, for Hausman, refers to 'that in virtue of which a cluster of items – a group of things or of already identified characteristics – cohere sufficiently to be recognized, identified, and subsequently characterized' (ibid, p.30-1). Form is an intelligible identity, it could for instance be a repeatable visual shape, but is not restricted to visual patterns, referring equally to, for example, propositions exemplified by certain sequences of sounds. Forms can be displayed with a degree of indeterminateness, for a Form to be absolutely definite is an ideal which need not be attained in order for it to be recognised. So, for example, the identity of an object which we call a cup has a degree of indeterminacy, the exact dimensions of an ideal cup could almost certainly never be specified. However, the coherence of the cluster of items which exemplify a Form must be sufficient to reveal an identity in difference.

For an object to exemplify a Form then implies some kind of coherence, a definiteness of focus which leads to recognition and anchors intelligibility. It is this definiteness, inviting cognition, which is valuable in helping us to make sense of, and convey information about, objects: the thing that is noticed and recognised as possessing a specific character must be at least minimally valuable in the sense of being selected for sustained attention. When a Form is newly exemplified, the occurrence breaks existing conceptual patterns, and it is this that allows us to recognise instances of radical, as opposed to numerical, novelty. The value of Forms then is functional. Form is a functional term to describe how things appear, and Hausman's use of the term can be clearly distinguished from other uses, for example, Plato's eternal forms. Forms in Hausman's sense are not eternal, they appear in time, as exemplified in structures, but their relation to temporality is different from that of structure, in that they point beyond specific instantiations. However, they differ from concepts, which also point beyond objects, in that they are not added to structures in order that we might manipulate them, but are rather just how the structures appear to be.

If we recognise radical novelty in initial apprehension of a new Form, we also try to make it intelligible. A created object can display radical, irreducible difference but also the presence of an identity which draws us to search for continuity between the creation and its context. Trying to understand something, or to find an explanation for an anomaly, usually involves finding relations between relevant connections that constitute some kind of system or coherent scheme, for example detecting connections between creations and their historical contexts to find a causal relation. And this is where a problem arises for an account of creativity: if continuity can be found, in principle prediction must be possible, and the appearance of radical novelty will turn out to be illusory. If prediction is possible, creativity is determined in some

way. Of course, deterministic accounts take many forms, but I take it that most would assert that what appears to be unique can be subjected to explanation in accordance with necessities or regularities that have just not been discovered yet, and therefore all things in the world are potentially predictable. If our starting point is the phenomenological *appearance* of something unprecedented and unpredictable, the charge that needs defending is specifically that made by psychological determinism, which upholds the view of consciousness as misled to some degree in regarding newness as something which is not reducible to the past. This charge is made by Maria Kronfeldner, in the paper *Creativity Naturalized* (2009). Kronfeldner presents two examples, the first of which aims to show that causal determinism is irrelevant to creativity and radical novelty, and the second that appearances of spontaneity are in fact illusory.

### A refutation of a naturalistic account of creativity

Kronfeldner holds that there is no reason to deny that creativity can be explained naturalistically, and that therefore it is susceptible to rational explanation. Whether or not an object is actually causally determined is, on her account, irrelevant to the possibility of an instance of originality and spontaneity in a created object. Kronfeldner's two examples, which are generally accepted as exhibiting creativity, are the creation of a new painting by Picasso, and Kekule's discovery of the benzene ring. I argue, counter to Kronfeldner's argument, that in the first instance, the Picasso painting, determinism is not shown to be compatible with creativity, and in the second, Kekule's discovery, she conflates inspiration with causal determinism, and has not shown that Kekule's idea could have been predicted from antecedent conditions. My conclusion is that the appearance of spontaneity has therefore not been proven illusory in this example.

In her first example, Kronfeldner considers the effect of the change that is brought about by an instance of novelty, and proceeds from the claim that determinism is not incompatible with change of this kind; determinism she argues 'does not imply that an observable change is already entailed in the antecedent conditions, the sum of causal factors leading deterministically to the change' (Kronfeldner, 2009, p.583). Kronfeldner's idea seems to be that the possibility of the observable change is entailed in the antecedent conditions, but that the instance of a new thing only occurs when there is the *actualisation* of a change. And, on her reasoning, creativity consists in making the possibility actual. So, in the example of a new painting by Picasso, even if the hypothetical all-knowing demon posited by Laplace knows that a certain painting, including all its precise characteristics, will be produced at any given time,

the painting still has to *become*, it isn't there yet. It is then the actualisation of the painting, carried out by Picasso, which adds novelty to the world, and Picasso is still responsible for the creative act. Since the occurrence of the painting is determined and known by the demon, it is not metaphysically new, but Kronfeldner argues that what matters is that it is new in this actual world.

I think this argument is unsuccessful. Even apart from the problems associated with the positing of a transcendent realm to which humans have no access, for the creative act of a new painting by Picasso to be subject to determinism, the exact painting would have had to exist, *as an idea* somewhere before Picasso painted it. So, in 'creating' the work, Picasso would merely have been bringing something already existing into actualisation. Creativity is, in Kronfeldner's account, reduced to 'making actual'. But, if this is right, anything coming into being, an email that I have just written or a sandwich that I have just made would count as creative in the same way as a new painting by Picasso. Kronfeldner then has no resources to identify Picasso's activity specifically as being creative. I suggest that the error in Kronfeldner's account results from reliance on an account of possibility that simply precludes a meaningful account of novelty, one in which possibility is taken as preceding actuality. If we want to preserve novelty, we need to distinguish between possibility understood merely as 'not impossible', and possibility understood in an ideal sense, which, as suggested by Bergson, must appear after actuality: 'As reality is created as something unforeseeable and new, its image is reflected behind it into the infinite past, thus it finds that it has from all time been possible, but it is at this precise moment that it begins to have been always possible, and that is why I have said that its possibility, which does not precede its reality, will have preceded it once the reality has appeared' (Bergson 2002, p.229). By focussing on change as the sole indicator of creativity, Kronfeldner fails to offer an interesting account of genuine newness, reducing it instead to a predictable, continuous deterministic process. And whilst purporting to uphold the artist's position as primary creative source, this account in fact reduces Picasso to something resembling the status of Master Craftsman. The fact of whether some creative outcome is determined is not then, in Kronfeldner's account, shown to be irrelevant to the occurrence of genuine novelty in creativity.

In the second example Kronfeldner focuses on causal determination. Kekule's revolutionary discovery that the structure of the benzene molecule forms a ring is generally taken as a paradigm example of creativity. In order for this idea to have been original and spontaneous in the mind of Kekule, it cannot have been explained by direct experience, knowledge, or learning that predated it. Although during the course of his life and achievements, Kekule would have



relied on his own and others' knowledge for his work, in the instance of this particular discovery, something was added 'out of nothing'. Kronfeldner assumes that Kekule himself did not know exactly what caused the idea in his mind, acknowledging that most cognitive processes are unconscious or automatic. But she claims that a report attributed to Kekule, which might or might not be accurate, suggests that although the idea was determined, the deterministic cause was outside the sphere of the particular problem concerning the benzene structure, and that therefore the idea was still an example of creativity. In this report, Kekule describes a dream in which atoms appeared before his eyes 'all twining and twisting in snake-like motion. But look! What was that? One of the snakes had seized hold of its own tail, and the form whirled mockingly before my eyes' (quoted in Kronfeldner, 2009, p.587). At this point, Kekule woke up with the clue to the solution of the puzzle of the benzene molecule in his mind. Kronfeldner writes 'the dance and the visualization of atoms played a crucial role in the association which had to take place for him to "see" that the benzene molecule can form a ring, like the snake' (ibid). She goes on to speculate that Kekule might perhaps have visited a zoo that morning and perhaps watched a dance performance later. These unrelated experiences might be said to have influenced the dream and caused the revelation, but because the hypothetical zoo visit and dance performance fall outside the context of knowledge and enquiry regarding the benzene molecule, the idea still counts as spontaneous and original, and therefore creative. The argument is that although a creative outcome precludes having been causally determined by something directly associated with the process of problem solving in that particular instance, summarised as learning and experience, determination from any other source would not count against its status as spontaneous and original. Kronfeldner argues that what took place was a coincidence between something cognitively 'out of nowhere' and a given problem space, and this coincidence was responsible for the element of surprise and practical unpredictability. As, on her account, coincidences are a kind of chance which is compatible with determinism, the argument is that; 'The real and the hypothesized coincidental factors that influenced Kekule in the story can lead to creativity, with all its characteristic surprise, even though at the same time they and the overall process are compatible with determinism' (ibid, p.588). To reinforce her claim that coincidence is compatible with determinism, Kronfeldner refers to the classic example of the chance fall of a stone from a roof that kills a man. The stone comes 'out of nothing' in the causal space surrounding the man. The man was predetermined to be killed by coinciding with a stone falling, by chance, out of nothing.

In this example, it can clearly be seen that if a man was determined to reach a geographical location 'A' at noon, and that if the material supporting a stone on a roof was determined to fail at one second before noon causing the stone to fall on point 'A', the man was predetermined to

be killed by its falling. Can we equally say that if a man working on a particular chemistry problem observes snakes and dancers moving in specific patterns, the result will correspond to the idea of a molecular ring structure? This seems unlikely. It is important to recognise that the radically new creation was not *the ring structure* but the *idea* of it in Kekule's mind. The movement of the snakes and dancers could not coincide with the idea of the ring structure because the idea didn't exist until Kekule thought of it. Out of the total of Kekule's visual and perceptual experiences on that day, the only possible reason to pick out the snakes and dancers as being significant would be in relation to the idea of the ring structure, an idea which did not yet exist. Kronfeldner is then projecting the intelligibility of the outcome backwards. Kekule was responsible for creativity in this case not in spite of, but because of his antecedent work on the chemistry problem, which enabled him to recognise the intelligible form when it spontaneously occurred to him. If the coincidence posited related instead to all the perceptions which Kekule had on that day coinciding with themselves, there would be no way of predicting which of them would influence him in any direction at all. Rather than explain creativity, by reducing it to coincidence Kronfeldner closes the argument down by suggesting that change 'just happens'. But as it is change brought about by creativity that she focusses on from the start, this is confused.

I do not think that Kronfeldner's argument justifies her claim that radical creativity can be explained naturalistically. This is clearly just one instance of an argument proposing a rationalist account, and does not rule out further instances of more successful arguments. But considering such challenges generally, any account denying the actuality of radical novelty has to explain away the conceptual gap that occurs between what went before and what now exists, and so the appearance of something radically new must be shown to be illusory. It is, however, unclear what kind of illusion this could be. An instance of novelty is not the appearance of one thing that is mistakenly taken for something else, what it is taken for is not something that already exists. We might be mistaken by not realising that what appears as new has been instantiated before, but this would be to deny novelty in this particular case, not the possibility of its existence at all.

If the appearance of novelty cannot be understood as something about which consciousness is deluded, the determinist might argue instead that all instances of novelty-presentation are located in a total deterministic network. However, the determinist would need to include the fact of this novelty-presentation and the subsequent change in available knowledge within his deterministic network. To take an example, with the discovery of the virus there was a change in biological knowledge. A transformation of consciousness was brought about by the inclusion

of the radically new thing, in this case an instance of an unprecedented and unpredicted life form, and *this transformation of consciousness itself* appears as an instance of radical novelty. The determinist would have to explain that this radical transformation is illusory and show how consciousness could rid itself of the illusion of the presentation of something disconnected from the deterministic scheme. If the determinist were to somehow intervene to correct this illusion, he would create a new transformation of consciousness, in itself another novelty-presentation. Unless he steps outside the deterministic network by resorting to a principle of creativity (and in so doing destroys his own argument), the determinist is stuck in an infinite chain of reflection.

Seen from a broader perspective, the determinist would have to explain the gap in knowledge surrounding the appearance of something apprehended as novelty for finite consciousness and the appearance as it would be understood within a deterministic network. An existent consciousness observes something as radically new, but a transcendent consciousness sees it as understandable within a causal chain. A gap then appears between finite intelligence that is taken by surprise, and perfect intelligence that is unsurprised and so not deluded. This problem is illustrated in Kronfeldner's example; what appears to finite intelligence to be a radically new painting created by Picasso, when seen by a transcendent existent turns out to be a continuation of what went before. The determinist could only banish the illusion of radical novelty by assuming an understanding that transcends his own intelligence as a finite existent, limited by and subject to historical conditions. In order to see how the appearance of radical novelty fits within the deterministic network, intelligence would have to assume a posture of omniscience. In denying spontaneity then, the determinist is at risk of avoiding one instance of irrationality by introducing something completely implausible, with an appeal for man to assume the perspective of an all-knowing God. So, the assertion that what appears to be an instance of radical novelty is purely an effect of the limitation of finite human knowledge does not appear to help restore rationality.

So far, the attempt to understand spontaneity rationally whilst also rejecting the claim that it is an illusion has not succeeded, and I argue that it is bound to fail. The artist begins the creative process without initial insight into how the data constituting the internal structure will be organised. The Form, as understood in Hausman's account, is not foreseen but discovered as he creates, and therefore must be discerned through a structure inherent to the totality of elements comprising the finished created object. The moment of insight or illumination in which the artist begins to realise the new structure is the moment when the radical break with the past is made possible. Explanations of spontaneity cannot avoid irrationality because

spontaneity eludes categories and principles, which as rational, take the form of constancy of regularity. They must impose repeatability or continuity on the phenomena they set out to interpret. Spontaneity, because it yields newness, resists continuities.

## The role of concepts

So, on Hausman's account, creativity is, by nature, arational. Does he then create an unnecessary difficulty by understanding it in terms of concepts such as structure and Form? There are certainly other ways of understanding the arational in general to be intelligible. Bergson, for example, argues that because discursive analysis is limited, a non-abstractive, non-discursive knowledge is required. However, a brief consideration of the way Bergson incorporates newness into reality helps to highlight the distinctiveness of what I propose can be achieved by using Hausman's concepts.

A truly intuitive metaphysics, according to Bergson, would not gather together the totality of things, but rather would give each thing an explanation that would exactly fit that thing alone. Each new problem demands a new effort, no important truth will be deduced from one already acquired; intuition 'is the direct vision of the mind by the mind – nothing intervening, no refraction through the prism, one of whose facets is space and another language' (Bergson, 2007, p.30). Intuition signifies immediate consciousness: 'a vision which is scarcely distinguishable from the object seen, a knowledge which is contact and even coincidence' (ibid).

What intuition refers to for Bergson, cannot be reconstituted with known components or abstracted from the whole of reality by understanding, common sense or language. The key to understanding intuition is that thinking intuitively involves thinking in duration. This contrasts with intelligence, which concerns itself with things, with the static, and then adds change as an accident. Intuition starts from the perception of movement as reality itself, the essential is change: 'Intuition, bound up to a duration which is growth, perceives in it an uninterrupted continuity of unforeseeable novelty' (Bergson, 2007, p.33).

An idea achieved intuitively begins by being obscure, but the intuition then achieves a different kind of clarity. Intelligence achieves the clarity which comes of understanding ideas that we already possess, that have merely been arranged in a new order, finding the old in what is new. However, intuition achieves the clarity of the 'radically new and absolutely simple idea' (ibid,

p.34) which, having no elements, cannot be constructed with pre-existing elements. Initially appearing to be incomprehensible, after acceptance it dissipates obscurities whilst itself being obscure. The problems, which seemed insoluble, then disappear. On this account, what we call obscure is merely that which cannot be reduced to generalities.

Bergson acknowledges that thought, whether reached by intellection or intuition, always becomes lodged in concepts such as duration, multiplicity etc. and he proposes the use of both metaphor and what he calls 'fluid concepts', to suggest that which cannot be expressed conventionally. We need a different order of concepts, representations that are flexible and mobile, reflecting a reality that is continuous, creative movement. Mind must transcend simple concepts in order to arrive at an intuition and instead create 'flexible, mobile, almost fluid representations, always ready to mould themselves on the fleeting forms of intuition' (Bergson, 2007, p.180). The suggestion of fluid concepts and use of metaphor certainly aligns with much of our discussion; a world in which there is radical creativity cannot be adequately understood using a precise and determinate language. What is required is a language that allows for subtlety, ambiguity and indeterminateness, for which metaphorical expression is well suited.

