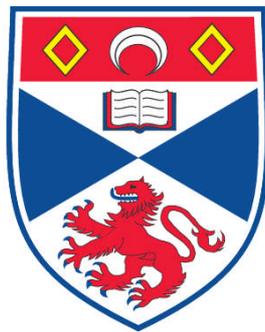


FREEDOM AS FAITH

Diana Aganey

**A Thesis Submitted for the Degree of MPhil
at the
University of St. Andrews**



2011

**Full metadata for this item is available in
Research@StAndrews:FullText
at:**

<http://research-repository.st-andrews.ac.uk/>

Please use this identifier to cite or link to this item:

<http://hdl.handle.net/10023/1883>

This item is protected by original copyright

**This item is licensed under a
Creative Commons License**

1. Candidate's declarations

I, Diana, hereby certify that this thesis, which is approximately 36,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in May 2010 and as a candidate for the degree of M.Phil. in September 2010; the higher study for which this is a record was carried out in the University of St Andrews between 2009 and 2011.

(If you received assistance in writing from anyone other than your supervisor/s): I,, received assistance in the writing of this thesis in respect of [language, grammar, spelling or syntax], which was provided by

Date 16/05/11 Signature of candidate

2. Supervisor's declarations

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of M.Phil. in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date 13/5/11 Signature of supervisor

3. Permission for electronic publication

(To be signed by both candidate and supervisor.)

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the electronic publication of this thesis:

Add one of the following options:

- i. Access to printed copy and electronic publication of thesis through the University of St Andrews.

Date 16/05/11 Signature of candidate Signature of supervisor

Freedom As Faith

Diana Aganey

Abstract

The belief in free will is something we are entitled to hold despite what determinism says. This notion however cannot be adequately explained or defended by traditional accounts of freedom amongst which, compatibilist and libertarian perspectives dominate the field of inquiry. I argue that an alternative approach is necessary to capture the full implications of what freedom as an idea contains and to establish this idea's validity, though one which exhibits none of the usual extravagances which philosophers so often pursue in their attempts at justification.

I would like to thank my supervisor, Professor Sarah Broadie for her acute understanding and careful suggestions which helped concretise my ideas and steer this work in the right direction. I am also very grateful to all those who supported me in this project –they know who they are.

Contents

1. Prologue	1
2. Chapter 1: The Gauntlet	
Part I – Introduction	3
Part II – Taking up the Gauntlet	
<i>Fatalism</i>	5
<i>Libertarianism</i>	8
Part III – The Alternative	12
<i>Preliminaries</i>	12
<i>Premises of the Argument</i>	12
3. Chapter 2: The Compatibilist’s Compromise	
Part I – The Unsuitability of Compatibilism	14
Part II – Compatibilist Solutions	14
<i>Strawson</i>	15
<i>Taylor & Dennett</i>	18
<i>Frankfurt</i>	24
Part III – Troubled by Compatibilism	32
4. Chapter 3: Kane’s Labyrinth	
Part I – Neglected Pathways	37
<i>Ultimate Responsibility</i>	37
<i>Plurality Conditions</i>	39
<i>Wills-setting</i>	41
Part II – The intelligibility Question & Kane’s Ultimate Proposal	43
<i>Kane’s Position</i>	44

Part III – Out of the labyrinth?	48
5. Chapter 4: The Myriad	
Part I – Retracing Steps	54
Part II –The Breakers of Freedom	55
Part III – Mending Freedom: The Alternative	60
6. References	71

Prologue

Condemned to death!

I have been living with this thought for five weeks now, always alone with it, always chilled by its presence, always bent under its weight!

Once, because it seems years rather than weeks, I was a man like other men. Every day, every hour, every minute had its idea... There was always a party going on in my imagination. I could think about whatever I liked, I was free.

*Now I'm a prisoner. My body is in irons in a dungeon, my mind imprisoned in an idea. A terrible, bloody, remorseless idea! I have only one thought now, one belief, one certainty: Condemned to death!*¹

The idea of one's freedom seldom provokes much deliberation in the mind of a typical human being unless of course, that freedom which so often is taken for granted somehow becomes threatened. One way in which this threat can arise, follows when an individual having done either a presumed or actual wrongdoing, suffers punishment in the form of coercive measures which aim to significantly curtail that individual's array of civil liberties as well as the general *sense* of freedom that standardly accompanies them. However, there is a yet more invasive way in which the sense of one's freedom can come under threat because it challenges the very notion of free agency and so throws doubt over the common assumption that persons ordinarily are morally accountable for their actions. Historically, such a way has been represented in differing guises through various *determinist* doctrines, which via their own unique methodologies essentially seek to establish one and the same austere conclusion: our sense of freedom is illusory. Despite our enjoyment of various civil liberties the larger picture of our reality suggests that we do not possess what is arguably the core kind of freedom, namely, one which sees our deliberations, choices, and subsequent doings as being products of one's will, a will shaped over time primarily by the individual himself, in other words a *free will* which forms the basis of this freedom and which determinism denies.

I argue that the full ramifications of such a position cannot be taken lightly when its assumptions are brought to their logical conclusion, and the passage presented above aims to demonstrate the general kind of worry that may surface, a worry that arguably initiates from a basic tension between the intuitive feeling that one is free and the reality of seeing this intuition undermined on account of its possible falsity.

¹ Victor Hugo, *The Last Day of A Condemned Man*; p.37.

In the case of the condemned man, he is overwhelmed by a certain despair that comes with knowing the irreversible finality that his life is about to end. The event is inevitable and the condemned's acknowledgement of this seals that inevitability in his mind which now fully recognises the fixity of his situation. Almost immediately, the awareness of one's former freedom intensifies, becoming more vivid and central; it is finally given the attention it deserves because at last a definite comparison can be felt between having it and having it taken away. Every living part of the condemned man is consumed by this one idea that he shall soon die and that nothing can be done to reverse this, more importantly, that there is nothing that can be done by *him* to change his circumstances – they are simply out of his hands.

Likewise, if the doctrine of determinism is taken seriously, a similar wave of despair and quite panic may be experienced upon acknowledgment that our lives ultimately unfold according to an inevitable course which the individual is largely impotent to influence. Under such conditions, it appears that there would be little to distinguish our situation from that of the condemned's with the exception, that he at least thought himself free before the sentence was passed whereas we strictly speaking would not even be entitled to that, since in our case the past, present, and future, would all be regarded as time-phases carried by the same causal chain (at least according to the *physical* version of determinism) which in theory renders our choices and subsequent doings to being its mere effects. In this case I believe it is we who would be imprisoned, mind and body, in a "remorseless idea" or better, in an inescapable conscious awareness that one is condemned to exist without genuine freedom over one's own self.

It is here then where the road begins to fork reflecting the conflict between commonsense intuition and philosophical theory. We think of ourselves as free agents who are creative, in control, and responsible for the majority of our decisions and actions, a thought plausibly fostered by the inner sense of freedom we feel, but this intuition must (i) be qualified and (ii) defended if there is to be any hope of preserving it against the determinist's assaults.² Such a task is by no means easy however, assuming we want our resultant theory of freedom to be comprehensive, viz., capable of justifying an account of agency that not only sees us as being responsible beings who act voluntarily, but also offers the possibility for intentional acts of *originality* in order to exclude the assumption that our lives are nothing more than the predictable outcomes of our environment or fixed nature.

Admittedly, not everyone seeks such a demanding theory of freedom and this perhaps is the main reason behind the diversity of approaches dealing with (i), broadly classed under Libertarian and Compatibilist perspectives together with some lesser known "alternative" or "non-standard" views. However I believe that a successful theoretical response should be determined by its ability to deal

² In particularly, the "hard" determinist.

with (i) and (ii) *jointly*, that is, not only should it provide a plausible explanation of the intuition in question, but it must also be capable of defending it in a way that would mitigate the threat from determinism without however the resulting account becoming obscure or incomplete.

In light of this condition, I contend that it is not possible to effectively satisfy both (i) and (ii) via a compatibilist account of freedom which (amongst other things) appears to leave itself exposed to the charge of incompleteness, about which more will be said in the main body of the work.