But if Bergson allows intelligibility in what is arational, what is missing from his account is an explanation of how intuitions might grasp creative acts which appear as radical breaks interrupting an otherwise intelligible world. Even if most processes occur as continuous developments, at some point in the development of a creative process there is a break in continuity, a break in the structure of the process; 'The initial ordering of experience of Cézanne before he began to paint is different from the ordering of his experience when he completed his painting, no matter how continuous was the actual process in which he engaged while painting' (Hausman, 1975, p.36). If radical novelty is recognised by contrast; the difference between exhibited structure and exemplified Form, and past structures and Forms previously exemplified, the tension that arises is a conceptual tension. The mechanism which Hausman provides helps explain what, in Bergson's account, is unexplained fluidity, and it allows the surprise and worth of the new intelligible Form, which might otherwise go unnoticed, to be exposed. Without drawing attention to this conceptual break, the potential for valuing and building on new advances is not fully realised.

Acknowledging a conceptual tension in the instance of an occurrence of radical novelty clearly does not entail that concepts already exist for what appears; the Form which makes the object intelligible appears at the same time as the structure, and intelligibility must be worked out. And in order to proceed with this kind of project, both Bergson and Hausman endorse the use

of metaphorical expression, in order that what is unprecedented can be acknowledged and made intelligible. To see how this might be achieved I will now return to the subject of metaphor, and show that even though we don't yet have concepts, this doesn't mean that intelligibility is ruled out in radically new objects.

### Metaphor as a model of the creative process

Metaphors can offer insight into the intelligibility of created objects for two reasons: firstly they can be used to signify the unique meanings within instances of radical novelty, for example in musical appreciation we use terms like 'tension', and 'sharp' or 'flat' notes to convey useful meaning without diminishing or levelling the character of the object being described. The metaphorical expressions retain a hint of figurative mystery. Secondly, metaphors offer a model by which we can mark out the fundamental features of novelty because they are themselves products of creative imagination and so they offer an example of how creative imagination can work. As the product of the creative process they reveal something of that process.

Metaphors convey meanings in a specific and unique way. To be understandable and shareable, meanings must be sufficiently familiar to be recognised. But precisely because original meanings introduce unfamiliar, not already shared elements, originative speech seems to be inconsistent with the function of language; to articulate and transmit identifiable shared meanings. If new meanings are to be expressed in creative speech, the implements it relies on must be used for new purposes and in new ways. If speech uses words as a tool, in the way paintings use pigment, and music uses sound, for example, in order to create we need to use the tools in a way that does not follow established habits of use for those tools. And in the case of new meanings, the new purpose can only be discovered as the tools are being refashioned. Their familiar known purposes must be set aside in order to achieve purposes so far unachieved. The speaker must then hint at what he says, forcing those words and expressions that *are* familiar into tension-ridden formulations. And, this is exactly what metaphor does, it calls attention to what, by using established concepts, is not the case, there is an essential element of negation which emphasises a tension between concepts that were familiar and their new usage. The negative aspect is crucial to discerning the new meaning, a meaning not exhausted by the combined terms understood by their literal meanings. The recipient of the metaphor is prompted by the literal inappropriateness to a shift of focus in order to apprehend new qualities and meanings, and crucially to make new use of what is familiar. The set of familiar meanings that make up a metaphor connect with each other in threads even though

there is no one thread which has characteristics common to the whole meaning. The threads exhibit broken connections and contrasts as well as commonalities. Metaphor demonstrates how through the tension and disruption, novelty can be manifest, and some kind of intelligibility sustained, and it does this by making use of the limited familiarity that persists.

In a metaphor we encounter parts functioning as constituents and this functioning and these constituents appear as immediately familiar, the parts present themselves for immediate acquaintance. The context in which they are presented is a structure that can be contrasted with distinguishable antecedent elements and continuities. The structure can then be further analysed as to the consequent elements, the new constituents presented in contrast to the antecedent elements are apprehended as transformations of those elements: Hausman writes 'Change is integral to the way the structure appears to cognition. The new structure is given as a complex of familiar meanings manifest in evolutionary change – from past familiarity through unfamiliarity to new familiarity' (Hausman, 1975, p.122-3). This combination of familiarity and unfamiliarity within the metaphor presents us with a challenge; the contrasts and continuities invite exploration. The complex of familiar meanings in their new structure call for analysis and jolt us into new ways of thinking what is possible. The change that ensues then is not only that which is created in the metaphor but extends to our thinking practice. But the very combination of familiarity and unfamiliarity grounds our thinking so that the result, whilst being a new transformation and a change in the way we think, is not sheer randomness or cognitive oblivion.

From this we can see why metaphor can illuminate the creative process – it is because the creative process shares the structure of a metaphor. The artist initially brings to the process elements that are familiar, technical knowledge and skill, life experience. And he puts this in tension with what is unknown, with what is not yet conceptualised but is being brought into being in a sequence of actions that he controls without foreseeing or pre-determining. The result is an object which manifests a coherence that demands our sustained attention. As observers we recognise it as a new thing with an unfamiliar structure, which somehow demonstrates familiarity. The work is complete in bearing its own necessity, none of the elements can be changed without affecting the whole and an understanding of the work demands the interaction of all the parts observed together.

A newly created object appears as a presence, but as indeterminate. The intelligibility it displays is autonomous – it compels attention regardless of any dependence on any other noticeable thing, or of any similarities in its properties that might be referred to, by which it might be

described. It cannot be described but must be distinguished on its own terms. An instance of created novelty with autonomous intelligibility does not appear to have come into being by chance, it has a necessity, but one which comes only from itself; 'It bears its own necessity, and it appears as responsible – responsible to itself but also as ready to be responsible for something else' (Hausman, 1989, p.49). It is because of this that it displays an incipient familiarity as well as an explicit unfamiliarity – it invokes anticipation of how it might fit into future systems and traditions. When new things come to our attention they appear as relatable, not in the sense of what went before but in terms of the relations to be discovered which originate with them – they appear as exemplary for constituting changes in existing systems or traditions. A radically new object then creates its own relevance. Whilst its value is perhaps hard to see when it appears because it may not fit perfectly into the present, it is constitutive of the future. The creation of *Lion Man* for example introduced a unique and unprecedented phenomenon, a physical object, symbolising ideas about the supernatural, indicating a system of relations that would develop and transform the lived human world. A creation is 'right' in relation to how it will fit in future systems. In creating novelty then, humans participate in the constitution of how the world will be in the future, pushing the boundaries of what is considered possible by creating objects which will be, to repeat Kant's word, exemplary.

The way new things create the systems into which they fit can be recognised by considering the work of Robert Rauschenberg who, in 1951, created a series of modular shaped geometric canvases, painted entirely in white. These all-white paintings denied all possibility of narrative or external reference. In his radical reduction of content, Rauschenberg can be seen as foreshadowing minimalism, which would emerge about 10 years later. In addition to the effect this work had on the art tradition, these paintings, which were generally misunderstood at the time, had a great influence on the experimental composer John Cage, who was partly inspired by them to pursue the corresponding notion of silence and ambient sound in music. In 1952 Cage created the three-movement composition 4'33", a piece of music lasting four minutes and 33 seconds, during which no instrument is played. This piece was equally controversial and again radically detached from previous musical compositions, but it continues to be relevant and to challenge existing ways of thinking; in 2004, the BBC Symphony Orchestra gave the first orchestral performance in the UK. This concert was broadcast live on Radio 3, and as an illustration of the radical nature of the work, this necessitated for the first time switching off the emergency back-up systems programmed to substitute content in the event of prolonged silence. And perhaps extending the influence further, Alastair Williams writes 'there is a nascent environmentalism and politics of the particular in the way that pieces such as 4'33" ... challenge us to open our ears' (Williams, 2002, p.233).



The structure of a creative object demands our attention so that we recognise it as something that *should have been* there – the context of what is presented in the new work demands it, although we could not have recognised this before the instantiation. What makes the radically new object seem familiar is the way we begin to see how it could form a new relevance and become part of a new system. Here we can see how working with the tension inherent in creativity – acknowledging that the results of creative acts are general as examples of their kind, whilst also being uniquely different – is important for understanding intelligibility in radical novelty: ‘Witness of radical change, then, is central, to this model, while apprehension of structure is the framework that makes this possible’ (Hausman, 1975, p.152). As can be seen in the examples of Rauschenberg and Cage, the new possibilities are not necessarily immediately seen or taken up, and it might be that only a few people initially see the potential. It may be that we need the art historians for example to draw attention to the gain in cultural knowledge and the expansion of human limits, showing us how radically different things are from their past. But a successful radically new, aesthetic creation will over time, demonstrate this leap, or advance in understanding, which suggests that the possibilities for creation are never exhausted.

In this account, the phenomenon of human creativity, the capacity to produce things that are new kinds rather than merely new singularities, is sustained. And this capacity can be made intelligible with reference to the tools that we refashion, the broken connections, contrasts and commonalities in the connecting threads of familiar meanings, the awareness of tension and contrast. The detail of creative process itself is not fully explainable, the speaker must ‘hint at what he says’, discovering as he creates. The artist retains responsibility for his creation in spite of not knowing what the result will be, his responsibility extends to working with the creative process and allowing it to develop without predetermination. So, the tension we outlined at the beginning has not been resolved, and it appears to be integral to creativity itself; attention to this illuminates radical novelty by allowing the complete break between the old and the new rather than trying to reconnect them. If we accept that an element of paradox is an inescapable feature of the world, working *with* it in the sense of observing the tension between the familiar and the unfamiliar, the stable Form and discontinuity, and making use of both intuition and rational thought, we can recognise that with the advent of a successful new created object, an advance is made, our world is enlarged, and the horizon of what it makes sense to think is expanded.

## CHAPTER SEVEN: THE CREATIVE PROCESS

In this project I have been looking at ways to understand what kind of entity the human is, and whether our capacity to live meaningful lives depends on there being something beyond the human against which we might measure ourselves. In the last chapter I concluded that humans are the kind of beings that have the ability to create radically new objects, and in doing so, we contribute to the sum of what there is; humans are then co-agents in the creation of reality, a realisation that has the potential to offer our lives a sense of purpose. I would now like to return to the question of answerability as described in chapter four. I argue that a close look at what takes place in the creative process itself reveals something which eludes rational explanation, a phenomenon against which we test our ideas and creativity, and which gives a sense of being beyond the human.

As before, I will restrict my inquiry to creative artists who initiate new conventions, creating work displaying originality and spontaneity, without in any sense meaning to suggest that this kind of creativity is the only interesting or valuable kind. The investigation will focus on the question of whether we are right to think of the human as the sole agent in the creative process. If we conclude that control is shared in some way, with what do we share it, and is our freedom compromised?

There is an air of mystery in the fact that in some instances of creativity, an artist sets out on a course, the endpoint of which cannot be foreseen. Along the way he considers whether each advance he makes is 'right', rejecting anything that doesn't stand up to this test. Perhaps he makes propositions and tries them out, or encounters problems which he works to resolve. But there are many unanswered questions surrounding the process itself: what motivates it, what drives it forward and to what end? And whilst I have argued that the agent is responsible for the created object, there is, in anecdotal evidence and philosophical argument, much to suggest that this responsibility is somehow additionally shared with something beyond, or other than the artist himself.

As an example of anecdotal evidence, Didier Maleuvre, commenting on the film *The mystery of Picasso*, writes 'the film shows that once the initial brush strokes are laid down, the painter seems to *obey* the painting. He is in dialogue with what happens on the canvas. ... unlike the laws of science or architecture or engineering, the laws to which the artist is beholden emerge from the work of art and do not precede it' (Maleuvre, 2006, p.219).

Another example can be noted in an interview with the artist Craigie Aitchison, a self-confessed atheist. When asked why so many of his paintings depicted the crucifixion, Aitchison replied 'I was always striving to get it *right*, so that I could then not paint crucifixions. But I could never do it' (Haste, 2014, p.181). From this interview, Aitchison makes clear that what he was working towards was not the creation of an object of religious significance, the subject matter was, according to him, incidental. His pictures were *of* the crucifixion but do not seem to have been *about* that historical or religious event. Aitchison was trying to get 'something' right but he could not articulate, perhaps because he did not know, what this was.

For Jackson Pollock, the creator is not separate from the work, the work does not come from him but rather he is immersed in what he makes: 'When I am *in* my painting, I am not aware of what I'm doing ... I have no fears about making changes, destroying the image, etc. because the painting has a life of its own. I try to let it come through. It is only when I lose contact with the painting that the result is a mess' (Friedman, 1995, p.100).

These examples suggest that something mysterious is experienced during the creative process. This suggestion is reinforced in a study detailed in the paper *Its Own Reward: A Phenomenological Study of Artistic Creativity* in which the authors report on an investigation into the subject through semi-structured interviews with 11 artists (Nelson, Rawlings, 2007). Among the essential constituents of the experience of the artistic creative process that emerged from the interviews, a key element was a feeling of loss of control: 'For some artists, this can result in feeling that they have 'connected' with a source outside immediate awareness—whether internal (for example, the unconscious) or external (for example, 'nature' or God)—and that they are 'receiving' the ideas from this source' (Nelson, Rawlings, 2007, p.230). The study also reports experiences of heightened mood:

a 'celebration' of existence ... felt as a strong sense of fulfilment or satisfaction, as a sense of connection with something beyond the self, or even paradoxically as a sense of tranquillity. The highness of mood can be experienced not only as propelling the artistic activity forward, but as guiding the direction of the work ... The deep engagement with the work provides an escape from analytical cognitive processes and from the standard experience of self-awareness. The sense of freedom can also emerge from a sense of dissolution of personal boundaries and sense of 'connection' with something beyond the self (ibid, p.232).

One small and admittedly limited piece of research, combined with selected quotes from artists, can do no more than suggest a possible subject for investigation, but it is this idea that I would like to consider. The philosophical debate has tended to focus less on this

phenomenological aspect than on the identification of characteristics and conditions of creativity. Monroe Beardsley, for example, specifically focusses on the part of the creative process that takes place between the inceptive element, the initial spur or stimulus for the work, and the final touch: 'between the thought "I may be on to something here" and the thought "It is finished"' (Beardsley, 1965, p.291). In his search for a naturalistic explanation, as opposed to an account reliant perhaps on divine inspiration or intervention of the muses, Beardsley outlines two possible theories which aim to illuminate what occurs in this interval of time: one which he largely accepts, which he calls the 'propulsive theory', and one which he rejects, the 'finalistic theory'. Both claim to explain, in different ways, what motivates or drives the artist during the creative process. The discussion presented by Beardsley and others surrounding this issue does offer some insight into this process, and it will be helpful briefly to consider their different views. But ultimately, none of them directly addresses the particular aspect that interests me; the apparent resistance experienced by the artist. This is an idea that I will return to.

According to finalistic theories, the controlling agent in the creative process is a final goal, towards which it aims. Berys Gaut is largely supportive of this view, arguing that creative processes are teleological on the grounds that they require the input of an agent: as agents necessarily act, and as actions aim at desired states of affairs, the creative process can be said to be end-driven (Gaut, 2010, p.1041). However, in the case of radical creativity as defined earlier, this theory seems problematic; if creation is taken to mean producing something unprecedented, the creative process cannot involve a known end on the grounds that if it did, the work would already have been created. As Vincent Tomas writes:

Unlike either the rifleman or the academic painter or writer, the creative artist does not initially know what his target is. Although he seems to himself to be "aiming" at something, it is not until just before he affixes his signature or seal of approval to his work that he finds out that this is the determinate thing he was all along "aiming" at, and that this was the way to bring it into being (Tomas, 1958, p.2).

For Tomas, creation itself, as origination, necessarily precludes a foreseen goal. Collingwood too rejects the idea of a teleological account because for him, creation is the expression of unexpressed emotions, emotions which are, as yet, unknown to the artist:

these things, in so far as they are works of art proper, are not made as means to an end; they are not made according to any preconceived plan. ... Yet they are made deliberately and responsibly, by people who know what they are doing, even though they do not know in advance what is going to come of it (Collingwood, 1958, p.129).

One defence of teleological accounts offered by Gaut against these objections cites the example of a creative engineer interpreting an architectural design for a new building. The claim is that creativity is demanded in determining the means to achieve the desired, specified end. Further, Gaut offers the suggestion that an artist might have a partially determined goal: 'a poet or painter may begin with only a rough idea of what she aims to write or paint, and her creativity consists in part in clarifying the goal that she is trying to reach' (Gaut, 2010, p.1041). But whilst the work of the creative engineer can clearly be regarded as teleological, it is confined to instances of a weak conception of creativity, and so does not do justice to the creation of something radically new. I think that Gaut's second argument also fails to make the case for the finalistic theory, because in this instance, even if a goal is partially determined in the form of a rough idea, it is the undetermined part which has the potential to display radical creativity, and the motivating control still lacks an explanation.