Libertarianism however in its traditional form arguably does little better, for despite being grounded on an intuitively persuasive set of principles, these in the past have come under scrutiny owing to their metaphysical vagueness and the difficulty of countering this objection has unfavourably led many proponents of the theory to focus almost exclusively on their defence at the expense of developing a more inclusive account of freedom.

It appears then that there is a plausible correlation between the nature of the account offered and the degree to which it can successfully be defended, or between (i) and (ii), and for this reason the principle aim of the paper will be to argue that what is needed is an alternative method to the standard ones proposed towards establishing a theory of freedom. Indeed, if we are ever to succeed in having anything more than an anaemic imitation of that noble human intuition, it is imperative I suggest that a cogent and theoretically encompassing alternative be developed to mirror that intuition faithfully.

What I have in mind here is an approach free from the idiosyncrasies and distortions fostered by the compatibilist's compromise or the metaphysical predicaments of the libertarian, presenting instead, a simple yet forceful defence of its principles which have the capacity to offer a truly comprehensive account of *freedom* as an idea valuable in its entirety and not merely from the consideration of securing agent responsibility. My own intention and hope is to be in a position to present the beginnings of such an approach here, though whether it succeeds in fulfilling the stated criteria I shall leave for the reader to decide.

The Gauntlet

I - Introduction

If there are certain basic conditions whose joint occurrence is logically sufficient for the occurrence of some event, then it *must* be the case that *if* these determining conditions jointly obtain the determined event occurs.³ There are *in fact* such basic conditions the determinist argues, and so the belief in free will is mere illusion. Determinism, broadly stated, is a doctrine which represents a body of related theoretical positions that hold to a greater or lesser degree that certain aspects of our lives are *inevitable* or beyond a person's control. Amongst its variants determinist positions can be *fatalist*, *theological*, *scientific* or *physical*, as well as *logical* and *psychological* but in every instance the same principle is maintained, namely, there is no free will of the kind ordinarily imagined.

With this assertion the determinist throws down his gauntlet and it is up to us to either choose to accept the challenge or refrain from it. It is my intention to accept it and see what can be done to remedy the blow cast, since I take the ordinary idea of a person's freedom (which in this present context is dependent on the notion of free will) a belief worth fighting for. I will begin then by presenting a basic determinist argument in the form of fatalism, followed by some preliminary responses to it. After that, I will introduce one form of libertarian perspective known as 'noncausalism' but will not dwell on it for long since I wish to later concentrate on a different kind of libertarian account which I find offers a more complete analysis of freedom than what has standardly been provided. Lastly, I will conclude the chapter by giving an outline of my own "alternative" position which will describe the type of freedom I believe human beings ought to possess and the way in which they could possess it without fearing determinism. Before proceeding however, I wish to explain the structure of the whole work by way of guideline for the reader.

There are four chapters, the first of which concentrates on introducing the challenge set by determinism and the typical responses used to counter it, concluding with my own proposal as regards this issue which will form the basis of my reasoning in everything that follows. The second chapter examines compatibilism as a key solution but argues that the image of freedom it conveys is simply too limited to be acceptable as a satisfactory account of our reasonable expectations. In light of what is perceived as the compatibilist's inaptitude then, chapter three extensively examines a more recent libertarian approach given by Robert Kane, and argues that a plausible theory of freedom must begin from principles similar to those evinced by the libertarian if it is to have any

³ Kane, p.6: *The Contours of Contemporary Free Will Debate*.

chance of sounding sincere. The fourth and final chapter, though departing from Kane's strategy quite markedly, uses some basic elements of his theory as a starting point from which to resume discussion of the alternative sketched in order to further develop its idea as a promising means of response to confront the determinist's challenge.

II -Taking Up The Gauntlet

Prior to the onset of more sophisticated kinds of deterministic theory that have gradually taken root, fatalism recedes much farther into history with one of its earliest forms presented in Aristotle's *De Interpretatione*. Since then, at least two other variants of this idea can be found, respectively known as *Logical Fatalism* and *Theological Fatalism*. For my part, I would like to concentrate on the Aristotelian version of the argument referred to by Mark Bernstein as *Representative Fatalism*, because it adopts the prevalent argumentative structure to be found in most subsequent fatalist accounts.⁴ I will also briefly discuss theological fatalism, though purely for the purposes of comparison since it adds a further interesting complication.

Fatalism

The fatalist asserts that everything which does occur was necessarily meant to occur and that which does not likewise necessarily just wasn't meant to be.⁵ The implications of this threaten human freedom, because the possibility of realising an alternative outcome is shown to be impossible.⁶ The fatalist's argument can be set out as follows:

1) There will be a terrorist attack on 12/12/2020 or there will not be a terrorist attack on 12/12/2020.

2) If there will be such an attack on 12/12/2020, then it was always true, factually true, that there will be such an attack on 12/12/2020. If however there will not be an attack on 12/12/2020 then, it was always factually true that there will be no attack on 12/12/2020.

⁴ I will be using Bernstein's formulation of Aristotle's account, though not his particular example. See Mark Bernstein: *Fatalism* p.70.

⁵ To clarify, fatalism can be distinguished from later forms of deterministic theory in that its account of necessity invokes neither *causation* nor *natural laws* as forms of justification.

⁶ Or if preferred, fatalism rejects the notion of *Alternative Possibilities (AP)* which is believed to be a key component of freedom.

3) If it has always been factually true that there will be a terrorist attack on 12/12/2020, then there has never existed a period of time when someone could have done something to prevent it. If however it has always been factually true that there will not be a terrorist attack on 12/12/2020, then likewise there has never existed a time when someone could have done something to make it happen.

4) Therefore, either no one (or nothing), at any time, could prevent the terrorist attack, or no one (or nothing), at any time, could make it happen that the attack does take place.

5) This implies that either, the occurrence of the attack is *necessary* or its nonoccurrence is a necessity.

6) Thus, fatalism appears to be true.

In the case of theological fatalism, the argumentative structure is very similar to the one just given with the exception that God is now introduced into the picture. The characteristic properties of omniscience, infinitude, and infallibility which a deity is said to possess, again seem to undermine the idea that human beings could be free in their actions. To use an example, if it is the case that God infallibly believed three hundred years ago that Molina would not go to this Friday's performance of *Don Quixote*, then it appears to be outside Molina's sphere of control to do anything to prevent this from happening. Put plainly, according to the theological fatalist she *cannot* fail to *not* go this Friday, since to assume otherwise is to assume quite implausibly that she has the ability to change a past belief in God's mind.⁷

Some Responses

There are at least two ways in which one could attempt to counter both types of fatalist argument.⁸ One of them employs the terminology of a *hard* and *soft* fact; the former refers to facts which in principle are incapable of being changed, in virtue of having obtained a 'rigid' existence, whilst the latter belongs to that category of facts which still have the propensity or *malleability* to change their content should some form of agency or circumstance intervene.⁹

⁷ The reason for the implausibility is that it would directly contradict some of god's assumed properties, especially those of omniscience or omnipotence, which logically at least undermines his status as a divinity while paradoxically implying hers.

⁸ Both are discussed in Bernstein's 'Fatalism', see, pp.71-80.

⁹ Bernstein uses the notion of *amenability* to characterise the difference between the two sets of facts, where hard facts are classified as "unamenable" to any of our powers and the soft conversely do possess that tendency, *ibid*, p72. I have however opted for '*rigid*' to define the nature of a hard fact, since I think it better describes its status of inalterability.

To clarify the distinction a little better two points require mention. First, this particular response questions the validity of the third premise in the above outlined argument from representative fatalism. That is, it questions the reasoning that it was *always factually true* regardless of what one does that (for example) a terrorist attack would take place, by assuming that it is possible to make it the case or rather make it a fact, that such an attack would *not* occur if for instance one were to successfully negotiate a deal with the terrorists which they resultantly came to see as being more advantageous than pursuing the planned course of action. Such a strategy then essentially presupposes that the event in question has the character of being a *soft* fact, since part of the nature of a soft fact according to Bernstein is that its determinacy is contingent on the events that occur *afterward* or more plainly, many soft facts become hard facts only after the stated event together with all other events necessary for its validation have taken place.¹⁰ The following example can help demonstrate this idea:

If Evgene Plushenko wins the 2010 Olympic gold medal for figure skating, and then *from* January 2014 onwards according to new regulations all Russian figure skaters are banned from participating in future Olympic competitions, this would (in the sense of it being a hard fact) make Plushenko the last Russian skater to have won the Olympic Games. However, what makes this event remain a soft fact between 2010 and just before 2014 is the possibility that at any time *within* this period, *if* the ban does not become introduced *then* it would no longer be a fact that Plushenko was the last Russian skater to have won the Olympic medal. Indeed, given this circumstance this would simply be an impossible claim to make, though it would remain a *hard* fact nonetheless that at anytime from 2010 if Plushenko does win the Olympics, he would always be the Russian champion *for that year*.¹¹

A second proposal to challenge the fatalist, in this case specifically theological fatalism, is to apply Hillary Putnam's argument for externalism in order to show that what God believes could itself be partially dependent on what a specific agent chooses to do.¹² In essence, Putnam's argument can be summarised by his provocative remark that "meanings just ain't in the head" suggesting that a speaker's private usage of a word like 'elm' viz., the belief one has concerning it, and its public or communal usage may radically differ, but that it is the latter which determines a word's meaning and consequently the speaker's belief state. Likewise, the alternative states that by rejecting the Cartesian conception of *belief* and appropriating Putnam's externalism, we are then able to say that the truth of any belief held by God is partially dependent on what the agent *actually* chooses to do. That is, if the determinants of a belief are actual external phenomena rather than a being's private

¹⁰ Ibid, p.72.

¹¹ For a different example see Bernstein, p.72.

¹² This is another response which Bernstein explicitly addresses before dismissing it, see p.77.

meaning, then regardless of what a deity may subjectively believe, that belief's content will largely be subject to what a person in fact decides upon which in turn implies the presence of a basic kind of freedom of action.

Considering the two responses, though they potentially offer some leeway to help overcome the extremities of the fatalist's position, ultimately I do not think they help dissolve the worry it causes. This I suggest is primarily because both solutions approach the problem rather abstractedly, which may not be the best means of dealing with it. The distinction posited between hard and soft facts may be a relevant one but perhaps not especially helpful; the ability these facts entail to change the course of certain events is a meta-ability, but one can side with the libertarian here and insist that what is needed is *actual* ability with all other forms simply being seen as insufficient to ground freedom.¹³ Sustaining this objection I assume is the intuitive idea that when we speak of a human being's *freedom* to choose or do something, we essentially want this concept to signify a *practical* and not merely theoretical capacity or for it to at least aim towards this. The second response likewise, is equally problematic for numerous reasons which cannot receive discussion here. However, amongst some immediate concerns worth mentioning rests the plausibility of externalism itself, whose validity must first be grounded before any assumptions against the fatalist can be made.¹⁴

Libertarianism

The libertarian argues that free will and determinism are irreconcilable; therefore, unlike the previous two solutions just examined, there is no intention to deal with a theory like fatalism per se but rather to introduce broad grounds from which a theory of freedom can be established.

Noncausalism is one such position since it aims to show how actions could be genuinely intentional rather than merely necessitated by causal circumstances, so to this attention will now be given.

¹³ I here develop an objection noted by Bernstein and others, see Bernstein, p74.

¹⁴ The disagreement between externalists and internalists remains a matter largely undecided to this date. However, even supposing that consensus as regards externalism's validity could somehow be reached, it is not clear whether the solution it offers against the theological fatalist is in fact a workable one. Here I am inclined to agree with Bernstein who suggests that an externalist conception of belief would only provoke controversy within the Judeo-Christian notion of deity because of the characteristics it attributes to God. Specifically, he argues that it might require one to relinquish the idea of a 'God' seen as possessing divine omniscience (ibid, p.77), I would further add that it may also require that we abandon the idea of his atemporality given that the view permits his beliefs to be partially determined by happenings in the actual world, and this raises a further issue of *how* temporal circumstances could influence that which is presumed to exist outside of time. One answer which could suggest itself then, may claim that God is not really atemporal – but then if he is neither atemporal nor omniscient as Bernstein implies, then what exactly does his divinity consist in? The difficulty of presenting a convincing response to such a question suggests to me at least, that externalism as a solution may not after all be a worthwhile one.

An account of freedom requiring that there be no cause present during the making and enactment of a free decision or action, or that at most actions may only be *nondeterministically* caused by other events, is seen as being 'noncausal' since it places no positive causal emphasis on the criteria governing a free action.¹⁵ The formal starting point for such a view is the idea that every action is or begins with a simple *mental action*, such as a decision or choice, which leads to the occurrence of other events, like the raising of one's arm for instance. This kind of action is known as a volition and describes an agent's willing, trying, or endeavouring to accomplish something – in this case, the movement of a given body part.¹⁶

Carl Ginet is amongst the view's defenders and in his 2002 paper provided a broad response to a series of common objections. Amongst these, is the criticism that this approach implies actions to lack explanation in terms of reasons, which Ginet addresses in the following manner: Is it the case, he asks, that we need a causal connection between desires, beliefs, and intentions of the agent (i.e. the explanans) and the subsequent action which they are seen to explain (i.e. the explanandum)? According to him the answer is straightforwardly negative. However, David Donaldson's challenge suggests that it should be stated positively.¹⁷

A causal explanation is crucial Donaldson maintains because it is possible for a person to (a) have a reason for doing some action and (b) to perform it successfully, though *not* for the reason one originally had but because another reason happened to introduce itself.¹⁸ It appears then that unless we allow the presence of a causal relation between reasons and actions, we cannot provide a satisfactory explanation of why a person performed a given action *because* of that particular reason rather than another. Put more formally, to derive the truth of "S did A for reason R", it must be assumed that R *caused* S's doing A.¹⁹

Ginet's response to this is to introduce a condition the obtainment of which would be seen as sufficient for the truth of the reasons explanation, but one which does not entail that the explanans in question *caused* the action explained.²⁰ This he provides by inserting a different set of linguistic apparatus into a conventional reasons explanation of the form "S A-ed in order to B" which once paraphrased yields "S's *purpose/intention/aim* in A-ing was to B". This latter sentence is then further restructured and fortified to yield Ginet's proposed sufficient condition to account for the truth of the original reasons explanation, viz., we are now left with:

¹⁵ I borrow Randolph Clarke's definition from his discussion in 'Libertarian Views: critical survey of noncausal and event-causal accounts of free agency' pp.357-8.

¹⁶ Ibid.

¹⁷ Ginet, pp. 387-8: *Reasons Explanations of Action: Causalist versus Noncausalist Accounts*.

¹⁸ Ibid.

¹⁹ Ibid.

²⁰ Ibid.

Concurrently with the agent's A-ing S *intended* of that A-ing that by it (and in virtue of its being an A-ing) he would B (or would contribute to his B-ing).²¹

An obvious concern however, is that introducing 'intention' into an explanation of reasons doesn't seem to cancel Donaldson's main assertion. More straightforwardly, the question that needs to be asked is whether the concept of intention is *satisfactory* to provide an unambiguous account of action? As Mele and I argue, it is not. The presence of a particular intention may certainly be used as a *possible* reason to explain the occurrence of a given action, but it clearly cannot guarantee that it is *that* particular intention which produced it. A hypothetical situation suggested by Alfred Mele can help elaborate this reasoning. He asks us to consider an agent – call her 'Alice' – who has two concurrent intentions, **P** and **K** when she goes to shut the window. We can say that **P** represents the intention to prevent a draft entering the room, while **K** is the intention to prevent a possible burglary. Now we are asked to imagine that a devious scientist is involved who makes it such that intention **P** has no effect on Alice's bodily movements, whereas **K** is left to function normally. In this case then, Mele insists that it is **K** which in fact explains Alice's action of shutting the window, but if this is true, then Ginet must be wrong to assume that the *mere presence* of an intention suffices to explain an agent's action.²² In other words, the concept of intention needs an additional feature for it to be capable of explaining why *this* and not some other intention is responsible for the action in question, and this feature according to Mele is causation. Consequently, he offers the following modification of the above sentence:

Concurrently with her A-ing S intended of that A-ing that by it she would B (or would contribute to her B-ing) *and* that intention caused her A-ing.²³

In response Ginet sees Mele's position as question-begging, in that it appears to merely assume that intention causes action when realistically we just don't know whether this is so. It may be that this is true, but as yet, we simply do not fully understand how mental states are realised and so whether they *do* in fact cause the relevant voluntary bodily exertions. For my part, though I am in agreement with this observation, I do however think that Ginet perhaps fails to appreciate the central motivation of Mele's objection. The worry is arguably that intention on its own just doesn't do enough to prevent the possibility of the subsequent reasons explanation being mistaken, whereas supplementing it with a causal element (even if the latter's validity has not been verified) *may* give it more weight. Hence, the assumption of a causal factor can be said to function as a

²¹ Ibid.