A different version of the finalistic theory is the idea that art involves a process of problem solving; the creative process consists in a series of problems and their controlled resolution. In support of this idea, David Ecker quotes the words of a selection of artists, among them Henry Moore, who speaks of 'the problems that have concerned me from time to time' (Ecker, 1963, p.284). Moore describes how sometimes he begins a drawing without a preconceived problem to solve but having started to draw 'a point arrives where some idea becomes conscious and crystalizes, and then a control and ordering begins to take place' (ibid). And again, Ecker quotes the painter Yasuo Kuniyoshi who declares 'Lately I have come to the stage where I actually take a problem and try to solve it' (ibid, p.285). According to Ecker, art history is formed by successions of qualitative solutions to creative problems, and from a historical perspective, perhaps we might call matters of composition, perspective, proportion etc. 'problems' which can be seen to have been overcome. Equally, problem solving seems an appropriate term to describe the work of technical agents or co-workers who assist an artist in realising an artwork. However, Ecker goes on to apply this terminology to the creative process as it unfolds: 'The artist ... exhibits deliberate control over his materials: he arranges qualitative *means* such as lines, colors, planes, and textures, to achieve his qualitative *end*, which we might name "cubist" "impressionist," or "expressionist"' (ibid, p.287). The problems which are posed are those occurring within the medium in which the artist works. The end in view is understood as a total quality, and the process involves making choices which proceed towards that resolution. Also defending this idea, Howard Gardner acknowledges that the problem to be solved in the creative process isn't necessarily something known to the artist in advance. He writes;

The artist frequently begins with a theme, emotion, mood, or quality which occurs to him; he explores his 'incept' primarily within his given medium and achieves a finished product only after considerable experimentation. He is free to select his own constraint and neither he nor his colleagues can anticipate his end product, nor even specify the problem being solved until the final stages of execution are reached (Gardner, 1971, p.106).

Within these accounts, and Ecker's selected quotes, it appears that the 'problems' referred to are primarily those presented by the qualities of the materials the artist works with, and the relationships between those materials. Ecker claims that the pervasive quality of the work acts as the control in solving the overall problem, which is the qualitative end; the process involves performing testing operations 'as component qualities (now here or yesterday there) are thought in relation to total and/or pervasive quality also empirically there' (Ecker, 1963, p.289). At the end of the process the work is 'judged complete - the total achieved - the pervasive has adequately been the control' (ibid). But this returns us to the point being discussed, the artist cannot know in advance what the qualitative end is. Equally difficult to understand is how a total quality can be said to pervade a work through inception to the point of recognition on completion, without being known in advance. It is conceivable that different qualities emerge, gaining or losing prominence during the process, and making the finished work vary considerably from the stages of its development. Concentration on the problems of how to manipulate any given medium seem to avoid the central question of the motivating force for the development of the creative idea itself, the question of what the material is being manipulated towards, the final idea or goal posited by these types of theory. And if in fact the suggestion of a final idea or goal is set aside in favour of an account of art as purely about the manipulation of the medium, it is difficult to imagine anything that might actually show up as a 'problem' at all. For an account of creativity to offer anything helpful, it must, I think, address more than the issues relating to the chosen medium. Above all, the conception of creativity as being driven by 'problems' seems both too limited, and too negative, to succeed as a general rule. Means-ends accounts of the creative process appear to be unsuccessful, particularly when used to explain the creation of radically new objects. Would an understanding of the motivating control as originating at the start of the process, as it does in what Beardsley terms 'propulsive' theories, offer any improvement?

In the 'propulsive' theory, 'the controlling agent is something that exists prior to the creative process, and presides over it throughout' (Beardsley, 1965, p.294). There are varying accounts of what this entity is. For Collingwood, as I have said, it is an unexpressed emotion which oppresses the artist so much that he works to discover what it is (Collingwood, 1958, p.110).

According to Collingwood, we can't say what the emotion is, except that it is the kind of thing that is expressed. What isn't clear from Collingwood's account however, is how we can know that the emotion that sets the process in motion is the same one that we recognise at the end of the process. And if we can't say this, we haven't gone any further towards discovering the motivating force for the creative process.

Vincent Tomas defends a different version of the propulsive theory, arguing that the fact that the artist can be said to be heading somewhere, without having a final result in sight, can be observed by the way in which he can say that certain directions are not right, so exercising critical judgement. Tomas distinguishes two stages in the creative process, the moment of inspiration and that of development. Inspiration may occur instantly, but in the moment of development, which may take years, the artist strives to find out what his inspiration is. He writes and re-writes, paints and corrects etc, until he reaches the point where he can say 'There! That's what I wanted to say, just as I wanted to say it' (Tomas, 1958, p.9). Control then consists in making critical judgements throughout the process, on what has been achieved. The artist knows what is right and wrong because whenever he goes wrong 'he feels himself being kicked, and he tries another way which, he surmises, trusts, or hopes, will not be followed by a kick' (ibid, p.13). On this account, Tomas claims that the artist is pushed from behind. The agent doing the kicking is the inspiration which is 'already there' in the creative process. I think Tomas introduces an important point in suggesting that some kind of intervention occurs during the process, but the claim that the initial inspiration acts somehow as the control in the form of 'kicking' is not convincing. Inspiration is understood simply as the moment 'when the new suggestion appears in consciousness' (ibid, p.9), and this moment is not reached by reasoning or exercise of skill, and it is not subject to our will, but might be 'an emotion, a phantasy, an unclear idea' (ibid, p.13). The inspiration, that which does the kicking, is an unknown and unexplained agent. In addition, Tomas claims that what the artist creates must be adequate to his inspiration, but if the artist doesn't fully know what the inspiration is, how could he judge adequacy?

Beardsley defends a version of Tomas's argument which he develops in an attempt to explain those elements which Tomas leaves open. Beardsley argues that works of art generate their own momentum; 'the crucial controlling power at every point is the particular stage or condition of the unfinished work itself, the possibilities it presents, and the developments it permits' (Beardsley, 1965, p.297). Like Tomas, Beardsley considers that the inspiration, what he calls the incept, exerts control. However, he does not claim that the same influence pervades the whole process. Each stage of the process affects the succeeding stage, but this does not

mean that the initial stage is dominant over the whole. The initial incept might even be lost during the development. The creative process is motivated by negative critical judgement, by the tensions that exist between what has been done and what might have been done, the unrealised potential for improvement. Beardsley makes an important distinction between his own claims and those of other similar theories; for him it is not that the artist is 'looking to see whether he is saying what he already meant, but that he is looking to see whether he wants to mean what he is saying' (ibid, p.299). Beardsley then develops this idea to explain how the artist reaches the point of declaring the work finished. Rather than judging it to have adequately represented the inspiration, suggesting a realisation of something already in his mind, he reaches the conclusion that in the creation of the work 'he had become able to mean something much better than he was able to mean a few months before, and that what he now was able to mean – that is, to make – was enough' (ibid, p.300). The idea that the artist himself is changed by or through the creative process seems right, and might be considered indicative of the claim made above, that the creative process can be understood as being, in some sense, a shared endeavour.

I think Beardsley's account is convincing as far as it goes. It does perhaps elucidate how some instances of creativity move forward, and it acknowledges that change comes over the artist during the process; the artist is not the same at the end of the process as he was at the beginning. And although I think Tomas is incorrect in assuming total and rational control on behalf of the artist as response to being 'kicked' by inspiration, I agree that there appear to be impulses contributing to the activity of the artist. These differing accounts all seem to be attempts at finding common ground in the methods used by artists to approach their work, in the way an artist gets from point 'A' to point 'B'. I think that this is unlikely to yield a positive result, particularly with reference to radical creativity, as I have defined it. The way an artist gets from the start to the finish of a project is, I suggest, infinitely variable. But, this is not to say that no common ground can be found in the creative process. What is shared can perhaps be recognised not by considering the method, but in a closer look at what the artist *experiences* during the process: the feeling of constraint, or resistance referred to at the start of the chapter. At the core of the creative process, the artist, whilst being a responsible agent, is apparently informed or illuminated by co-operating factors, and in this sense his action is collaborative. This is of interest for two reasons: first because focussing on this aspect changes how we understand the creative process itself, and second because if we accept the actuality of resistance as integral to the process, we can allow that the resulting works of art are possible sources of knowledge. I will return to this second point.



## Creativity as a collaborative process

All the theories discussed here agree that during the process the artist exerts some kind of critical control, in the sense that he assesses whether each action he performs achieves a result that is 'right'. But without an envisaged goal or standard to compare it against, this is yet to be explained. And intriguingly, as I claimed earlier, created objects appear to pose constraints as they develop, the constraints emerging out of the created object at each stage. I do not mean by this the intentional constraints which may be chosen by, or imposed on, the artist; the intention to paint using a restricted colour palette for example, or constraints that emerge because of the era during which the artist is active. Rather, that during the creative process, the work itself seems to resist certain actions or decisions, perhaps certain lines or colours in a painting are somehow just wrong, or particular words in a poem demand to be replaced, and the constraints involved appear to be such that they function independently of thought. To offer an example, in an interview about her new book *Milkman*, Anna Burns claims that the fact that her characters don't have names is due not to an arbitrary decision, but to the fact that 'the book didn't work with names ... In the early days I tried out names a few times, but the book wouldn't stand for it' (Leith, 2018). These constraints do not seem to consist in rational choices, but rather to be imposed by the emerging creative objects themselves. In the case of *Milkman*, Burns seems to be reckoning with the failure of the characters to work with names, rather than making a decision that the book should be written this way. Even in retrospect, she does not appear to have a rational explanation of the 'demand' imposed by the characters, and she makes no claim to understand the value of them remaining nameless. And yet the move was clearly not unreflective. Burns can be said to have responded within a context that was not entirely orchestrated by her. This suggests that although the artist is ultimately responsible for the creation, artistic control is perhaps better considered as a collaboration with something 'other than', or beyond, the artist. Any account of the creative process must, I think, at least allow for this experience.

A further, and clear illustration of this idea of resistance located in the created object can be seen in the novel *At Swim Two Birds* by Flann O'Brien (2001). First published in 1939, the story is told in the first person, the narrator is a student at University College Dublin, and the book centres around his thoughts, activities and literary works. The narrator is writing a novel about a man called Trellis, the owner of the Red Swan Inn, who neglects his business to write a novel about the moral decline in modern Ireland, intended to be a cautionary tale. In order to convey

his message, Trellis decides to include detailed descriptions of the moral sins he refers to, as carried out by the characters in his novel, some of whom he invents and some he 'borrows' from other literary sources including Irish folklore and Celtic mythology. However, although these characters are subject to Trellis's plans while he is awake, as soon as he falls asleep they rebel, acting in accordance with their own wills, independently of him, and developing personalities which do not follow the roles he has set out for them. In order to pursue their own interests, the characters conspire to drug, and ultimately to do away with, Trellis. In a further twist in the story, Trellis, overwhelmed by attraction for one of his characters Sheila Lamont, rapes her and she produces a child, Orlick Trellis, who in turn writes a novel about his father. Trellis had created the character of Sheila to exemplify beauty and goodness, and as a result of Trellis's actions she dies of shame. The book is written in a wide variety of styles, including middle-Irish romance, Dublin working-class idiom, and folk-stories, it includes extracts of poetry, letters and reference books, and presents parodies of many different language styles and jargon. The book is an example of a stream of consciousness novel, focussing on the mind of the narrator, but the divide between reality and fantasy is continually confused. Describing the unnamed narrator in the novel, Anne Clissmann writes: 'He tends to translate events into literary clichés, to see things in terms of a previous literary style, and then, almost unconsciously, to parody that style and place it in immediate juxtaposition to another parodied style' (Clissmann, 1975, p.88). What is clear is that O'Brien offers an illustration of how his characters, and those of his fictional character Trellis, act independently of the writer and at times take control of the creative process, presenting a parody of the phenomenon described by Anna Burns as quoted above. The idea that literary characters have a life independent of the context of the work for which they were created is described by E M Forster, who writes:

The characters arrive when evoked but full of the spirit of mutiny. For they have these numerous parallels with people like ourselves, they try to live their own lives and are consequently often engaged in treason against the main scheme of the book. They "run away", they "get out of hand"; they are creations inside a creation, and often inharmonious towards it; if they are given complete freedom they kick the book to pieces... (Forster, 1937, p.92).

*At Swim Two Birds* is intentionally a work of meta-fiction, a self-conscious parody of literary styles, which has as a central theme the relative importance of reality, imagination, mimesis and myth. The idea of independence and rebelliousness as displayed by created characters was one that O'Brien had considered in previous literary works, and was clearly of interest to him. That he considered it a successful subject to parody, indicates his confidence that the phenomenon would be immediately recognisable to his audience.

Whether or not it was O'Brien's intention, the novel as a whole can be viewed as a metaphor for the creative process; the story does not progress in an orderly linear or methodical fashion, all the threads are interconnected, it twists and turns back on itself, it is unpredictable and the author faces continual resistance from the created material throughout. The illustration of the resistance experienced by the fictional author can I think be taken as an indication of the fact that O'Brien recognised this experience as his own. However, the artist, in this case O'Brien, has to direct the process, he has to find a way through and to battle with the characters, finding a way to get it 'right'. The book was highly praised by Graham Greene and James Joyce, and although initially on publication it was a financial failure, it went on to be described in the *New York Herald Tribune* as 'a comic masterpiece of modern Irish literature that is as dazzling as the aurora borealis and twice as difficult to describe' (quoted in Clissmann, 1975, p.82). In spite of its reputation for being a difficult, intellectual book, it went on to reach the top of the best-seller list in Dublin (Clissmann, 1975, p.80). Ultimately then, what results is a highly original and acclaimed work for which O'Brien takes responsibility.

The blurring of the boundaries between reality and fantasy, evident in this novel, reflects something that is missing from the accounts of the creative process which we have considered, and perhaps reveals why they have proved unsatisfactory. As already acknowledged, the artist starts a process whose end he does not foresee, and his total control is thus called into question. But in addition, his ideas seem to take on a reality which constrains him even further. An ontological gap appears between his idea, and the created object that he brings into being. By concentrating only on what motivates the artist to get from idea to finished artwork, as evident in the theories we have looked at so far, this vital element is overlooked. The artist must contend with something other, not just in the sense of taking inspiration or being influenced in a certain direction, but in the sense of real collaboration. O'Brien's work, understood as a metaphor, helps to demonstrate this phenomenon; the novel depends on the basic intelligibility of the idea of unpredicted relationships between the author and his characters, between the characters themselves, and of the resistance they present to the author in his attempt to create the work. But in order to get a real understanding of this phenomenon of resistance, I want to go beyond metaphor, and to consider in more detail the nature of the aesthetic object, and how it develops. And while it is perhaps most clearly demonstrated in this work of literature, the idea of resistance seems to apply equally to creative processes in other art forms.

## Imagination and the unreal object

If the initial concept or inspiration for the creative process takes place in the mind of the artist, the question appears to be about a process in which something that has imaginary form can be made real in a way that results in a transferrable 'real' object, something that can be communicated between the artist and the audience. This difficulty shows up in the theories of creativity considered above. A different instance of a similar problem was highlighted earlier in Sartre's account of imagining his return to Paris on leave but not being able to 'realise' it: 'All in all, that leave formed a whole – a solid, round shape which I glimpsed from afar and thought in January I'd be able to appropriate. But in the end I managed only to aim at that object and miss. At the very moment when I thought I almost had it, it escaped me' (Sartre 1984, p.197). As we have seen from the aesthetic theories discussed above, it is not clear that a general rule can be found to explain how an idea, which is at best vague, can result in a finished material object. The starting point and the finishing point seem to be two very different kinds of thing. So perhaps we need to consider what kind of thing the aesthetic object is: whether it is the mental, imaginary initial idea; or the finished physical object, the painting, the musical score etc.; or indeed whether it is something else altogether.