²² Ibid, p.389 (I have slightly altered Mele's example).

²³ Ibid.

theoretical tool which *completes* an ordinary reason's explanation.²⁴ At this point I will end my discussion of noncausalism, but before I do I would like to say a little about the potency of libertarian perspectives in general.

Much paper has been spent over the years trying to fight off opposition to the theory, and though this is normally an expected consequence of holding a conflicting position, sometimes however it could be an indication that something is either not working or lacking within the theoretical armour of that position itself, or perhaps both. Ginet seems to betray a similar line of thought when towards the end of 'Can the Will be Caused' he writes:

"The argument above only removes another of the confusions obscuring the free-will-determinism problem. The *real question*...is whether a vast addition to our knowledge about the physical causes and effects of the minuter internal processes of our bodies could possibly turn out to be incompatible with regarding any of the behaviour of those bodies as expressing wills.."²⁵

I take this paragraph to imply that despite the employment of various ingenious theoretical endeavours to secure an account of freedom, the most fundamental questions remain unexplored because they are still largely unexplorable. What this means then, is that the libertarian's primary mode of explanation is highly dependent on conjecture – a tendency that is prevalent amongst other theorists too, but which arguably affects libertarians more since they strive to provide a maximal account of free agency.²⁶ To make up for the lack of *evidential* support then, the libertarian relies on some questionable metaphysics in order to secure the theoretical basis of his position only to later spend time and energy on having to defend that same metaphysics. A distinctive consequence of all this, is that the majority of contemporary libertarian perspectives, despite their admirable intentions, fail to offer a truly comprehensive account of human free will, especially, one that is sensitive to an ordinary person's intuitive understanding of "freedom" as a practical everyday concept.

²⁴ Ginet however expressly disagrees with such a conclusion since he believes that in the event of it being discovered that no causal connection exists between intentions and actions, this would not lead us to abandon giving reasons explanations of the kind he suggests, or that the resultant account would be incapable of explaining an agent's actions (see, p.390). Again, this may be true, but considering that such a discovery has not been made yet, it seems more problematic to insist that an explanation in terms of intentions *alone* is better than one which is also accompanied by a causal component to help justify the relevance of a given intention.

²⁵ Ginet, p.55.

²⁶ I specifically refer here to the desire on behalf of the libertarian to defend a notion of freedom whereby agents are seen as possessing a will, capable of producing 'original' action which is seen to stem from the will's ability to be a cause, rather than an intermediary effect between the agent's fixed nature and environment, and the action which follows. See Honderich, 2002, p.473.

III - The Alternative

Preliminaries

I wish to now provide an initial sketch of a position that I will go on to defend in the last chapter of this work, though whose influence may be felt throughout all its subsequent parts. Beginning first with what should be apparent; *freedom* and *free will* are not necessarily identical concepts. 'Freedom' commands a wider definitional scope than 'free will' because it bears the potential to refer to a variety of different things and the capacity to apply to different species, such as dogs for instance which can be regarded as having freedom in the sense that they may sleep and eat whenever the mood takes them. 'Free Will' on the other hand is an idea that theoretically identifies two distinctively human features, firstly, a *will* and secondly, the *ability* to use that will as a medium for implementing various freely formed purposes. To this I will add a third feature or consequence of having such a will, namely, it *represents* the existence of a special kind of freedom which again seems to be exclusively human in nature in that it asserts a propensity to be creative and spontaneous. In this specified context then, freedom and free will become synonymous terms whereby the presence of the latter, establishes the reality of the former.

Premises of the argument

The following reasoning structure is merely provisional, aiming only to state the main premises of the alternative which will later be systematised into a more formal argument.

1: Freedom takes two forms: The first, I will call *organic freedom* the second I will refer to as *acquired freedom*. Organic freedom is what human beings are arguably endowed with from birth, and implies a natural limit with respect to our capacities. That is, it exists subject to the constraints of the laws of nature together with those of a human being's distinctive psycho-physical tendencies. Acquired freedom on the other hand, finds its origin exclusively within a person's will, and in order for it to be attainable that will must be free. I take this form of freedom to be superior in kind to its organic variant.

2: When a human being deliberates on a certain course of action, he first and foremost encounters the organic form of his freedom. This freedom presents him with a particular set of parameters within which he is potentially free to make his decision. Now, within these parameters he may either decide to obey the dictates of his formed nature or to try and overcome their influence through sheer strength of will so that with time he succeeds in *recreating* his character almost completely. If he fulfils this endeavour, then such a person can be said to have reached the acquired form of his freedom by rejecting the particulars of his organic state.

3: To me it seems intuitive to argue that human beings are in possession of an organic kind of freedom, but that they also have the capacity to reach a more elevated form of existence. Since one's acquired freedom is a direct consequence of having relinquished the particulars of the organic freedom, and because the transition itself arguably demands effort on behalf of the acting agent, this then secures that freedom's status as being superior in worth. It also implies however, that though everyone has the potential to reach it, not everyone will.

4: Now, a fair objection at this point would be to state that the argument does nothing to show how we can establish the possibility of this acquired freedom in the first place, especially in light of the determinist's challenge. In response, I add my fifth and central premise.

5: I began this chapter by stating that determinism argues free will to be an illusion. To this I reply, perhaps so, but ultimately we can never know either way. My belief is that any argument one suggests in defence of freedom, regardless of how cogently it is presented, will always be vulnerable to an attack. My proposal then is that we trust our intuitions²⁷ by taking a decided leap of faith to believe in our freedom regardless of what the determinist says, for it is better to believe that we are free than to reconcile ourselves to thinking that we are not.²⁸ The matter arguably does not permit the attainment of certainty or proof; why not choose to believe then? By doing so we gain more than if we do not, for one thing, we would be leading our lives according to the conviction that one is *free*, a conviction which seems fundamental for making our experience of living to be meaningful and exciting rather than futile and mundane.

Now, if freedom essentially relies on faith so that it can never be known with certainty whether or not we in fact have it, what is stopping us from choosing to believe rather than doubt? For if we believe and this belief (on some level impenetrable by us) happens to be *true*, then we will have done the right thing. If however, it happens to be *false*, I argue that one is none the worse for it, since at least then one has succeeded in *feeling* free. True, in such a case the feeling would indeed represent no more than an illusion (without our acknowledgement of this of course), but I ask, is it not better to be able to possess at least the illusion of freedom, than nothing at all? For an illusion consisting in a feeling is still worth something, it is still the feeling of freedom however false, and belongs exclusively to him who has taken the necessary leap of faith thereby securing himself this small victory over him who feels nothing.²⁹

²⁷ For those of us that have them of course, which I insist is anyone who has ever made a decision or realised some desirable action.

²⁸ This argumentative strategy takes its inspiration from Pascal's wager concerning the existence of god.

²⁹ In this final part of the argument, I am implicitly assuming that because one may never be in a position to find out whether one's leap of faith was justified, it seems reasonable to argue from the point of view of

The Compatibilist's Compromise

I - The Unsuitability of Compatibilism

The compatibilist maintains that the free will issue is no black and white matter and argues that it is possible to find a middle ground on which the determinist and libertarian can coexist. Indeed, a reasonable question that suggests itself at this point, is why we cannot just side with the compatibilists and leave it at that? For my part, I have contemplated this question for some time, struggling to find the right answer. The difficulty here is that on immediate viewing, it does seem that the issue has been effectively dealt with, appeasing all parties involved. Yet upon closer inspection something feels amiss, or at least ought to feel amiss I find, though what exactly is very hard to pin down. Nevertheless, I shall do my best in this and the next two chapters to do just that, by offering three broad but related reasons for why I believe the compatibilist's "solution" to be largely bogus. These reasons are as follows: (1) the account of freedom suggested is *extremely* limited. (2) Consequently, there is no possibility for persons to be *ultimately* responsible. (3) Focusing primarily on the responsibility component of free will makes the resultant account of freedom *incomplete*.

I will address each of these reasons separately starting with (1) in the present chapter, and dealing with the other two in chapters 3 and 4. But before commencing with this task, I would like to introduce three prominent compatibilist arguments to ensure that this position receives a proper hearing before the sentence is passed.

II - Compatibilist Solutions

Compatibilist accounts are diverse and many, but the one thing the majority of them have in common is the idea that the kind of freedom we want or need in order to keep the notion of responsibility intact does not have to be relinquished even if determinism is true. One of the central preoccupations of the compatibilist is to show that the concept of *choice* together with that of *alternative possibilities* or AP, where the former is seen to entail the latter, can still be preserved.³⁰ The other related concern, is to show how the notion of *voluntariness* could still be applicable, since without actions being voluntary the choices made cannot be assumed to be genuine ones. These

simple human psychology and wellbeing, that it is better to live *with* the feeling of freedom (whatever the truth concerning this issue may ultimately be) than to have nothing at all.

³⁰ There are exceptions however; Harry Frankfurt does not see AP as necessary for grounding free will, which will be discussed in due course.

then are the issues that interest compatibilists, and though the above aims may not always be stated explicitly, they are nevertheless the motivational force behind all manner of defence strategies.

Strawson

In his *Freedom and Resentment* essay, P.F. Strawson offers an argument from simple commonsense without resort to fiddly logical and/or metaphysical justifications that are standardly preferred.³¹ At the heart of Strawson's approach lies the conviction that the backbone of social practices – as those of rewards and punishment, praise and blame – is comprised of basic *human attitudes* which he argues "express our natures" and without which these social practices would not exist. Amongst the various human attitudes, we find most commonly those of approval and disapproval as well as of gratitude and resentment, to take just a few, and their presence is both natural and integral in our everyday lives.³²

Now it may be asked what all this has to do with establishing the compatibility of free will and determinism? According to Strawson, everything. It is in these attitudes that he thinks a reconciliation can and ought to be found because it is they which form the basis of the dispute between the "pessimists" and "optimists" or more standardly, the incompatibilists of libertarian origin and the compatibilists.³³ To understand this better we need to take a closer look at his argument.

The larger portion of Strawson's paper is spent examining the attitude of *resentment* and the conditions which lead either to its presence or withdrawal. Resentment is viewed primarily in terms of how one reacts to the projections or characteristics of other's wills towards *us*, so, for example, if I am sitting on a bus and someone happens to fling their arms mid conversation causing my glasses to come crashing down, it is left up to me to determine whether I should feel resentment in this situation by considering whether the action was intentional or accidental, in other words, whether the person used their will deliberately to inflict damage on me. This analysis of resentment is then supplemented with two further categories which help set the framework for his explanation, consisting of (I) *Participant reactive attitudes* (II) *The objective attitude*.³⁴ (I) describes the natural range of human reactions to the good/ill will/indifference of others towards us as shown in their

³¹ I could of course, have started with A.J. Ayer's or Hume's compatibilist accounts, but my aim here is to discuss contemporary rather than classical contributions.

³² Strawson, p.25.

³³ Because the libertarian believes that the truth of determinism leads to the irrelevance of such concepts as moral obligation and responsibility, and to the meaninglessness of practices as those of punishment and blame, for this reason Strawson permits himself to call this position pessimistic (see, p.1).

³⁴ Ibid, pp.6-10.

attitudes and actions. (II) being the antithesis of the participant attitudes, is effectively defined by the absence of any reactive feelings of the kind available in the above example.³⁵

Of main importance here, is that the objective attitude carries much significance for Strawson. First, he sees it as an essential commonly-reverted-to means when dealing with those members of society who may be classed as 'abnormal' or 'immature' adding that, under such circumstances "We shall not feel resentment against the man he is for the action done by the man he is not; or at least we shall feel less."³⁶

Secondly, and more importantly, he is keen to stress that we often do not flinch from employing the same objective attitude towards the behaviour of those standardly considered as 'normal' or 'mature' as a form of "refuge" or an "aid to policy" and this he argues, may present a possible cause for concern.³⁷

Subsequently, two complimentary sets of questions are posed:

(a) *If determinism is true would/should it lead to the end of reactive attitudes such as those of gratitude, resentment, and forgiveness; of all reciprocated adult loves; of all the essentially personal antagonisms?*

(b) *In other words, could or should the truth of determinism lead us to regard others whom we consider 'normal' in the objective way?*³⁸

According to Strawson the answer is a categorical No! To take resentment for example, it is implausible to assume that the truth of determinism would result in every intended injury being looked upon as an act of ignorance or seen to contain "overriding reasons" in virtue of which our resentment would either be mitigated or removed altogether. Specifically, it certainly wouldn't and shouldn't lead to persons being viewed through the "objective attitude" telescope that ordinarily stays reserved for the abnormal, and only occasionally becomes applicable towards other members of society as a temporary coping strategy. The simple fact according to Strawson, is that human beings are committed to maintaining an active interest in ordinary attitudinal exchanges, and this commitment being both "thoroughgoing" and "deeply rooted" implies that the possible truth of a theory such as determinism, can never seriously threaten what is effectively a way of life.³⁹

So this then is the plane of friction over which libertarians voice their discontent. Their anxiety consisting, as Strawson sees, in the perfectly legitimate concern that by accepting determinism we effectively accept a world dominated by the objective attitude where this attitude is no longer

³⁵ Ibid, p.10.

³⁶ Ibid,p.8.

³⁷ Ibid,p.9.

³⁸ Ibid, pp.10-11.

³⁹ Ibid, p.11.

restricted to those special cases, but delivered *universally* as the primary means of ‘dealing’ with offenders. Yet it is precisely here where the theories of determinism and libertarianism ought to merge, according to Strawson. The determinist must not turn a blind eye to the centrality and resilience of our human reactive attitudes, and the libertarian on his part must accede that it is this and only this narrow-mindedness exhibited by the theory that makes the view it espouses seem so deficient and hostile. However, once both sides acknowledge the perennial relevance of our human attitudes, the cause of the controversy between them is arguably resolved.⁴⁰

To conclude, for Strawson determinism does not present a challenge because the reactive attitudes we have are not likely to become extinct and replaced by the objective standard of assessment. And since it is in virtue of these attitudes that moral assessment becomes and remains possible, the libertarian can rest in peace that such practices as those of praise, blame, and punishment, originating from the notion of agent accountability, will continue to thrive given what we understand about the tendency of human nature.

Before I leave this particular discussion, I would like to comment a little on its overall effectiveness. First of all, I will say that of the subsequent two approaches I am still to present, this I feel to be the most acceptable because of its lack of pretension. That is, it does not realistically try to solve the problem posed by the determinist, but rather to deny its practical relevance. This I admit it does with plausible intuitive merit, by frankly asserting the resilience of our day-to-day social interactive habits. Having said this, two things remain worrisome. Firstly, I do think that Strawson’s approach somewhat trivialises the libertarian’s intentions, arguably, the aim is not *just* to secure the social relevance of responsibility and the practices it fosters, but to establish its *actual* reality. Secondly, I suggest that the main assumption concerning the innocuousness of determinism with respect to our daily practices is also highly questionable. It is not unreasonable to imagine that if tomorrow it somehow became known that a significant bulk of our actions proceed from causes other than those originating in the agent’s will, that we would in time feel inclined to adapt our behavioural attitudes to fit this new epistemic revelation. This point is especially appropriate I believe, and one that cannot be explained away by Strawson’s appeal to human nature or the inherent preposterousness of even suggesting such an alien measure, unfortunately though, it cannot be further dealt with here.⁴¹

⁴⁰ Ibid, pp.21-5.