It initially appears that the aesthetic object must be the completed material work - when we see a finished painting we assume we are looking at an aesthetic object. However, there are reasons to question this assumption. In some art, performance art, or poetry for example, there is no physical product; if a book of poetry is burned in a fire, the poems are not destroyed with it. And artworks seem to have at least some properties which are not directly perceived; commenting on Duchamp's work *Fountain*, Danto writes 'the work itself has properties that urinals themselves lack; it is daring, impudent, irreverent, witty, and clever' (Danto, 1981, p.93-4). These properties then, together with the material object, are constitutive of the work itself, even though not perceptible. Danto argues for a distinction between two perceptibly identical works, one of which is considered a work of art, and one, its material counterpart, which is not. He writes: 'a work whose material counterpart consists of three thumbtacks may have abysses of meaning to which the appropriate aesthetic response might be a religious-cosmic shudder' (ibid, p.105). For Danto, what gives a work of art its identity is interpretation 'nothing is an artwork without an interpretation that constitutes it as such' (ibid, p.135). Presumably then the artist must in some way interpret his own work, in order to recognise it as having an identity as a finished project. This idea is interesting, if not fully explained in Danto's account, but hints at a different kind of status for the aesthetic object.

This approach signals a move away from formalism towards idealism in art, an idea suggested by Lambert Weising: 'Formal aesthetics grasps the aesthetic object purely as a structural entity, introducing into aesthetics a phenomenological-positivist view that rules out any other deep "underlying" dimension' (Weising, Roth, 2016, p.28). By contrast, for idealism 'the beauty of a form is based on the content expressed in that form, that is, on a relationship that extends *beyond the work*, and not on a relation immanent to the work. An idealistic understanding of beauty is bound up with the meaning, the content of the work' (ibid, p.24).

If we deny, with Danto, that the finished material product, the painting or the novel for example, holds the status of aesthetic object absolutely, we must consider an alternative. And perhaps the aesthetic object is better understood as a fluid, non-linear process rather than a static object. The level of thought, consciousness, and active imagination which is integral to the creativity and receptivity of new intelligibility seems to demand this. Ideal theories of art have been advanced by writers such as Croce and perhaps Collingwood, who states: 'the work of art proper is something not seen or heard, but something imagined' (Collingwood, 1958, p.142). The thinking is broadly that 'In order to reach the distinctively aesthetic, we must ignore the surface elements, which can equally be found in non-artistic or practical contexts, and go straight to the mind, which organizes them' (Wollheim, Eldridge, 2015, p.25). Ideal art theories have generally been considered to fail on two accounts: if we make the work of art an inner or mental entity, there is no object which is shared by the artist and the audience, and therefore no link between the two; and secondly, they ignore the significance of the medium. The justification for these objections has been made extensively by Wollheim in his essay *Art and its objects*, and rather than rehearse the argument, I would like to bear them in mind whilst considering a different account, that put forward by Sartre. Whilst, together with Ideal theories, Sartre upholds the idea of the aesthetic object as a mental entity, he does not claim to offer a complete theory of art, and indeed if considered as such, his idea is problematic in several areas. But in what it does suggest, I think Sartre's account gets closer to an account of creativity that can address the question which I have raised: how the artist works towards something he doesn't yet know, whilst maintaining his freedom, particularly in the face of the apparent constraints we have highlighted, and the suggestion of the collaborative nature of the process. As the suggestion of resistance and collaboration concerns the phenomenology of creativity, Sartre's account might offer some illumination. His claims for the ontological status of an aesthetic object entail a detailed and distinctive account of imagination, and how it is bound up with freedom. Through imagination we are able to see what is not, and have a sense of what might be; 'To be free is to be able to step out of the stream of reality, to resist responding

immediately to what reality puts before us, and to consider how we should act' (Hopkins, 2016, p.92).

In common with the theories of Croce and Collingwood, Sartre rejects the idea of the aesthetic object as material artwork, and so avoids the problem of how to get from the imaginary to the real, something that he maintains cannot be achieved. However, for Sartre, the aesthetic object has a particular mode of existence, which he describes as 'irreal'. It is an irreality because it is an object of our imagining consciousness, rather than perception. A grasp of Sartre's theory of imagination is necessary to an understanding of his thoughts on aesthetic objects and experience. This theory is given in *The Imaginary*, and I will very briefly highlight some important points here, not with the intention of providing an explanation of it, or even of defending the theory in its entirety, but rather to explain how this way of understanding the aesthetic object might contribute positively to the question in hand.

The first, and perhaps most contentious, point to note in Sartre's theory is that he makes a sharp distinction between perception and imagination. Perception and imagination are taken as two distinct attitudes of consciousness, principally on the ground that they are mutually exclusive; you cannot both see and imagine the same object at the same time. In support of this distinction, Sartre draws out some contrasts between the two attitudes: unlike perceptual objects, imagined objects are present, but out of reach, they cannot be touched or moved except in an imaginary sense; in perception, knowledge is built up gradually, I think I perceive something, I see one profile at a time and I may be wrong about what I think I see. By contrast, in imagination, the object is given all at once, and I am not mistaken about what I imagine, it is what I think it is. All consciousness is intentional, but whilst perception posits an object as existing, imagination posits it as negativity; it can be given as non-existent, absent, or existing elsewhere for example. These positional acts are constitutive of consciousness, an object is not given to me first and then, in addition, noted as being absent, an object is given to me *as* absent. Whilst perception is consciousness of an object-as-existing, imagination is consciousness of the object-as-imagined – imagination involves consciousness of an object under a mode of negation. Objects of imagining consciousness are described by Sartre as 'not sensible, but rather quasi-sensible things' (Sartre, 2010, p.125). In the case of an imagined object, the object is not available for perception and so I make use of matter that can act as what Sartre terms an analogon, in order to make the object present; the intention is directed at an absent or non-existent object *via* an analogical representative. The analogon (for which I will use the more familiar term analogue), might have physical or psychic content; a photograph, a portrait or a mental image, described by Sartre as matter borrowed from the world of things

and 'matter' borrowed from the world of consciousness. It is important to recognise however that for Sartre, what he often refers to for convenience as the 'mental image' is not an entity that could be studied separately, there are actually, for him, no mental images, but rather acts of consciousness. When we imagine an object, we are not thematically conscious of an *image* of the object, we are conscious of the object itself as-being-absent. When I imagine my cat, it is actually my cat that is the object of my imagination, not a representation of my cat. If I describe how the object appears to me, I offer a description of a cat, I do not describe an image of a cat. Imagination, for Sartre, should be understood as 'a conscious and spontaneous act by which the imagining subject quasi-observes an irreal object' (Elpidorou, 2011, p.20).

With this explanation in mind, we can now consider what kind of thing a work of art is for Sartre. Essentially, the aesthetic object is irreal. The assumption that an artist has an idea and then realises it is, on Sartre's account, incorrect. It is an error to think that the artist can start with a mental image that is incommunicable and from that, produce an object which is available for contemplation to all who access it. There is then, for Sartre, no passage from the imaginary to the real; 'what is beautiful ... is a being that cannot be given to perception and that, in its very nature, is isolated from the universe' (Sartre 2010, p.189). Clearly there *are* real objects, like paintings and musical performances and sculptures, but these things are not themselves the subject of aesthetic appreciation. Instead they have an intermediary role which is to facilitate awareness of what is imaginary, they are the material stuff which animates consciousness. So, in the creative process, the artist produces this material stuff, which becomes the material for the analogue of the imagined aesthetic object. The painting then is 'a material thing *visited* from time to time (every time that the spectator takes the imaging attitude) by an irreality' (ibid, 189-90). The analogue functions as that which resembles, or stands-in for, an object that is absent or non-existent, the irreal objects that are manifested through the canvas do not exist either in the painting or in the world, but are apprehended by the imagining consciousness. Similarly, novelists, poets and dramatists use verbal analogues, and on the stage an actor's body becomes an analogue. The analogue is intended to *represent* the absent or non-existent object, but not to perfectly reproduce it; 'some knowledge came to interpret it and to fill in the gaps' (Sartre 2010, p.50). This reinforces Sartre's distinction between perception and imagination.

To clarify: in Sartre's account, aesthetic experience is made up of three distinct objects: 1. the physical object, for example the canvas; 2. the intentional object, the depicted subject or object, 'that which makes a work of art *of* or *about* something' (Elpidorou, 2011, p.21); 3. the aesthetic object, the irreal object of aesthetic appreciation. Whilst the nature of the aesthetic

object 'depends both on the material analogon and on the depicted object, it is reducible to neither' (ibid).

Imagination, for Sartre, rather than being controlled by perception of external objects is thought constituted and so internally generated and spontaneous: 'What we seek through the picture is not Pierre such as he could have appeared to us the day before yesterday ... it is *Pierre in general*, a prototype that acts as a thematic unity of all the individual appearances of Pierre' (Sartre 2010, p.50). What is demanded then is not reproduction but closer to resemblance; 'The intuitive matter is chosen for its relations of equivalence to the matter of the object' (ibid, p.51). In its intention to represent its object, the analogue is not considered to be neutral, but demands of its audience a response; 'spontaneity of consciousness is strongly solicited' (ibid, p.50), the qualities depicted in a portrait for example are projected beyond the picture to the real person; 'the person in the painting solicits me gently to take him for a man. Likewise, if I know (*connais*) the subject of the portrait, the portrait will have, before any interpretation, a real force' (ibid, p.22).

Sartre is here principally interested in aesthetic experience, and some further explanation is required if we are to attempt an application of it to the creative process. One immediate aspect which needs further consideration is how the spontaneous imagination combines with knowledge, belief and feeling in the development of the aesthetic object. Before making this attempt however, a potential challenge needs to be addressed, which is the charge that Sartre's account is too restrictive in that it appears to refer only to representational art.

If we accept Sartre's claim that the material artwork is *never* the direct object of aesthetic appreciation, how are we to understand non-representational art? Works that have no recognisable content are often considered to refer, if at all, only to themselves, and in instances of this kind, the way the work of art is made and its purely visual aspects might be considered the only significant characteristics. A strictly formalistic account is in clear contradiction with Sartre's position on what the object of aesthetic appreciation is. But as indicated above, Sartre's argument aligns with ideal theories, and as such I argue that his account of the central role of the imagination does not necessarily fail when applied to non-representational art.

The concern that Sartre can accommodate only art which is representational in some sense, is reinforced by his definition of one component of aesthetic experience as 'the depicted subject or object ... that which makes a work of art *of* or *about* something'. But the key here is to note that the aesthetic object is constituted by both the beliefs *and* the feelings of the imagining



subject. The image is a 'synthesis of affectivity and knowledge' (Sartre, 2010, p.72). Feelings are intentional states of consciousness, to admire is to admire *something*. But in admiring something, I do not intellectually judge it to be admirable, rather I am conscious of it *as admirable*, that is how I make sense of it. Feeling might be understood then as (potentially) a non-intellectual kind of knowledge. In a state of cognitive-affective consciousness, the object of my consciousness is at the same time the object of knowledge and of feeling, the affective qualities 'entirely permeate the object' (ibid, p.69). Sartre clearly states that affectivity does not rely on representation; 'One affirms that there must always be a representation to provoke the feeling. Nothing is more false' (Sartre 2010, p.70). An image can appear as a representation within which an affective consciousness can be determined, but equally an affective consciousness, or feeling, can be provoked in the absence of any representative object. As an example, Sartre describes a consciousness which is inspired by the hands of an absent acquaintance as 'consciousness of something fine, graceful, pure, with a strictly individual nuance of finesse and purity' (ibid). The value property is represented in affective experience without representation of any non-evaluative features: 'These hands are given to me *in their affective form*' (ibid, p.71). It seems then that there is nothing to preclude non-representational art from acting as the animating matter inspiring the imagining attitude. If beliefs and feelings equally constitute the aesthetic object, in a work with no representational content feelings will predominate. Describing how Sartre's account might be defended in the example of the paintings of Mark Rothko, Elpidorou states that what we are inspired to imagine in these works are 'objects with indeterminate visual properties but with specific emotional textures' (Elpidorou, 2011, p.27). As intentional, both image and affect must be directed towards something, so in a Rothko painting, what the affective character motivates us to experience must be an object beyond the canvas. This transcendent object might appear *as* familiar, *as* unsettling, oppressive and so on. As Elpidorou writes; 'there seems to be no difficulty in accepting that a Rothko – or in fact, any non-representational painting – motivates us to give rise to a novel unreal object ... an object which ... resides neither in the painting nor anywhere in the world' (ibid, p.30).

Experience of art, according to Sartre, might tell us about ourselves, and can be understood to convey feelings or emotions. But by separating imagination from perception, does Sartre deny art the capacity to tell us anything about the world? Sartre claims that in imagination 'I will never find anything there but what I put there' (Sartre, 2010, p.9).

## Aesthetic experience and the creative process as a way of learning.

From the perspective of the observer, a meaningful interaction with art often leads to an ability to see or think differently, the world seems to have changed for us, and we know something new. And I have suggested that during the creative process, the artist changes; as Beardsley proposed, there is the potential for an artist to come to mean something better than he was able to mean before. If the artist doesn't know in advance what he will create, he can be understood to have learned something about the world as a result of the process:

in creating a visual artwork, the work is formed on the basis of an interpretation in a physical medium produced by ... a unique finite agent or creative ensemble. This means that through this new addition to the world the world is transformed, however slightly ... By creating this transformation, the maker adds to his or her consciousness of both the represented item and the medium of representation and is thence put in a new relation to both factors, to himself or herself, and to the world in general (Crowther, 2009, p.31).

According to the theory we have been considering, the aesthetic object is not the material stuff of the artwork, but rather is located in the imagination. If the artist is initially inspired, and then works on the inspiration in his imagination, are we right to think that he can learn anything he doesn't already know? Equally, in aesthetic experience, the observer is prompted by the material object to apprehend an irreal object with imaging consciousness. Mere perception of the painted canvas will not convey the whole story. But can the activity of his own imagination change the world for him?

In general terms, we find out about the world by perception, both by direct observation and through observing our affective response to what we perceive. I see a fierce wolf, I experience the feeling of fear, and my fear signals that I should be alert to danger. This response forms part of the process of learning that wolves are dangerous. What we perceive constrains what we can know or feel in relation to the perceived object. But imagination, particularly if understood as distinct from perception, appears to be unconstrained. Sartre denies that we can learn anything about what we have imagined on the grounds that we can only imagine what we already know or believe. Clearly by imagining we can learn *something* about the world, at the very least we can learn that we are able to imagine. But for Sartre, an imagined object cannot teach us anything about that object *as imagined*. In addition, the status of the imagined as irreal leads him to deny that imagined objects stand in causal relations. If something we imagine cannot *cause* a feeling, it cannot, by this route, lead to new understanding. Imagination is an act of

consciousness directed to an absent object, and something that doesn't exist cannot logically be understood as the cause of an affective response; 'For us, who have distinguished from the outset between the real imagining consciousness and the irreal object, it is impossible to admit a causal relation that would go from object to consciousness. The irreal cannot be seen, touched, smelled, except irreally. Reciprocally, it can act only on an irreal object' (Sartre 2010, p.136). In the face of this claim it would seem that genuine affective responses cannot be *caused* by imagining consciousness, and so aesthetic objects if understood as irreal, cannot teach us about the world through affective response.

However, the claim that we can only imagine what we already know or believe must be challenged, particularly with regard to our experiences of creativity; 'since obviously artists do constantly strike out on unsuspected paths and bring out undreamt-of forms and ideas, it follows that their way of making runs ahead of the act of knowing' (Maleuvre, 2006, p.196). In perceptual experience, affective response is generally triggered when something shows up as significant or striking, and it is usually the result of an occurrence that we did not foresee and are not conscious of causing or controlling. As a result, we learn something new. And, in contradiction to what is suggested by Sartre, it appears that this phenomenon is also possible in reactions to imaginings, and this is evidenced by some instances of imagined situations where relevant physiological responses are engendered. An example of this is given by Murray Smith who writes: 'Imagine gripping the blade of a sharp knife and then having it pulled from your grip, slicing through the flesh of your hand. If you shuddered in reaction to the idea, you didn't do so because you believed that your hand was being cut by a knife' (Smith, 1995, p.118). In these cases, imagination *does* seem to resemble perception; if I imagine a situation, this imagining is often accompanied by undeniably real bodily responses which are of the same type as those accompanying perceptions of the same type. These physiological responses to an imagined situation, like those of a perceived situation, do suggest a reaction to something unexpected and uncontrolled. There have been various attempts to understand this phenomenon in order to make it consistent with Sartre's account, and I will very briefly mention one of them here.