⁴¹ Strawson explicitly address this type of objection on p.13 of *Freedom and Resentment* when he acknowledges that some may feel that the central question here concerns what may be the *rationally* correct thing to do, should determinism prove true? However the merit of such a question is immediately dismissed by him for its failure to “grasp the purport of the preceding answer” that is, “the fact of our natural human commitment to ordinary inter-personal attitudes” a commitment that “is part of the general framework of human life [rather than] something that can come up for review as particular cases can come up for review

A notably different route taken to resolving the tension has been offered by Christopher Taylor and Daniel Dennett in *Who's Afraid of Determinism? Rethinking Causes and Possibilities*. The authors' approach essentially relies on the use of possible worlds discourse to demonstrate (amongst other things) how we may continue to meaningfully employ the notion of *alternative possibilities* or AP as I shall refer to it, despite what determinism argues. To see this we need to take a closer look at their argument.

The incompatibilist accepts the following two assumptions; 1) in a deterministic universe it is not possible to utter sincerely "I could have done otherwise". 2) In such a universe one can never really take credit for having caused an event, since all events are supposed to be predetermined by conditions present during the universe's birth.⁴² As a result, there is an understandable tendency to want to reject deterministic theories because they entail the absence of possibility to do *otherwise* in a given situation, and because they undermine the idea of agent causation which sees persons as capable of producing various events in the world in a way that makes our actions instrumental to their occurrence, rather than merely subsidiary. According to Taylor and Dennett however, these fears can be alleviated with some careful analysis which reveals a way to preserve AP while also permitting a non-trivial sense of agent causation to exist.

The following reasoning is provided to demonstrate the retainment of AP:

1) Take ϕ as representing the sentence "Joubert lands the quad"⁴³

2) Now, let \mathbf{K} stand for the set of physically possible worlds which are *identical* to the actual world at some time t_0 prior to the landing of the quad, *and* let us assume that Joubert fails and that determinism is true.⁴⁴

3) If this is the scenario then the incompatibilist's fear is justified, since by matter of fact, ϕ cannot be said to obtain or be true of any world in set \mathbf{K} ($\sim\exists x \phi$), because \mathbf{K} contains only one world, viz., the *actual* one.

*4) However, the consequences are misleading, and the reason for this is that 2) relies on a *narrow method* of selecting \mathbf{K} , though one which arguable need not be observed each time. All that is required to remedy the seeming lack of AP, is to admit into \mathbf{K} worlds which despite being very similar

within this general framework". I maintain though that Strawson's faith in the durability of this 'commitment' might be called upon by the determinist, to supply additional justification.

⁴² Taylor & Dennett, p.257.

⁴³ I hope the reader does not mind too much my preferred reliance on figure-skating terminology in helping to illustrate examples.

⁴⁴ I.e. fails to land the quad in the actual world.

to the actual one, nevertheless differ from it at t_0 in some microscopically imperceptible ways. According to this new scenario then, we may discover that ϕ is possible after all or that, $\Delta x \phi$, despite determinism obtaining.⁴⁵

The idea introduced here is that “possible” does not have to be interpreted exclusively as “X is possible *given that the conditions are identical*.” In other words, it is being questioned whether when one talks of having alternative possibilities, the referent of *possibility* must always be the narrow method. So when for example I say “I could have aced this maths test” (having just received a B) the interpretation of my belief in thinking there was an alternative open to me, is *not* explained by the idea that provided there is a set of identical circumstances to the actual ones, I know there exists the possibility that I could have gotten an A. Rather, the authors contend that when we speak of possibilities, we often tacitly assume that given circumstances *very similar* to the actual ones, there is the distinct opportunity for things to have turned out otherwise than they actually did.⁴⁶

Regarding the second related concern, namely, *causation*, here Taylor and Dennett suggest that the problem lies in our tendency to conflate causal *necessity* with causal *sufficiency*, pointing out that determinism primarily concerns the *sufficiency* component whereas causation in general assumes a necessity factor.⁴⁷ This implies then, that determinism’s possible truth would have but a minor effect on the validity of our causal assessments and to understand this point better, we can use the following example:

1) According to determinism the precise conditions of the universe following the big bang (call them **C**) were causally *sufficient* to produce WWII (call this **W**).

2) According to Taylor and Dennett though, there is no good reason to think that this means **C** caused **W**, for despite being sufficient for it, **C** is not strictly speaking necessary.⁴⁸ That is, the

⁴⁵ For a near duplicate of this explanation see Taylor and Dennett, pp.266-7.

⁴⁶ In stating this assumption, Taylor and Dennett particularly wish to attack John Austin’s claim that when one talks of possibilities one means “conditions as they precisely were” since several lines later, he appears to contradict himself by talking of conducting “further experiments” to hole a previously missed putt for instance, in order to satisfy the belief that there was opportunity for a different outcome to occur. The accusation waged here is that Austin is guilty of equivocation when discussing possibilities, since the idea of staging further experiments implies the use of near-duplicates of the putt etc., rather than having identical ones. But if that’s the case, then it appears Austin is not as committed to conditions being *as they precisely were*, and would admit “X holes the putt” as possible, if in situations *similar* to the actual one X actually succeeds in doing this (See pp. 258,267). However, of relevance here is the consideration that Austin’s remark about doing further experiments is open to an alternative interpretation, one for example where it could be taken to imply the introduction of identical samples if we assume something like an ideal world where such a thing may be possible. In this situation, it would be possible then to argue that the opportunity of holing the putt implies the presence of identical conditions.

⁴⁷ Ibid, p.271.

⁴⁸ In honesty, I find myself troubled by this assertion, that is, if I understand it correctly. In what sense exactly do the authors mean that **C** despite being sufficient is “hardly necessary”? Arguably, in one very fundamental sense it *is* necessary, for it has all the desired qualities to produce **W** in the event that no other (necessary)

possibility that WWII may still have broken out even if the universe had settled into a different pattern of conditions (i.e. those which lacked sufficiency but possessed necessity instead) shouldn't be ignored. Amongst some of the *necessary* causes of the war then, we may find for instance, Germany's desire for Lebensraum; Britain's and Russia's underestimation of Hitler; and more fundamentally the birth of Hitler.⁴⁹

3) As soon as we admit there could be other (in this case necessary) causes at play apart from the sufficient ones, the incompatibilist's fears appear groundless since there is nothing in our causal explanation that alludes to the multitude of factors present within the universe billions of years prior to the occurrence of **W**. The incompatibilist therefore, misunderstands the nature of causal inquiry if he thinks that under determinism **C** causes or explains **W**.⁵⁰

Having explained Taylor's and Dennett's solutions to AP and causation, I will now mention two further points and some concluding remarks made by the authors in support of their position before stating some general thoughts as regards this approach.

First, with respect to causation in general; determinism it is stressed, is quite compatible with there being no single qualifying cause of a given event, but rather a "confluence of dozens of factors that jointly sufficed" [to cause X] though none of which individually prove essential for it. The implication here being that no 'real' or 'ultimate' cause may be found when examining the occurrence of a particular event – an idea which the following law school riddle aptly demonstrates:

Fred is hated and wanted dead by all in the French Foreign Legion outpost. In the night before his trek across the desert, Tom poisons the water in his flask; then Dick oblivious of Tom's sabotage pours out the poisoned water and fills it with sand granules. Lastly, Harry arrives to pierce a hole in the flask, so that all the "water" may eventually seep out. The next morning Fred begins his long journey and eventually discovers his flask to be almost empty, filled with nothing but a little sand at

factors obtain to produce it. That is, the implication is not that **C** happens to be the sole bearer of qualities capable of causing **W** but rather, in the event that other circumstances do not materialise the sufficiency of **C** alone will be enough for its realisation. Furthermore, even supposing that there *are* other circumstances present that could just as easily lead to **W** but that **C** is present too, in this situation I would argue that we could never know exactly whether it was **C** or those other circumstances that ultimately resulted in the occurrence of **W** (perhaps it was a joint effort?) and for this reason it appears rather clumsy of Taylor and Dennett to downplay the potential importance that **C** carries.