If we uphold the denial of a causal role for irreal objects, but wish to understand the phenomena of affective response to imagining, we might claim instead that these responses are imaginary responses; we imagine a certain object and then imagine a response to it. In places Sartre appears to indicate agreement with this view; in a passage on suffering he writes 'The distress is indeed there, without doubt, but it *confronts* the victim as imaged, inactive, passive, irreal' (Sartre 2010, p.143). His acceptance of the imaginary response view is far from

certain, at a different point in the same text he says of imaginary feelings 'I do not mean that the feelings are themselves unreal, but that they never appear except in the face of unreal objects' (ibid, p.145). But leaving this aside, some serious objections arise; what would be the link between an *imagined* response and an *actual* physiological change? Secondly, it is not clear why imagined responses would correspond to imagined objects in the way they appear to. In the case of genuine imagination, rather than mere flight of fancy, it is not the case that we imagine, for example, the sight of a fierce wolf running towards us and at the same time experience a real affective response of complete unconcern. If we were able to do this, imagination would have no relation to our knowledge of the world at all, something that our lived experience contradicts. If, as we suggested, Sartre would agree that imagination teaches us something about the world, even though denying that it teaches us about the world as we have imagined it, the outcomes of imagination must correspond, at least minimally, to the nature of what it investigates. As Robert Hopkins writes 'The underlying demand here is for affective imagining to be constrained in some way. Only activities and processes that are subject to constraint can be possible sources of knowledge' (Hopkins, 2011, p.111).

Hopkins offers a possible explanation for the correspondence of imagined object and imagined affective response (Hopkins, 2011, p.100-117). Affective states, fear or pleasure for example, are directed at some object: I do not just fear, I fear something. The object of my fear is presented in a certain way, to be frightened of something is for it to be presented as *frightening*. And although I can imagine *that* I am frightened, an instance of propositional imagining, to imagine feeling frightened experientially, from the inside, the object must be imagined appropriately for the response to be imaginable. So, if we make a distinction between merely entertaining the possibility of something and actually experientially imagining it, we find that imagining is constrained in the way we required. Not only does this explain how responses would be appropriate, it also suggests that we might genuinely learn from feelings that are only imagined by, for example, being surprised at their depth, or by the recognition that we have imagined an extreme reaction to something we initially entertained as fairly benign. This suggests an answer to some of the objections to the imagined response account, although as we have said, it is unclear whether Sartre would have found it persuasive. A genuine concern would, I think, still remain as to whether it is in fact possible to affectively imagine a feeling, from the inside, without actually having it. That discussion is beyond the scope of this project. The imagined response view is not the only available answer to the problem, but if, with Sartre, we are committed to the denial of a causal role to imagined objects, it might offer a way forward.

However, perhaps a more productive approach is to consider Hopkins' further question, which is whether Sartre is necessarily committed to a denial of the view that we respond to imagining with genuine affective responses. Sartre holds that the irreal object of the imagining consciousness, as a non-existent object, cannot be the cause of anything. But by distinguishing, as Sartre does, the act of imagining from the imagined object, he potentially leaves open the possibility that the imagination itself can have a causal role. The causal relation would need to reflect something of the nature of the imagined object: 'the idea is that we are (e.g.) disgusted *because we imagine something disgusting*' (Hopkins, 2011, p.115). So, imagining something causes us to feel a certain way because of the nature of what we imagine. The effects caused by the imagining reflect the properties of the object presented, and causality can be attributed to the imagination rather than the irreal object.

In addition, it can be recognised that by imagining we can learn from relational properties of objects: 'although we cannot learn anything of an object's primary or secondary qualities from mental images of it, we can learn tertiary and relational properties of objects in this way' (Elliot, 2006, p.31). If, as Sartre contends, what we imagine is only what we know or believe, this still allows for us to combine any number of imaginary things that we know or believe, and observe our genuine reactions. The synthetic nature of imagination suggests that we might be surprised by our reactions to new combinations which generate entirely new imagined entities. We might, for example, imagine a relationship between two fictional characters, without fully grasping its significance. When the broader consequences of the new situation occur to us at a later stage, we might well be surprised and learn something new about those characters. It is also conceivable that a painter might be compelled to repeatedly include a particular object in his work without fully understanding what it is about the object that compels him.

Further to this we might also understand the imagination as having the capacity to reflect *more* than the properties of the object presented. Poellner suggests this possibility in the case of what he terms 'affective-evaluative misrepresentation of possible future states of affairs' (Poellner, forthcoming), in which there is an inconsistency between the value a desired object is imagined to have, and the value it actually does have. In anticipation of a certain event, my representation of the details of it come to include value experiences which I then mistakenly attribute to the event itself. This might occur as a result of the desire for individual fulfilment: 'I attach my affective image of 'fulfilment', originally independent of it, to that representation' (ibid). In these cases, recollections of past experiences, which are associated for me with a momentary experience of fulfilment, bring about a transference of that value experience to the anticipated event. It is not the representation of the event itself that has given rise to the

evaluative experience, but rather my act of imagining in which I have included a transferred experience of value: 'it is not difficult to see how someone's strong affective image based on a (non-evaluatively speaking) vague recollection of a childhood summer holiday in the country might be transferred to his current representation of a particular country mansion – "like that one we used to go to on holidays" – such that the latter then strikes him as a highly desirable thing to possess' (ibid). In this example an appraisal is based not only on what the imaginer put in imagination – the country mansion, but on additional information non-thematically transferred. If the imagination is capable of misrepresentation which misleads the imaginer, there seems to be more going on in imagining than Sartre allowed for.

The distinction of the act of imagining from the imagined object which leaves open the possibility that the imagination itself can have a causal role does not, in itself, appear to conflict with Sartre's own position, although his claims for the absolute spontaneity of imagination are perhaps challenged. It does however offer a way for imagination to be both constrained and to have the potential to tell us something about how the world is by engendering real feelings and emotions. The significance of allowing that imagination does have this capacity is identified in Smith's claim: 'imagination is a precondition for progressive social change. It is only through imagination that we can fully grasp the experience and predicaments of groups other than those to which we belong, and so it is central to our rising above self-interest, be it individual or tribal' (Smith, 1995, p.123).

That Sartre's account might justifiably accommodate imagination, and by extension art, as a source of knowledge does not of course warrant by itself the position he proposes in separating perception and imagination - his arguments for doing so are beyond the scope of this discussion. But for our purposes, the requirement to allow for art to be a source of knowledge does not, I think, mean that we have to rule his distinction out.

To return to the question raised at the start of this chapter relating to what happens in the creative process, I have acknowledged that Sartre does not put forward a complete aesthetic theory. In *The Imaginary* what he is interested in is the aesthetic experience. Applying these ideas to the creative process necessarily involves speculation. But I suggest that doing this with the aim, not of establishing what Sartre might have said, but rather to consider what might be plausible, will be productive.

## An account of the creative process based on Sartre's aesthetic experience

If we understand the aesthetic object as an object of imagining consciousness, the idea starts and finishes in the imagination. The artist works to make an analogue of something which is formed in the imaginative consciousness. Belief and knowledge are constitutive of the imagining consciousness, but the work is not a purely intellectual endeavour. It is a synthesis of belief, knowledge, intentions, judgements, emotions and feelings, given form within the productive imagination. The analogue which is produced by the artist may also incorporate an affective dimension, feelings and kinaesthetic sensation for example, it can be constituted by both physical and psychic content to enable it to manifest the aesthetic object. As imagination is thought constituted and not controlled by perception or any external object, the analogue does not perfectly replicate what it represents. And as internally generated, imagination is spontaneous and self-determining, unlike the analogue. But if a work of art is to be made freely available, if it is to appear in the public domain, as Kathleen Lennon writes: 'It must *suggest the form* which imaging consciousness can take. The implication seems to be that the individual reading of the artwork must be, in some sense, demonstratively anchored in what has been presented to us.' (Lennon, 2015, p.43). The analogue can perhaps be understood, to reflect a point made earlier in the context of metaphor creation, as *hinting* at what it stands-in for.

It is important to emphasise that although much of what Sartre has to say about imagination focuses on the positing of an object as absent, his account is not restricted to representing that which has already been experienced. The imagination is productive to the extent that it introduces irreality and novelty into the world, and the artist aims to represent that in an analogue. The inclusion of novelty can be seen in Sartre's description of the act of reading as one of directed creation, 'an absolute beginning ... brought about by the freedom of the reader' (Sartre 2001a, p.34). Incidentally, this signals an avoidance of one of the main objections to Ideal art theories in general: that if the work of art is an inner or mental entity there is no link between artist and audience. In the act of reading where we might think the creative work has already been done, Sartre writes of the reader; 'the work exists only at the exact level of his capacities; while he reads and creates, he knows that he can always go further in his reading, can always create more profoundly, and thus the work seems to him as inexhaustible' (ibid, p.33). The aesthetic object is not then fixed in the material object, Sartre writes: 'the literary object though realized *through* language, is never given *in* language' (ibid, p.32). And clearly here imagination is linked with freedom through the intentional structure of consciousness; through imagination consciousness detaches from the world 'through the various objects which it produces or reproduces, the creative act aims at a total renewal of the world. Each painting,

each book, is a recovery of the totality of being. Each of them presents this totality to the freedom of the spectator' (ibid, p.42). Our freedom then is the condition of our ability to imagine; we are able to think not only of how things are, but of what is possible. Consciousness enables us to deny the object of perception and to posit in its place an irreality that arises out of its negation; 'This denial, Sartre writes, is constitutive of our freedom; it is essential to who we are. To be free or, what amounts to the same thing, to be a consciousness, is to be able to imagine' (Elpidorou, 2011, p.19).

An understanding of the creative process based on Sartre's account circumvents the problem of how the initial idea or inspiration in a creative project becomes realised into something which has a totally different nature, how a mental idea becomes a physical object. It upholds the freedom of the artist as the condition for his ability to create, and the freedom of the aesthetic object because it is not confined to the material object. It gives a creative role to the audience which seems to fit with the phenomenon of aesthetic experience, the sense that we are more than passive observers when we interact fully with art: imagination, writes R.K Elliot 'refers us to our freedom, for in imaginal experience we are at least implicitly aware that what we are perceiving we are also spontaneously creating and maintaining in being' (Elliot, 2006, p.49).

If we pursue the thought of the creative act however, some further concerns need to be addressed. Firstly, by locating the aesthetic object in the imagination, Sartre appears to make the material object, the chosen medium of the artist, incidental, the second of the objections levelled against the Ideal art theories referred to above. If the aesthetic object is not the material artwork, the implication is that Picasso would have created exactly the same aesthetic object had he made *Guernica* out of wool, or composed a symphony instead of a painting. In addition, if we allow that the creative process takes place in the imagination, at what stage is the analogue created? If the material object is representative of the aesthetic object in any sense at all, the creative work would surely have to have been done first, which returns us to the situation we are trying to understand, how the artist proceeds without knowing what he will create until he has finished the project. And in addition, the experience of resistance cited in the accounts at the beginning of the chapter remains unexplained. If the creative work is done internally by the artist, purely by his imagination generating its own real objects, his capacity to be enlightened by the process would appear to be curtailed; to repeat the claim quoted above 'Only activities and processes that are subject to constraint can be possible sources of knowledge' (Hopkins, 2011, p.111). Therefore, as I suggested above, the pure spontaneity of the image, argued for by Sartre, is not sustainable.



I think, however, that these concerns can be dealt with largely by building on, rather than abandoning, Sartre's idea. Although Sartre considers the aesthetic object to be located in the imagination, the work of art itself consists of several parts, one of which *is* the material object. Interaction between the artist, the idea, the material work and the audience, necessarily takes place if, as we have said, the work is to be made publicly available. The imaginary content must be externalised, it must be freed from the exclusivity of the artist's mind as part of the creative process. Externalisation of the imaginary content is achieved during the creation of the material object. If the material object acts as an analogue for the aesthetic object when the work is complete, this does not preclude it from being an integral part of the creative process itself. In the creative process, the artist uses imagination to synthesise beliefs, knowledge, feelings and emotions, but these must be expressed in order for the creative process to move forward and develop. The material object and the specific medium, the paint, canvas, words, are then understood as tools for the expression and development of the ideas, and as part of the creative process. The artist sets the process in motion, working to produce an object that he simultaneously observes by stepping back from it. Seeing it from a distance he might be genuinely surprised by it: 'This is the sense in which the work of art gives me the unwavering deep sense of standing before something that is definitely not me, nor of my own making' (Maleuvre, 2006, p.153). The finished painting, the story, or the dance, becomes the focus, or anchor point, for further development once it appears in the public domain, in the consciousness of the individuals who engage with it. In the process of externalising his ideas, the artist frees them from the confines of his own mind with the aid of the artistic medium, he animates them and they gain independence from him, taking on a life of their own. Having done this, as bearers of independent life, they display the characteristics of living individual consciousnesses, resisting and surprising him, constraining his actions and ideas. This principle of animation becomes visible in the comparison of a formulaic lifeless work of fiction compared to a great and enduring literary work – as Maleuvre writes: 'This is creativity: not to bully matter or sound or meaning into whatever form the mind fancies, but to attend diligently to the concrete accidental development of the work' (ibid, p.200).

## Conclusion to chapter seven.

At the start of this chapter I suggested that something mysterious takes place during the creative process, and Sartre's account, I think, leaves that possibility open:

Tintoretto did not choose that yellow rift in the sky above Golgotha to *signify* anguish or to *provoke* it. It is anguish and yellow sky at the same time ... it is anguish become thing, an anguish which has turned into yellow rift of sky, and which thereby is submerged and impasted by the qualities peculiar to things, by their impermeability, their extension, their blind permanence, their externality, and that infinity of relations which they maintain with other things (Sartre 2001a, p.3).

In Sartre's theory, the work is not limited to the art object, but becomes part of the totality of the universe against which it appears. What is manifested through the material artwork is 'an irreal ensemble of *new things*, of objects that I have never seen nor will ever see but that are nonetheless irreal objects, objects that do not exist *in the painting*, nor anywhere in the world' (Sartre 2010, p.190). The synthesising role of imagination implies an opening out of experience beyond rationalisation. Our imagining consciousness is directed to objects accompanied by emotional textures and affective feelings which alter the course of our thinking and present us with challenges. Our powers of imagination and interpretation appear to come up against real constraints that are of a greater magnitude than mere decisions concerning colours, textures, form etc., the restraints are not just matters of rationality or decision making, there seems to be far more at stake. The elements of surprise or fascination experienced in an interaction with a work of art do not seem to be fully accounted for in descriptions of technical skill with particular properties of materials, or as reactions to represented content, for example.

If we return to the example of *At Swim Two Birds*, we can see a metaphor for how this source of resistance might be understood. In this work, O'Brien creates his characters in a way that makes them come alive. As the characters receive from him the 'anima', the life-force necessary to bring them into full being, they gain independence, and as they achieve this, the artist necessarily loses control. Having granted them freedom, he must in a sense respect their independence and allow them to develop their own natures. But if we take this metaphor as suggestive of something actual, it appears that the artist, in collaborating with what is not limited to himself, can create something that has its own way of prompting and constraining experience - as if it has an unfolding life, it appears to us to intervene independently – where this 'life' is one that the artist did not directly foresee and define. Beardsley suggests this when he writes of the artist; 'the powers he works with are, in the end, not his own but those of nature' (Beardsley, 1965, p.303). This nature is, I suggest, the individuality of the created object and the extent to which it is given life by the artist, breaking away from his control, and is then allowed to live. This directly responds to the concern of Maleuvre who writes:

Current aesthetic sensibility ... prefers to read art forms as acts of personal or social self-representation, vehicles of cultural understanding ... all of which assumes that artistic expression lies within the ambit of articulate ideas ... We have shaped our language to speak learnedly of what is interesting about artworks but we are mute as to what makes them fascinating (Maleuvre, 2006, p.194).

The artist then works in collaboration with his creation. The work is no longer constrained by what appears in the material object of the artwork, or to what is in the mind of the artist or the spectator of the finished piece. The 'anima' which is given by the artist in creation, becomes the force which can convey the aesthetic object, understood as initiated in imagining consciousness, via the material object, between artist and imagining spectator. And this is the source of fascination, we are fascinated by encountering an unexpected appearance of life, where this is more than whatever the artist could have expressed, and more than what is already encoded in the culture. Art points beyond the realm of articulate ideas, and self-assured reasoning, towards ideas that are beyond saying. The artist makes contact with aspects of reality that enable life, and they can play a part in the created work despite not having been deliberately put there by the artist. Certainly not all, but some works of art do give us the experience of connecting with a living presence, and although this is not fully open to explanation, the reality of these instances is hard to deny. Maleuvre describes art as encounter, the artist comes up against the unexpected, suggesting that in creating we are capable of estrangement from our own subjectivity: 'The artist sacrifices what he is and what he knows for the sake of encounter. For without encounter, art is the application of rules ... it waits on the self ... By contrast, an artist must be always ready ... to let reality shatter his notions' (Maleuvre, 2006, p.198). The artist must respond to the artwork on its own terms while at the same time guiding it, and jointly pushing the boundaries of what we can know and understand. Creativity involves an encounter with these limits. The newly created object appears as something that *should have been there*, creating its own relevance, and to return to Bergson, it also creates its own possibility: 'As reality is created as something unforeseeable and new, its image is reflected behind it into the infinite past, thus it finds that it has from all time been possible, but it is at this precise moment that it begins to have been always possible' (Bergson, 2002, p.229).

Here then we can see the beginnings of a response to the question raised in chapter four: whether our lives can be understood as having an ultimate meaning or purpose in any sense, and whether there is anything to which we are answerable. I suggested that there was a way in which we could experience a sense of being answerable to something greater, without necessitating a transcendental realm. In the acts both of creating and responding to aesthetic objects we encounter resistance that appears to come from outside or beyond us. Does this

indicate a realm beyond the human? Perhaps. Or, is this unknown resistance with which we collaborate, content which is buried so deep within us, it falls well below our conscious level? That too is possible. The resistance takes the character of the unknown. And in creativity we push against this to test whether it is just unknown at this moment, or intrinsically unknowable. We come up against something that resists our control, and the process becomes one of collaboration, we collaborate with the unknown. This is why I suggested that the quality of passivity in a mystical experience needs qualification; in the creative process the unknown, that which resists, is not experienced as altogether passive, it calls for our attention and imposes demands. In a successful creative endeavour we work with it, allowing it to shape the process. If we approach this phenomenologically, we find Heidegger's advice relevant:

The achieving of phenomenological access to the entities which we encounter, consists rather in thrusting aside our interpretative tendencies ... which conceal not only the phenomenon of such "concern", but even more those entities themselves as encountered of their own accord in our concern with them (Heidegger, 1962, p96).

An understanding of this resistance perhaps can only come through our engagement with it rather than by rationalisation. As Merleau-Ponty writes:

The gesture of reaching one's hand out toward an object contains a reference to the object, not as a representation, but as this highly determinate thing towards which we are thrown, next to which we are through anticipation, and which we haunt (Merleau-Ponty, 2014, p.140).

Our purpose is to explore and expand the boundaries of what there is and what we can know. In our appeal to this resistance for a sense of rightness or completion in creating, we acknowledge that it is not only other than, but somehow greater, than us. And in the created object, as an exemplar of freedom and self-determination, we find something against which to measure ourselves.

I do not, of course, suggest that creativity is the only way to encounter something to which we might be answerable, but merely that it is particularly in an exploration of this act, which is fundamental to human nature, that we begin to recognise this as a possibility. The type of experience which, in chapter four, I described as mystical, is usually associated with religion. And as such, these experiences are often held as having the potential to make a significant change to how we relate to all aspects of life, understood perhaps as somehow salvific. Mark Johnston writes 'Salvation, understood as the goal of religious or spiritual life, is a new orientation that authentically addresses the large-scale defects of human life, and thereby

provides a reservoir of energy otherwise dissipated in denial of, and resistance to, necessary suffering. Salvation, so understood, is not the mere feeling or conviction that you are “saved.” It is a new form of life’ (Johnston, 2009, p.16). But if we look back at the characterisation of mystical experience, we find that the six qualities I identified: Ineffability; noetic quality; transiency; passivity; surprise and scepticism, can similarly be identified in the particular instances of aesthetic experience that I have been analysing. And because the creation and appreciation of art is so central to the human lived experience, this suggests that a life of radical freedom, in which we must create our own purposes and take responsibility for our own lives, exercising our capacity to contribute to the sum of what there is, pushing the boundaries of what is known and what it makes sense to think, can indeed be both liveable and meaningful.

## CHAPTER EIGHT: CONCLUSION

Throughout this project my overall intention has been to articulate a defence of existentialism as a route to an ethical and meaningful life, whilst holding on to two particular strands of thought: one is the phenomenology of the relevant areas under discussion, and in this respect I am broadly in agreement with William James who I quoted earlier: 'For what seriousness can possibly remain in debating philosophic propositions that will never make an appreciable difference to us in action? And what could it matter, if all propositions were practically indifferent, which of them we should agree to call true or which false?' (James, 1985, p.444); and the second is a commitment to a qualified realist position. This position holds that there is a way the world is which is independent of human consciousness, but that reality, understood as everything that exists, includes the contribution of human consciousnesses, their subjective feelings and experiences. Therefore, it is not the case that everything that exists is composed of matter or energy, as is claimed by the materialists. A thorough elucidation of this realist position and its commitments is beyond the scope of this project, but is something which I intend to pursue in the future.

It is not the case that our absolute freedom and autonomy is universally acknowledged in philosophical debate. A much more prevalent assumption is that ultimately, scientific knowledge of the brain will allow a complete and fully objective understanding of the human mind, revealing that we are ultimately controlled by our evolved genetic programming. The thesis that the brain is equivalent to the mind is typically accompanied by the materialist assumption that everything is potentially understandable in a natural-scientific manner. If the brain and the mind are equivalent, everything, including our decisions, beliefs and emotions are controlled by complex neurophysiological processes. Eventually, it is argued, science will be able to gain a complete understanding of this subject, and so humans will, in theory, have the capacity to predict the outcome of any possible situation.

In the introduction I claimed that pre-reflectively, it seems to us that we do have freedom to choose the form our own lives will take so that it accords with our fundamental principles and beliefs, and that we are not unavoidably compelled in any particular direction. The existentialist claim 'existence precedes essence' is one that I consider to be generally in line with our pre-philosophical conception of ourselves. If this intimation of freedom reflects the way things really are, choice is an inescapable concept. Accepting that, to a significant extent, who we are is who we have chosen to be, demands intellectual rigour, honesty and responsibility. I suggest that an autonomous, free state is what remains when we subtract assumptions concerning

scientific discoveries that have yet to be made, and which do not accord with our experiences of what it is like to be human, or religious beliefs that involve a leap of faith. I am not, of course, discrediting the adoption of religious belief, but defend the thought that ultimately this amounts to a choice, which we freely make.

Further, I propose that if there is no real possibility of human autonomy, many of the wider concerns about whether we can find an ultimate meaning to life are hypothetical. Unless we really are autonomous, the question of the meaningfulness or otherwise of life is limited in scope, and can ultimately be delegated to something or someone other than us. We might perhaps find meaning in an everyday sense within our individual or community projects, but if we stand back from these and look at human life as a whole, the question ‘what is the point of that?’ will have a differing degree of importance in accordance with our fundamental beliefs about our level of autonomy. So my first task has been to defend the intellectual credibility of a theory of human autonomy, and for this I looked at Sartre’s conception of freedom. Sartre offers a radical account in defence of this position: for Sartre we cannot escape an understanding of ourselves as free. This account is detailed and complex and introduces many counter-intuitive propositions. But there is a vibrancy in Sartre’s thinking which makes his ideas stand out, offering the potential to change the way we see the world and our place in it.

For Sartre, we are free because we are humanly conscious, and consciousness, being empty and translucent, cannot be influenced by anything outside itself, without that influence being experienced. Sartre makes some particularly interesting arguments as to why the things about our lives which we can’t change, our place, our past, our environment, still do not diminish our freedom. Humans then, on this account, can be understood as beings who are self-determining. For our autonomy to be complete, our actions must be responsive to reasons that are determined by ourselves, and which apparently justify the choosing of them, rather than being taken randomly or by chance. I have argued that a modified reading of Sartre’s account makes his theory of human autonomy intellectually credible.

But I also proposed that we have an *obligation* to take up our freedom – to live as beings who are absolutely certain of our status as autonomous – in order to take responsibility for the choices we make. This has consequences for the way we regard others; their freedom to choose must logically be honoured. And there are wider consequences, we cannot at any point hand over responsibility for our decisions and actions to a higher authority; responsibility rests with us alone.

Sartre's outlook on humanity is not generally considered as an example of moral virtue. For Sartre, nothing justifies any decision I make except the fact that it is mine. I am free to choose who I will be without having to answer to anyone or anything except myself. And whilst there are reasons to think that Sartre was in fact concerned with ethics, he did not develop a complete theory. He certainly did not rule out the possibility of living ethically as absolutely free beings, clearly stating in *Existentialism is a Humanism* that 'nothing can be good for any of us unless it is good for all' (Sartre, 2007, p.24). Sartre's understanding of the absolute value of freedom is oriented towards experiential completeness of oneself and others. But there are concerns which need to be addressed surrounding, for example, his conception of the other, as described in *Being and Nothingness*, as a locus of conflict and also as an object which is merely useful for my self-justification. With this in mind, I suggested that Kant, who also upholds freedom as the source of all value, initially seems more promising. For Kant, if you are rational, you must value freedom, freedom and rationality are synonymous, and it is rationality that confers personhood. If you act in accordance with morally right thinking, you are beyond reproach. Rational ideas alone perform an action guiding role.

Kant, in theory, provides strong guidance on how to live ethically. But in practice, I am not convinced that it gives us enough. What seems to be missing from Kant's account is a place for empathy. Kant excludes emotions from his account which, I argue, reduces the capacity for 'seeing' a situation as being in the moral domain: 'moral perception is not conceived as something to be naively taken for granted, to be presupposed as a mere given, but is conceived as an accomplishment, a complex feat, and an utterly precarious one at that' (Vetlesen, 1994, p.6-7). Moral obligations must extend to whole persons and real-life situations. So, to achieve a satisfactory account of ethics grounded in radical freedom, I have argued that we need to move beyond both Kant and Sartre.

Reflecting on the question of empathy and the need for moral perception, it seems that we can identify parallels between sense perceptions and some emotions, and an exploration of this claim offers insight into how emotions can be understood as having the capacity to give us direct access to values. Emotions then are, potentially, ways of discovering the world. This allows us entry into the moral domain which I have argued, following Vetlesen, requires a sequence of perception, judgement and action. Empathy involves a direct experience of the other, taking account of his values: 'So long as we identify the world presented in our own consciousness with *the* world, we fail to realize that our values are just one limited set of the world's values' (Sprigge, 1988, p.255).



Taking up our freedom then does not necessitate an isolated, amoral and individual existence. But if we are free, autonomous beings, then we face some serious decisions about what kind of people we want to be, and how we are to make the best use of our lives. Without an overall plan laid out for us, or any rules to obey, against what standard can we assess our lives and our actions? If we want to be involved in projects that matter, we need something against which to measure what we do. For this to be effective we arguably want it to be something beyond our limited human perspective: 'If significance is to be conducted back along the chain, its source must be "beyond the human", for whatever remains within the precincts of the human always inspires the question of its own significance' (Cooper, 2012, p.271).

After considering what kind of meaning we might be hoping to find, I argued that even if we could establish some kind of intrinsic meaning held by the totality of all that there is, this fact could not have any real meaningful effect on us, it wouldn't have the capacity to change how we live or how we view the world. However, if we take a more pragmatic approach, we might find that our lives have meaning in a collective sense as a species, in the work of both expanding what is knowable, and contributing to the sum of what there is. This purpose is necessarily linked to the question of measure or answerability. So, can we consider ourselves answerable to anything or anyone beyond the human realm if we are at the same time free and autonomous beings? David Cooper considers that as a source of measure is not available from within an uncompensated humanist position, such a position is not liveable.

So my two related areas of interest are the extent to which humans *can* find meaning, in both the expansion of our knowledge of what exists and the contribution to the sum of existing things, and the question of whether we can find a sense of measure or answerability, which would enable us to assess our lives, from within a position of uncompensated humanism.

My challenge to Cooper's position is that experience does suggest some kind of answerability can be detected which, whilst allusive of a realm beyond the human, can be interpreted in a variety of ways, some of which remain firmly within the human world. In order to get a closer understanding of what form this answerability might take, I first consider a relatively common experience, that of being overcome by awe at the top of a mountain, and set this against a definition of perhaps rarer, but certainly real, experiences which can be defined as mystical. Whilst mystical experience is often associated with, and used to affirm, traditional religious belief, it seems that the defining qualities can be applied without distortion to this more commonplace experience. Having found much shared ground between these two types of

experience, I suggest that something similar can also be found within the paradigmatically human endeavour, the creative process.

Taking up the first area of interest, expanding our knowledge of what is, led me to an analysis of the potential of metaphor. If we are dealing with things that we don't know, we struggle to communicate in literal language and existing concepts, and so metaphor presents itself as a uniquely useful tool. Metaphors have been used traditionally to convey spiritual truths and perhaps to reveal hints of things that are otherwise beyond our human capacity to comprehend. But a study of the various available metaphor theories reveals that there is much more to metaphor than this.

Creative metaphor theories can be sub-divided into two categories, those that consider the purpose of metaphor to be world disclosing, and those that propose it can also play a role in world constitution, having the potential to exceed the boundaries of what *is*, to create something entirely new. When functioning as a cognitive instrument, a metaphor can certainly generate new knowledge, the composer and the audience, when fully engaging with it can gain knowledge which neither of them previously had. Metaphor can then deepen and enrich our cognitive grasp of the world. And I argue that this new knowledge should not be understood as merely a new way of carving the world.

According to Max Black, metaphors produce genuine semantic innovation by connection and contrast – importing word meanings and oral traditions from specific cultures over many generations. By changing the relation between the primary and secondary subjects, new knowledge and insight can be generated. Similarities between the two subjects are not only discovered, but generated. Metaphors then create new things that it makes sense to think. Merleau-Ponty develops the idea of creative innovation, referring to 'the occult trade of the metaphor [which] is able to institute relations that the thinker can grasp only afterwards ... and in a certain sense despite himself' (quoted in Vanzago, 2005, p.431). This account recognises that metaphor is not limited to intellectual content which is fed in by the composer and audience. Metaphor incorporates what matters to me, my values, imagination and feelings, and these elements have a semantic role. Imagination both sees and produces likenesses – and the image is understood as an emerging meaning. Metaphorical meaning is not a fixed, separate entity, it is better understood as an event, forming part of 'a great chain of meaningful being'. Metaphors give us access to a bigger context for feeling and enrichment than literal descriptions can offer, and this can ultimately liberate us from a referentially static or fixed conception of the world, a world that literal language enforces 'allowing not only a changed

metaphor-induced perception of the world, but more extensively, a changed ontology, indeed a changed world' (Hagberg, 2001, p.376).

So, rather than merely deriving from some antecedent meaning, metaphor can be original, its dynamic capability bringing about new possibilities of significance and asserting itself as 'as an autonomous and irreducible dimension of language and thought' (Franke, 2000, p.146). For Hans Blumenberg, absolute metaphors both resist and exceed our grasp, reaching beyond linguistic theory into the arena of experience, the enquiry becoming a collective adventure. Metaphor then can be understood both to disclose aspects of reality and contribute to it. It is interpreted subjectively, but grounded in objectivity, and whilst possessing symbolic truth-value it is shared in common language. The newness referred to is not just a new perspective, but something unprecedented.

After an analysis of these metaphor theories, I concluded that metaphorical expressions can be understood as having the potential to create their own significance by designating something unique, extralinguistic and extraconceptual. The referent is a new object which had no place in the intelligible world before the creation of the metaphor.

This raises some questions as to what it means for humans to create something radically new, and how an unpredicted and unprecedented object might be recognised and made intelligible. From the starting point of an acknowledgement that radically new objects do appear to come into being as a result of human activity, I set out to show that arguments claiming that this appearance can always, either actually or potentially, be traced to a naturalistic explanation, have so far not succeeded. Such accounts set out to show that the appearance of radically new objects is either illusory or open to rational explanation.