⁴⁹ The authors of course make a valid point, but it seems that the assumed necessity conditions which make the idea of human agency possible could prove unhelpful if it is argued that their existence is also the outcome of deterministic causes.

⁵⁰ For a different example of this see p.272.

the bottom – not water, and not even poisoned water. He dies of dehydration. Who caused his death?⁵¹

Second, in order to reinforce the importance of *causal necessity* as an additional consideration, we are asked to reflect on a typical model of computer chess, since computer programmes are perfect paradigms of deterministic virtual universes. Next, we are told to imagine two chess programmes, A and B where B is seen as winning contrary to expectations, and so naturally provoking our curiosity as to what could have caused the sudden victory. Merely claiming that B won because of the initial state of the computer sheds very little light on the matter, for though the totality of this virtual universe at prior moments may have been *sufficient* for the win, we still want to know which features in particular were *necessary* for it.⁵² Could A have been infected with some kind of virus causing it to make a false move where it would normally have gained an advantage? Or perhaps B's victory is purely coincidental? Further yet, maybe B harbours an "idiosyncratic island of brilliance" which once understood will allow us to know which future circumstances would enable it to have more wins?⁵³

The point of emphasis here is that when we are faced with these kinds of scenarios, the natural inclination is to seek out a cause *other* than the one given by the underlying sufficiency conditions of the system or programme in question. Our natural tendency to want to pinpoint a necessary feature that could justly be attributed with having produced B's win is just as important, as knowing that the totality of the system which B occupies is sufficiently conducive to guarantee its victory. By knowing the necessary causal circumstance, we are undeniably placed in a better epistemic state to judge the quality of the win and to be capable of making future predictions. Causal necessity it would seem then, is a factor that continues to be of great significance even if we do occupy a deterministic universe, because it happens to be the one we consult when seeking out answers.

⁵¹ Ibid.p.266. Reluctance to veer off course prevents me from offering a comprehensive rebuttal of the above scenario. Nevertheless, it should be obvious I think that one way of selecting a culprit would be to assess the relative actions done to the flask in order of their believed degree of malice, and the expected probability of success in accomplishing the intended aim. So, in this case, it is possible to assume that Tom's action was the most malicious and likely to have the highest probability of achieving its aim, since poison victims seldom recover when stranded in the middle of a desert. Harry's action conversely may be viewed as the least spiteful, for there is still a reasonable chance that despite the pierced flask Fred may still have had enough water leftover to get him through the journey (if only it wasn't replaced by sand of course) or, perhaps, if Fred had been of stronger constitution, he may have managed without water long enough to find his way out of the desert. The other alternative in this situation is to identify the trigger of his death, which we know was dehydration, this would certainly rule out Tom as the culprit even though his action may be judged as the most potentially harmful. The point of all these solutions however is simply to show that though an event *can* be necessitated by more than one cause, this does not always or even often imply that a single 'core' or *primary* cause doesn't exist to account for it, it only suggests that this cause is usually difficult to isolate.

⁵² Ibid.p.272.

⁵³ Ibid.

The authors conclude their discussion with one final example aimed at reaffirming their approach's ability to ensure the survival of AP together with that of agent causation:

A man is falling down an elevator shaft with no idea which of the possible worlds in a given set he in fact occupies. The one thing he does know is that within *all* the worlds in the set, he will assuredly land at the bottom of the shaft in virtue of the law of gravity. The landing then becomes *inevitable*. But perhaps *death* is not inevitable? There could after all be possible worlds in which he lands virtually unharmed, those in which he finding himself in a crouching position for instance rather than lying prostrate.⁵⁴

The suggested alternative of obtaining an optimistic outcome implies that it is neither meaningless nor futile to plan one's actions according to the assumption that living is possible. And though it may still be the case that no sufficiency conditions exist to *guarantee* survival, at least one can improve one's chances by taking whatever steps might be necessary to realise this. In other words, under a deterministic system we appear to have discovered some "elbow room" to manoeuvre.⁵⁵ Before closing the discussion the authors reminisce on Robert Kane's remark about our human desire to take full credit as the creators and causes of change in the world, with respect to this they remain adamant however, insisting that even something so relatively minor as the present essay could not have obtained without their active *participation*, viz., without their effort. This then offers plausible grounds for assuming actions to contain at least some "origivative value" for without the agent's *input* it is safe to say that no essay could have been produced, and more fundamentally, the world as we know it would have turned out markedly different had human effort not been exerted.⁵⁶

Taylor and Dennett clearly regard it an attribute of their method that they are able to employ possible worlds discourse so effectively. I however remain sceptical about the merits of this technique. Specifically, it remains unclear to me whether anything more substantial is being achieved via possible worlds criteria, as opposed to using standard modality assertions. If the goal is simply to show a logical way in which an alternative outcome could be feasible, then traditional modality statements arguably accomplish this already. We could just as easily say "*had circumstances been slightly different Q would have been possible*" instead of supplementing the sentence with the notion of a possible world, as if that world described some ontologically distinct

⁵⁴ Ibid.pp.272-3.

⁵⁵ Ibid.

⁵⁶ Ibid.

entity. Rather, the very concept of such a world remains to be both puzzling and vague, and though useful in certain circumstances⁵⁷ here seems to have been stretched far beyond its capacity.

It seems that when speaking of human freedom, what is desired is *real* rather than merely logical or hypothetical possibilities. That is, we do not want these possibilities to be tucked away behind a theoretically contentious notion of a possible world, but to have them here, right here in the one we occupy right now with the circumstances being just as they are we want to have the confidence that there was always an opportunity available for things to have turned out differently than they did. A possible world I argue cannot compete with the actual one, the actual world is all that matters because from the practical perspective it is the only one we happen to exist in.⁵⁸

Supposing however we indulge Taylor's and Dennett's argument by assuming the possible worlds discourse to yield a satisfactory account of AP, is there anything troublesome about such a move? My inclination is to say that potentially there could be, and the cause stems from the authors' requirement that we abandon the *narrow method* when selecting our sets of possible worlds, which may provoke a personal identity issue. For if the circumstances in a given possible world in which let's say I don't die after falling down an elevator shaft, happen to be *very similar* but not identical to those in the actual world, then the question to ask oneself here is whether I am still the same person viz., whether the slight shift in circumstances manages to leave my identity unscathed? It could be argued that one's identity remains constant throughout all the possible worlds one may inhabit, however, there is also opportunity to contend that if the possible world in which I don't die lacks absolute identity with the actual one, then so do I – or rather why not me too? And if so, then who or what survives the fall? Put differently, in what sense did I, Diana Aganey, have AP if the survivor in question is (1) In a different world from the actual one and more importantly (2) In that world circumstances are just that little bit different, giving cause for thinking that the 'Diana' of that world is also quite different from the Diana in the actual one.

Of course more could be said about Taylor's and Dennett's approach, but I will close this discussion now with one final thought. The authors believe their account is capable of satisfying Kane's "origination" requirement to a considerable degree, and this I think is not an entirely mistaken assumption. In one important sense, we remain being the principle initiators of change in this world which can be viewed as our *accomplishment*, however, given that the current position accepts the possible truth of determinism, our role as 'initiators' becomes more questionable. How much of what I did was due to my own initiative and how much was a direct result of covert

⁵⁷ Ordinarily in the study of formal axiomatic systems of modal logic.

⁵⁸ I do not wish to enter the debate here between possible world realists such as David Lewis and those who hold a deflationary view of them like the one just described. What matters here is that the absence of consensus regarding this subject matter is acknowledged as a possible means of challenging Dennett's and Taylor's account of AP.

deterministic influences remains unknown. What if what I take as being my initiative is really nothing more than a necessary outcome of such influences? The point is that compatibilism seems to offer at best a very *restrictive* notion of origination which fails to capture what libertarians like Kane mean when they talk of our human need to be the causes and creators of change. In the end however, I believe that what all this comes down to is personal preferences and whether one is content to settle for such a humble depiction of human freedom as given by the present theory, and I for one am not.