To count as radically new, an object must at least instance irreducible and unprecedented difference. It must also be unpredictable, appearing in a reality that was not prepared for it, instancing a spontaneous and radical break from a traceable past. A new thing therefore cannot be identified by its relation to what is already known. But a characterisation can be based on an understanding of the organisation of intelligible objects in general, which I argue, following Hausman, are encountered as structures - the assembled characteristics of an object - and forms - that by which a cluster of items cohere sufficiently to be recognised, identified and subsequently characterised. Recognition of a new thing comes from noting a new structure which exhibits a new form. When a form is newly exemplified, existing conceptual patterns are broken and it is this that allows us to recognise radical novelty. Having recognised it as such, we

have to make it intelligible, and here the continuity we usually seek in looking for intelligibility is necessarily missing. Any account which offers a naturalistic explanation must explain away the conceptual gap that appears between what went before and what now exists.

Having countered one specific instance of a naturalistic argument which claims that we are determined, but that this does not negatively affect our potential for creativity, I argue that such arguments are bound to fail. In a creative project, the creator often begins without a clear idea of how the data constituting the internal structure will be organised. The moment of insight in which the new structure begins to be realised is the moment when the radical break with the past is made possible. Any deterministic account introduces the potential for predictability and traceability, which does not address this radical break, but just assumes that it cannot be absolute, denying spontaneity in favour of potential rational explanation. But the spontaneity instanced in radical novelty eludes categories and principles, because these necessarily impose repeatability or continuity on the phenomena they set out to interpret. Spontaneity, because it yields newness, resists continuities.

However, the intention is not to do away with concepts altogether. If, as I have just claimed, it is by Hausman's concepts of structure and form that a new thing is recognised, the tension between exhibited radical novelty and what previously existed is a conceptual tension. The concepts exemplified do not already exist, but make the object intelligible by appearing together with it. Hausman's conceptual mechanism allows the surprise and worth of the new intelligible form to be exposed. This is important for the potential for valuing and building on new advances.

So, if we need concepts to convey intelligibility, but the concepts surrounding radically new objects do not pre-exist them, we need a tool which goes beyond literal language in order to communicate in this arena. Once again, we are returned to the use of metaphor, for two reasons: firstly, metaphors can be used to signify the unique meanings within instances of radical novelty, whilst retaining a hint of figurative mystery; and secondly, because they are themselves the products of creative imagination, metaphors can reveal something of the creative process because they share the same structure. The artist brings familiar elements to the creative process, his skills, his cultural heritage etc., and puts them into tension with what is unknown. The result is an object which manifests a coherence that demands our sustained attention. We recognise a new thing, with an unfamiliar structure, and yet it somehow demonstrates familiarity. The intelligibility displayed by the created object is indeterminate and autonomous, it compels attention regardless of any dependence on any other noticeable thing,

or of any similarities in its properties that might be referred to, or by which it might be described. And it displays an emerging familiarity as well as an explicit unfamiliarity, invoking anticipation of how it might fit into future systems and traditions. A radically new object then creates its own relevance. The structure of a creative object demands our attention, we recognise it as something that should have been there. What makes it seem familiar is that we begin to see how it could form a new relevance, and become part of a new system.

My search for answerability also takes place within the creative realm, and for this I focus on the creative act itself. What I am particularly interested in is the sense of resistance which artists appear to experience when creating, and the suspicion that this in some sense comes from the work itself rather than from the artist. It seems as if, although the artist is responsible for the work, he cannot do exactly what he chooses with it, the work itself seems to resist his efforts and alter the course of the process, resulting in a created object which has the capacity to surprise even the artist himself. This mysterious phenomenon must to an extent remain a mystery, but it does nevertheless bear a closer inspection.

For the created object to have this kind of potential it must be considered as more than merely a material, physical object. Sartre's account of the aesthetic object as having three distinct components, the physical object, the intentional object and the irreal object of aesthetic appreciation suggests how this might work. It allows an understanding of the aesthetic object as located in the imagination, and as a synthesis of belief, knowledge, intentions, judgements, emotion and feelings. Imagination is internally generated and spontaneous, but the idea itself requires anchoring in the physical in order to appear in the public domain. The finished material artwork becomes the focus or anchor point for individuals who engage with it. But having been animated by the artist, it takes on a life of its own and so gains independence. This 'anima' is then perhaps what allows artworks to fascinate and surprise, we find beauty in what surprises us because we lose our subjective grip on reality and focus on that which is definitely not us: 'Who shall argue that experiencing aesthetically is central to who we are and may become as human beings? For in this experience, accessible to every person, is the humanizing cord that links us to the world. This is the coming of world into life, the coming together of our selves and being. We all sit at the edge of the most profound experience. No need to climb mountains' (Ginsberg, 1986, p.71).

And so, it seems, life does have a purpose. We contribute as co-creators, existing at the boundaries of what is known and extending outwards to what is yet to be known as well as what is yet to be created, enlarging our world and expanding the horizon of what it makes

sense to think. And for this to be possible, we must be beings who are radically free and autonomous, in order to be capable of spontaneous action. But in addition, I think we can experience some kind of answerability, to that which resists us, to which we appeal in the act of judging a created work 'right' or 'finished' for example. In the paradigmatic human activity of aesthetic creation, we measure our efforts against this unseen presence and test whether we can do more, whether we have gone wrong somewhere and need to make adjustments. What this entity is, is more easily experienced than articulated. Although there are many who believe that it must be something deep within our subconscious, I am inclined to agree with Sartre, who would deny this possibility on the grounds that there is nothing *in* consciousness. But the fact that this possibility remains, together with the acknowledgement that the phenomena occurs within the human activity of creation, directly responds to the concern that whatever falls within the precincts of the human must necessarily inspire the question of its own significance.

In the process of aesthetic creativity, I have argued that we can detect a sense of resistance, against which we measure what we think and how we act. That which resists appears as something which is beyond and above us, or to use Sartrean language; it is given as 'over against me', or as distinct from me. And if we refer back to chapter four where I outlined the six qualities which define mystical experience, there is a strong sense that these qualities are present here too, suggesting that this phenomenon is shared by experience extending beyond the art world. Indeed, creativity is a feature of the development of new ideas and developments in any realm. Humans are naturally creative creatures, we need to explore, to build on and develop what we know and understand. And in this process of exploration, the experience of resistance can provide a framework within which our lives feel grounded. It is this sense of something greater than us, to which we are connected, that I suggest makes human life both liveable and meaningful.

## BIBLIOGRAPHY

- Anderson, T. C. (1979). *The foundation and structure of Sartrean ethics*. Lawrence: Regents Press of Kansas.
- Anscombe, G. E. M. (1958). *Modern Moral Philosophy*. *Philosophy* 23(124) 1-19.
- Antliff, M. (1999). Rhythms of Duration. In: Mullarkey, J., ed., *The New Bergson*. Manchester University Press, pp.184-208.
- Aristotle. (1998). *Metaphysics*. Transl. Lawson-Tancred, H., Penguin Books.
- Avis, P. (1999). *God and the Creative Imagination. Metaphor, symbol and myth in religion and theology*. Routledge.
- Baldwin, T. (1980). *The Original Choice in Sartre and Kant*. *Proceedings of the Aristotelian Society*, Volume 80 (1), pp.31-44.
- Beardsley, M. (1965). *On the creation of art*. *The Journal of Aesthetics and Art Criticism*, Volume 23 (3), pp.291-304.
- Beauvoir, S.d. (1976). *The ethics of ambiguity*. Transl. Frechtman, B., New York: Kensington Publishing.
- Berger, J. (1996). G. Bloomsbury.
- Berger, J., J. Mohr and S. Blomberg. (2010). *A seventh man: a book of images and words about the experience of migrant workers in Europe*. New ed. London; New York: Verso.
- Bergson, H. (2002). The Possible and the Real. In: Ansell Pearson, K., Mullarkey, J., eds., *Key Writings: Henri Bergson*. New York: Continuum.
- Bergson, H. (2007). *The Creative Mind*. Transl. Andison, M. L., Dover Publications.
- Black, M. (1993). More about Metaphor. In: Ortony, A., ed., *Metaphor and Thought*. Cambridge University Press, pp.19-41.
- Blumenberg, H. (1993). Light as a Metaphor for Truth. In: Levin, D. M., ed., *Modernity and the Hegemony of vision*. University of California Press, pp.30-62.
- Blumenberg, H. (2010). *Paradigms for a metaphorology*. Cornell University Press: Cornell University Library.
- Boden, M. A. (2004). *The Creative Mind: Myths and Mechanisms*. London: Routledge.
- Brandon, S. (1969). The weighing of the soul. In: Kitagawa, J. M., Long, C. H., eds., *Myths and Symbols. Studies in honour of Mircea Eliade*. University of Chicago Press, pp91-110.
- Campbell, J. (2002). *The inner reaches of outer space. Metaphor as myth and as religion*. California: New World Library.
- Campbell, J. (2013). *Thou Art That. Transforming religious metaphor*. California: New World Library.

- Catalano, J. S. (1974). *A commentary on Jean-Paul Sartre's Being and nothingness*. University of Chicago Press.
- Clissmann, A. (1975). *Flann O'Brien. A critical introduction to his writings*. Dublin: Gill and Macmillan.
- Collingwood, R. G. (1958). *The Principles of Art*. New York: Oxford University Press.
- Cooper, D. (2012). *The Measure of Things. Humanism, humility, and mystery*. Oxford: Clarendon Press.
- Crowther, P. (2009). *Phenomenology of the visual arts (even the frame)*. Stanford University Press.
- Danto, A. C. (1981). *The Transfiguration of the Commonplace: a Philosophy of Art*. Cambridge, Mass; London: Harvard University Press.
- Dean, R. (2013). *Humanity as an Idea, as an Ideal, and as an End in Itself*. Kantian Review Volume 18 (2), pp.171-195.
- Deleuze, G. (2014). *Difference and Repetition*. Transl. Patton, P., Bloomsbury.
- Detmer, D. (1988). *Freedom as a value. A Critique of the ethical theory of Jean-Paul Sartre*. Open Court Publishing Company.
- Dissanayake, E. (1991). *What is art for?* University of Washington Press.
- Dokic, J., Lemaire, S. (2013). *Are emotions perceptions of value?* Canadian Journal of Philosophy, Volume 43 (2), pp.227-247.
- Ecker, D. W. (1963). *The artistic process as qualitative problem solving*. The Journal of Aesthetics and Art Criticism, Volume 21 (3), pp.283-290.
- Elliott, R. K. (2006). *Aesthetics, Imagination and the unity of experience*. Ashgate.
- Elpidorou, A. (2011). Imagination in non-representational painting. In: Weber, J., ed. *Reading Sartre. On phenomenology and existentialism*. Routledge, pp.15-30.
- Ferraris, M. (2015). *New Realism*. Bloomsbury.
- Feyerabend, P. (1999). *Conquest of abundance: a tale of abstraction versus the richness of being*. Transl. Terpstra, B., University of Chicago Press.
- Formosa, P. (2011). *Is Kant a Moral Constructivist or a Moral Realist?* European Journal of Philosophy, Volume 21 (2), pp.170-196.
- Forster, E. M. (1937). *Aspects of the novel*. London: Edward Arnold.
- Franke, W. (2000). *Metaphor and the Making of Sense: The Contemporary Metaphor Renaissance*. Philosophy & Rhetoric, Volume 33 (2), pp.137-153.
- Friedman, B. H. (1995). *Jackson Pollock: Energy Made Visible*. New York: Da Capo Press.
- Gardner, H. (1971). *Problem-Solving in the Arts and Sciences*. The Journal of Aesthetic Education, Volume 5 (1), pp.93-113.



- Gardner, S. (2009). *Sartre's Being and nothingness: a reader's guide*. London; New York: Continuum.
- Gaut, B. (2010). *The Philosophy of Creativity*. *Philosophy Compass*, Volume 5 (12), pp.1034-1046.
- Gellman, J. (2009). *Jean Paul Sartre: The mystical atheist*. *European Journal for Philosophy of Religion*, Volume 1 (2), pp.127–137.
- Ginsberg, R. (1986). Experiencing aesthetically, aesthetic experience, and experience in aesthetics. In: Mitias, M. H. ed. *Possibility of the Aesthetic Experience*. Dordrecht: Martinus Nijhoff, pp.61-78.
- Guyer, P. (2000). *Kant on freedom, law, and happiness*. Cambridge: Cambridge University Press.
- Hagberg, G. L. (2001). Metaphor. In: Gaut, B., Lopes, D., eds. *The Routledge Companion to Aesthetics*. Routledge, pp.371-381.
- Haitos, A. (2010). *Possibility, Novelty, and Creativity*. *Stance*, Volume 3 pp.9-16.
- Harries, K. (1979). Metaphor and Transcendence. In: Sacks, S., ed., *On Metaphor*. University of Chicago Press.
- Haste, C. (2014). *Craigie Aitchison. A life in colour*. Lund Humphries.
- Hausman, C. R. (1961). *Art and Symbol*. *The Review of Metaphysics*, Volume 15 (2), pp.256-270.
- Hausman, C. R. (1975). *A Discourse on Novelty and Creation*. The Hague: Martinus Nijhoff.
- Hausman, C. R. (1989). *Metaphor & Art. Interaction and reference in the verbal and nonverbal arts*. Cambridge University Press.
- Heidegger, M. (1962). *Being and Time*. Transl. Macquarrie, J. and Robinson, E., Blackwell Publishing.
- Hopkins, R. (2011). Imagination and affective response. In: Weber, J., ed. *Reading Sartre. On phenomenology and existentialism*. Routledge, pp.100-117.
- Hopkins, R. (2016). Sartre. In: Kind, A., ed. *The Routledge Handbook of Philosophy of Imagination*. London; New York: Routledge.
- James, W. (1985). *The Varieties of Religious Experience*. Penguin Classics.
- Jaspers, K. (1969). *Philosophy, Volume 1*. Transl. Ashton, E. B., Chicago; London: University of Chicago Press.
- Johnston, M. (2009). *Saving God: Religion After Idolatry*. Princeton, N.J.: Princeton University Press.
- Jollimore, T. (2009). "Like a Picture or a Bump on the Head": Vision, Cognition, and the Language of Poetry. *Midwest Studies In Philosophy*, Volume 33, pp.131–158.
- Jones, G. E. (1980). *Sartre, Consciousness, and Responsibility*. *Philosophy and Phenomenological Research*, Volume 41 (1/2), pp.234-237.
- Kane, R. (2009). Responsibility, Luck and Chance: Reflections on free will and indeterminism. In: Watson, G., ed., *Free Will*. Oxford University Press, pp.299-321.
- Kane, R. (2005). *A contemporary introduction to free will*. New York; Oxford: Oxford University Press.

- Kant, I. (2000). *Critique of the Power of Judgement*. Transl. Guyer, P., Cambridge University Press.
- Kant, I. (2012). *Groundwork of the metaphysics of morals*. Transl. Gregor, M., and Timmermann, J., Cambridge University Press.
- Kant, I. (2013). *The Metaphysics of Morals*. Transl. Gregor, M., Cambridge University Press.
- Kaushall, J. N. (2018). *Can Art fight Fascism?* Philosophy Now, Volume 129, pp.14-16.
- Kirkpatrick, K (2013). *Jean-Paul Sartre: Mystical Atheist or Mystical Antipathist?* European Journal for Philosophy of Religion, Volume 5 (2), pp.159-168.
- Korsgaard, C. M. (1996). *The sources of normativity*. Cambridge University Press.
- Korsgaard, C. M. (2000). *Creating the kingdom of ends*. Cambridge University Press.
- Korsgaard, C. M. (2009). *Self-constitution: agency, identity, and integrity*. Oxford; New York: Oxford University Press.
- Kriegel, U. (2016). Perception and Imagination: A Sartrean account. In: Miguens, S., Preyer, G., Morando, C. B., eds., *Pre-reflective Consciousness*. Routledge, pp.299-321.
- Kronfeldner, M. E. (2009). Creativity Naturalized. The Philosophical Quarterly, Volume 59 (237), pp.577-592.
- Leith, S. (2018). *Pretentious, impenetrable, hard work ... better? Why we need difficult books*. The Guardian. Available at: <https://www.theguardian.com/books/2018/nov/10/anna-burns-milkman-difficult-novel>. [Accessed 05.06.2019].
- Lennon, K. (2015). *Imagination and the imaginary*. New York: Routledge.
- Livingston, P. (2009). Poincaré's 'Delicate Sieve': On Creativity and Constraints in the Arts. In: Krausz, M., Dutton, D., Bardsley, K., eds. *The Idea of Creativity*. Leiden; Boston: Brill, pp.129-146.
- Maleuvre, D. (2006). *The Religion of Reality. Inquiry into the Self, Art and Transcendence*. The Catholic University of America Press.
- Marcel, G. (1965). *Being and Having*. Transl. Farrer, K. London: Collins.
- Merleau-Ponty, M. (1968). *The Visible and the Invisible*. Transl. Lingis, A., Northwestern University Press.
- Merleau-Ponty, M. (2004). *Basic Writings*. Baldwin, T. ed, New York; London: Routledge.
- Merleau-Ponty, M. (2014). *Phenomenology of perception*. Transl. Landes, D., New York; London: Routledge.
- Moran, R. (1989). *Seeing and Believing: Metaphor, Image, and Force*. Critical Enquiry, Volume 16 (1), pp.87-112.
- Morris, K. J. (2008). *Sartre*. Blackwell Publishing.

- Murdoch, I. (1970). *The Sovereignty of Good*. London: Routledge.
- Nelson, B., Rawlings, D. (2007). *Its Own Reward: A Phenomenological Study of Artistic Creativity*. Journal of Phenomenological Psychology, Volume 38 (2), pp.217-255.
- Nozick, R. (1981). *Philosophical Explanations*. Harvard University Press.
- O'Brien, F. (2001). *At-Swim-Two-Birds*. Penguin Classics.
- O'Neill, O. (1975). *Acting on principle: an essay on Kantian ethics*. New York; London: Columbia University Press.
- Pelser, A. C. (2014). Emotion, Evaluative Perception, and Epistemic Justification. In: Roeser, S., Todd, C., eds., *Emotion and Value*. Oxford University Press, pp.108-122. Available at <https://0-www-oxfordscholarship-com.pugwash.lib.warwick.ac.uk/view/10.1093/acprof:oso/9780199686094.001.0001/acprof-9780199686094-chapter-7> Accessed 20.09.2019.
- Pillow, K. (2001). *Jupiter's Eagle and the Despot's Hand Mill: Two views on Metaphor in Kant*. The Journal of Aesthetics and Art Criticism, Volume 59 (2), pp.193-209.
- Plato. (1996). *The dialogues of Plato / translated with comment by R.E. Allen. Vol.3, Ion, Hippias Minor, Laches, Protagoras*. New Haven; London: Yale University Press
- Poellner, P. (2014). Value. In: Luft, S., Overgaard, S., eds., *The Routledge Companion to Phenomenology*. Routledge, pp.297-306.
- Poellner, P. (2015). *Early Sartre on Freedom and Ethics*. European Journal of Philosophy, Volume 23 (2), pp.221-247.
- Poellner, P. (2016). Action, Value, and Autonomy: A Quasi-Sartrean View. In: Baiasu, S., ed., *Comparing Kant and Sartre*. Palgrave, Macmillan, pp.132-157.
- Poellner, P. (2016a). *Phenomenology and the Perceptual Model of Emotion*. Proceedings of the Aristotelian Society, Volume 116 (3), pp.261-288.
- Poellner P (Forthcoming). *Indistinctness in Value Experience*.
- Ricoeur, P. (1978). *The Metaphorical Process as Cognition, Imagination and Feeling*. Critical Enquiry, Volume 5 (1), pp.143-159.
- Sartre, J. (1965). *Nausea*. Transl. Baldick, R., Penguin Books.
- Sartre, J. (1984). *War diaries: notebooks from a phoney war: November 1939-March 1940*. Transl. Hoare, Q. London: Verso.
- Sartre, J. (2001). *The Psychology of the Imagination*. Transl. Warnock, M., London: Routledge.
- Sartre, J. (2001a). *What is Literature*. Transl. Frechtman, B., London; New York: Routledge.
- Sartre, J. (2004). *The Transcendence of the Ego*. Transl. Brown, A., London; New York: Routledge.
- Sartre, J. (2005). *Being and Nothingness*. Transl. Barnes, H., London; New York: Routledge.

Sartre, J. (2007). *Existentialism is a humanism*. Transl. Macomber, C., London; New Haven: Yale University Press.

Sartre, J. (2010). *The Imaginary*. Transl. Webber, J., London; New York: Routledge.

Schumski, I. (2017). *The Problem of Relevant Descriptions and the Scope of Moral Principles*. *European Journal of Philosophy*, Volume 25 (4), pp.1588-1613.

Simont, J. (1999). Sartrean Ethics. In: Howells, C., ed., *The Cambridge Companion to Sartre*. Cambridge University Press, pp.178-210.

Smith, M. (1995). *Film Spectatorship and the Institution of Fiction*. *The Journal of Aesthetics and Art Criticism*, Volume 53 (2), pp.113-127.

Snvarr, S. (2010). *Metaphors, Narratives, Emotions. Their Interplay and Impact*. Amsterdam: Rodopi.

Sprigge, T. L. S. (1988). *The Rational Foundations of Ethics*. London; New York: Routledge & Kegan Paul.

Sprigge, T. L. S. (2006). *The God of metaphysics: being a study of the metaphysics and religious doctrines of Spinoza, Hegel, Kierkegaard, T.H. Green, Bernard Bosanquet, Josiah Royce, A.N. Whitehead, Charles Hartshorne, and concluding with a defence of pantheistic idealism*. Oxford; New York: Clarendon Press; Oxford University Press.

Strawson, G. (2009). Impossibility of moral responsibility. In: Watson, G., ed., *Free Will*. Second Edition. Oxford University Press, pp.212-228.

Stein, E. (1989). *On the problem of Empathy*. Springer Science and Business.

Stern, R. (2012). *Understanding moral obligation: Kant, Hegel, Kierkegaard*. Cambridge University Press.

Taylor, C. (1982). Responsibility for Self. In: Watson, G., ed., *Free Will*. Oxford University Press, pp.111-126.

The Telegraph. (2011). *Professor Sir Michael Dummett* - Culture Obituaries. Available at: <https://www.telegraph.co.uk/news/obituaries/culture-obituaries/8981654/Professor-Sir-Michael-Dummett.html> . [Accessed 17.09.2019].

Thomas, A. (2011). *Alienation, objectification and virtue*. In: Webber, J., ed., *Reading Sartre: on phenomenology and existentialism*. London; New York, Routledge, pp.161-179.

Tomas, V. (1958). *Creativity in Art*. *The Philosophical Review*, Volume 67 (1), pp.1-15.

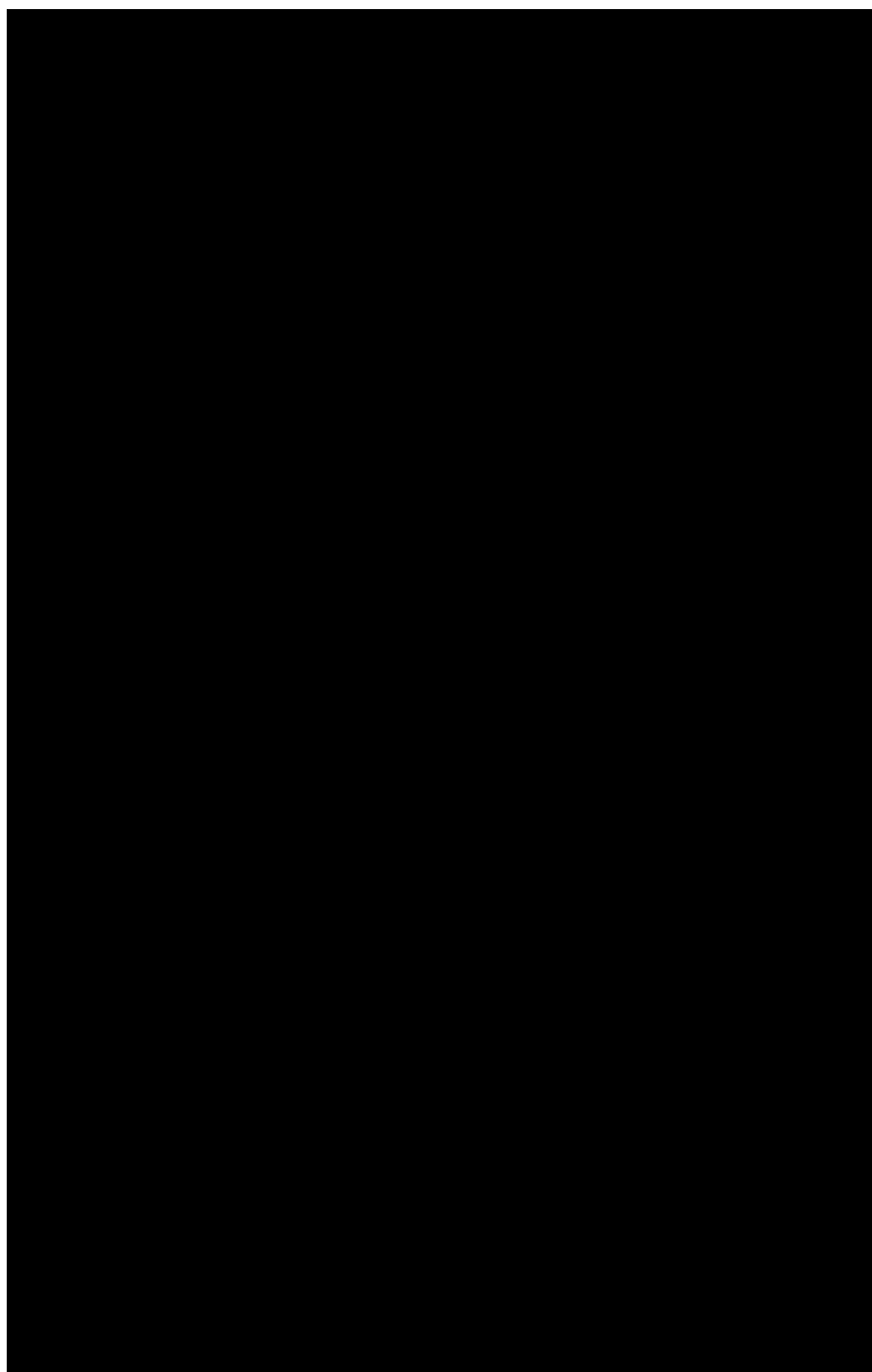
Tracy, D. (1978). *Metaphor and Religion: The Test Case of Christian Texts*. *Critical Enquiry*, Volume 5 (1), pp.91-106.

Vanzago, L. (2005). The Invisible and the Unpresentable. The Role of Metaphor in Merleau-Ponty's Last Writings. In: Tymieniecka, A. T., ed., *Logos of phenomenology and phenomenology of the logos. Book One. Analecta Husserliana, (The Yearbook of Phenomenological Research)*, vol 88. Springer, Dordrecht. pp.429-440.

- Vetlesen, A. J. (1994). *Perception, Empathy, and Judgement. An Inquiry into the Preconditions of Moral Performance*. Pennsylvania State University Press.
- Walton, K. (1993). *Metaphor and Prop Oriented Make-Believe*. *European Journal of Philosophy*, Volume 1 (1), pp.39-57.
- Watson, G. (2004). Two Faces of Responsibility. In: Watson, G., ed., *Agency and Answerability: Selected Essays*. Oxford: Clarendon, pp.260-288.
- Whitehead, A. N. (1967). *Science and the Modern World*. New York: The Free Press.
- Wiesing, L., Roth, N. A. (2016). *The visibility of the image: history and perspectives of formal aesthetics*. Bloomsbury. [online]. Available at: [http://encore.lib.warwick.ac.uk/iii/encore/record/C\\_\\_Rb3354092](http://encore.lib.warwick.ac.uk/iii/encore/record/C__Rb3354092) [Accessed 19.09.2019].
- Williams, A. (2002). Cage and postmodernism. In: Nicholls D., ed., *The Cambridge Companion to John Cage*. Cambridge University Press.
- Williford, K. (2011). Self-Consciousness and the Ego. In: Webber, J., ed., *Reading Sartre. On phenomenology and existentialism*. London; New York: Routledge, pp.195-210.
- Wolf, S. R. (1993). *Freedom within reason*. New York; Oxford: Oxford University Press.
- Wollheim, R., Eldridge, R. T. (2015). *Art and its objects: with six supplementary essays*. [online] Second edition. Available at: <https://0-doi-org.pugwash.lib.warwick.ac.uk/10.1017/CBO9781316286777> [Accessed 24.07.2019].
- Wood, A. W. (2008). *Kantian Ethics*. Cambridge University Press.
- Zahavi, D. (1999). *Self-awareness and alterity: a phenomenological investigation*. Northwestern University Press.
- Zahavi, D. (2014). *Empathy and Other-Directed Intentionality*. *Topoi*, Volume 33, pp.129-142.

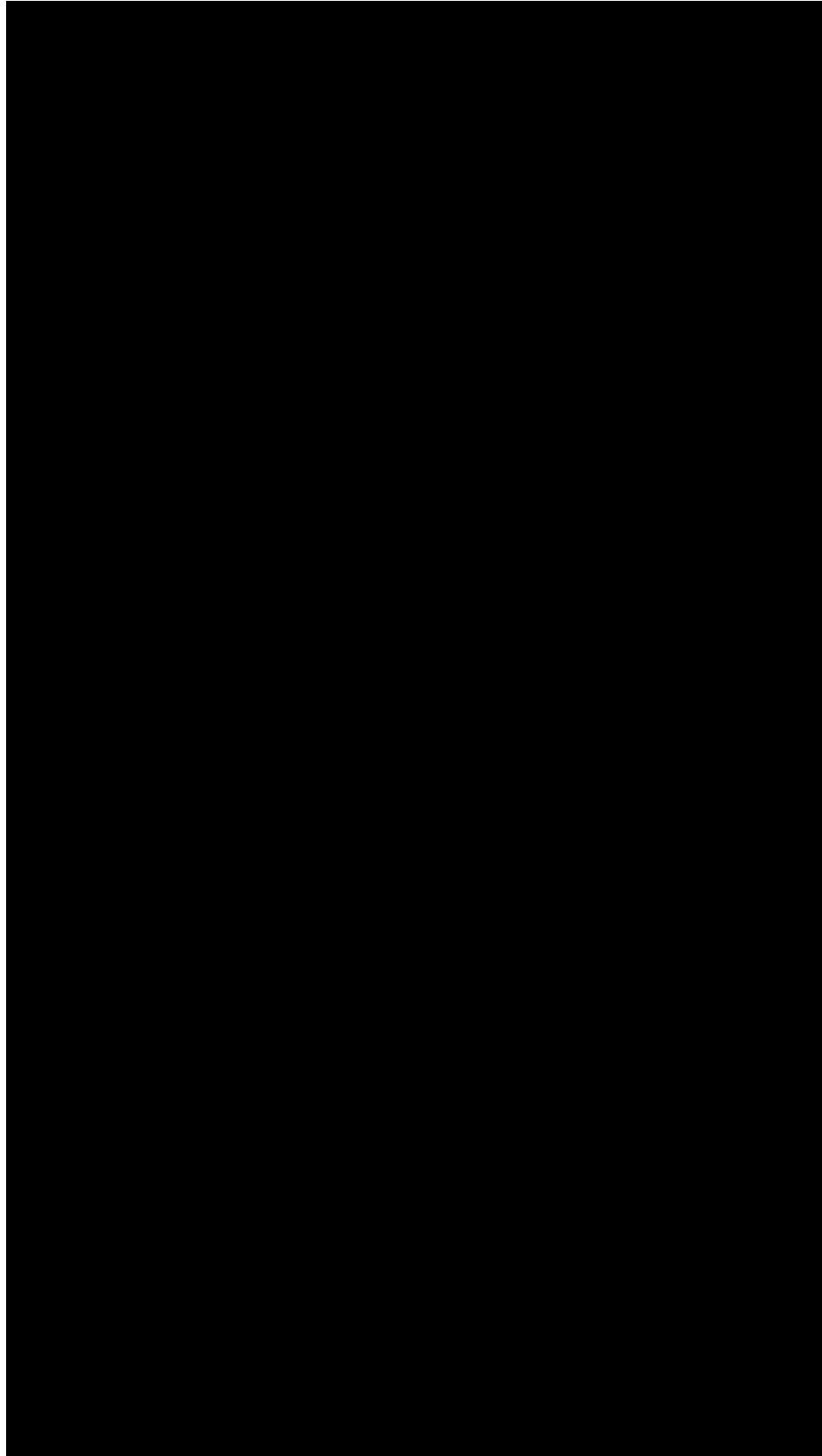
## FIGURES

Figure 1: Papyrus of Ani



©Trustees of the British Museum

Figure 2: The Lion Man



*©Trustees of the British Museum*