Frankfurt

A final approach which I think requires attention because it continues to be classed as one of the most enlightening contributions to the free will debate, was offered by Harry Frankfurt in the late 1960's and early 70's.⁵⁹ In two complimentary papers Frankfurt addresses the familiar issues occupying compatibilists; freedom of the will and moral responsibility. I will begin by discussing the latter proceeding then towards an explanation of the former.

Let's indulge our imagination in the following fictional scene: The evil witch Magdalene wishes Prince Charming to kill his bride to be. She is set on using all her wiles and her dark powers to get him to do this, but is reluctant to intervene just yet seeing that Charming is still in the midst of deliberation as to whether or not he should execute this action.⁶⁰ Magdalene resolves to interfere *only* in the instance that Charming decides against the murder, should this happen, she will employ a magical spell with the power to force him to kill his bride by making him reverse his original intention. Regardless of Charming's decision then, the witch will inevitably get her way. However, as it turns out she needn't have worried since Charming resolves to go ahead with the murder of his own accord and executes it successfully. His bride is dead, Magdalene is pleased, and more importantly Charming is satisfied as well.⁶¹

⁵⁹ There is some disagreement though as to whether Frankfurt's is a genuinely compatibilist approach. He asserts himself openly to be "neutral" as regards the problem of determinism, and the focus of his discussion is on qualifying agent responsibility rather than explicitly seeking a route to uniting determinism with free will. Nonetheless, his argument sheds valuable light on the ways one could respond to the above issue, and for this reason I will regard him as an *indirect* compatibilist (See p.95 of *Freedom of the will & The concept of a person* for some of his remarks on determinism).

⁶⁰ A tale where the Prince himself possibly carries a wicked streak I will accept may seem slightly counterintuitive for those raised on the fairy tales of Hans Christian Anderson, but for the purposes of making a philosophical point I would urge such readers to put their ordinary intuitions aside in this instance. If it helps at all, we can assume Charming's inclination towards malicious intent to have been fostered by doubt concerning his bride's loyalty.

⁶¹ For Frankfurt's original example see pp.835-6: *Alternative Possibilities and Moral Responsibility*.

According to Frankfurt the above scenario demonstrates perfectly how moral responsibility could still be applicable despite the absence of alternative possibilities or AP. For here we have an instance of someone having performed a given action – in this case a murder – which was willingly chosen, and yet, the fact remains that this “choice” obtained within circumstances that had a fixed outcome. Since Magdalene is the one with ultimate control over Charming’s actions, for this reason he is said to lack AP but this need not imply that he thereby forfeits his chance to be a subject for moral assessment. The action performed is still *his* because it proceeded from his own volition rather than having been forced, even though it could have been forced had he decided otherwise than he did. On this basis Frankfurt contends that in any such similar situation, it would be “unreasonable” to excuse the agent for his action or to deny him the praise which it would normally call for on the mere assumption that he *could not have done otherwise*; true, the action undertaken strictly speaking may not have been “up to him” but it was “in a way up to him” whether he largely acted from his own initiative or merely as a result of external interference.⁶²

In a subsequent article titled *Freedom Of The Will And The Concept Of A Person* the question of human free will is examined at length. A starting premise of Frankfurt’s is that the will of a human being unlike that of any other animal, is distinguished by its capacity for forming second-order desires a special variant of which he calls second-order *volitions*.⁶³ Desires of the first-order, may include such things as wanting to lose weight or deciding to go for a walk, and are often followed by some action which aims to fulfil them. Second-order desires are different in that they result from “reflective-self evaluation” whereby the agent determines to herself which desires she in fact wishes she had and those which she wishes she didn’t.⁶⁴ For instance, *wanting to want* to stop smoking is a desire of the second-order, which if the agent means seriously to fulfil all the way to action, more properly represents a second-order volition.

Second-order volitions lie at the heart of our personhood because they reveal the intention to make a given second-order desire into one which also expresses the agent’s *will*. Having said this, it is of course possible though “unlikely” that there may be agents with only second-order desires but no corresponding second-order volitions, and in such cases the being in question could not be viewed as a person but rather seen to be a *wanton* given the absence of these volitions. To clarify,

⁶² Ibid, p.836.

⁶³ On Frankfurt’s view nothing but a human being may be accredited with possessing a *will*, where this is characterised by the capacity for self-reflection leading to the formation of second-order desires. Desires of this kind can take two forms: they can either represent the agent’s wish to simply experience what it is like to have a certain desire, or they can represent the agent’s wish for a certain desire of his to become *effective* in that it provides the impetus for future action and so comes to constitute his will. It is this latter sense of a second-order desire that Frankfurt identifies as volitional and a distinguishing mark between human and non-human creatures (See, Frankfurt, pp.84-86).

⁶⁴ Ibid, pp.82-3.

the following example can serve to demonstrate the difference between the two categories of being:⁶⁵

Through the concept of ‘addiction’ let’s assume that we have before us two alcohol addicts both of whom are driven towards drinking in virtue of the same underlying physiological condition and where both cyclically succumb to their desire for the spirit. What makes one of them a *wanton* addict and the other a *reluctant* or “unwilling” one is dependent on their attitude towards their disease. The wanton addict is characterised by his complete lack of concern as to which of his competing first-order desires win out. That is, he simply does not stop to think which of them he would prefer to constitute his will.⁶⁶ Whether or not his desire to drink overpowers his desire to refrain from doing so, he remains indifferent to the result so that whichever desire comes to define his will the wanton’s attitude will remain passive as regards the outcome. Conversely, the unwilling or reluctant addict does care about which of his competing first-order desires come to constitute his will. For this reason he additionally forms a volition of the second order expressed by the desire to not want to want to drink which becomes the desire he wishes to be *effective* in him and the one he hopes will be representative of his will.

The distinction between a wanton and a person allows Frankfurt to explicate how it could be plausible for someone to maintain that what he does goes against his will, and therefore how such a being may either be wholly or partially cleared of blame. The idea is that when a person forms a given second-order volition regarding his conflicting first-order desires, through this very action he associates himself with the one rather than the other of these desires reflecting the process of “identification” and “withdrawal” which signals his affinity with the one kind of desire, and a rejection of its opposite.⁶⁷ When the desire from which one firmly resolved to withdraw continues to dominate, that is when someone like the addict may justifiably say that his compulsion to drink proceeds against his will, or that his addiction is not of his own free will.⁶⁸

⁶⁵ I am using a slightly different example from Frankfurt’s – the specifics of the situation don’t matter here as long as the difference between a *person* and *wanton* is made apparent. For Frankfurt’s demonstration see pp.86-89, *Ibid*.

⁶⁶ Likewise, should he happen to have second-order desires resulting from social conditioning for instance, these too would be treated with the same indifference he shows towards his competing first-order ones. What is key is that the wanton will never develop these second-order desires into volitions.

⁶⁷ *Ibid*, p.88.

⁶⁸ The above describes a typical case. We have an addict whose desire to keep drinking remains the stronger of the two despite his resolve to withdraw from it after having formed the necessary volition. It is of course an interesting pair of questions to pose how exactly despite having identified oneself with the one kind of desire and withdrawn from the other, the addict continues to experience these relapses and why Frankfurt thinks his will may be presumed absent when he acts against his second-order volitions? Both questions may benefit more from a psychological rather than strictly philosophical analysis, having said that, with respect to the latter, I am prepared to resist established psychological and physiological opinion and argue that there is no

Freedom of the will then according to Frankfurt consists in the ability to make second-order volitions or *preferences* concerning one's first-order desires.⁶⁹ Where this capacity is seen waning or absent, the being in question may rightly be judged to lack a will that could be called free. However, freedom of the will is to be distinguished from freedom of action, the former, as earlier examples suggest need not include the latter in order to enable a person to become the subject of moral approval or disapproval.⁷⁰ A domestic animal such as a dog for instance can be said to have freedom of action in that it can see its desires reach fruition, though not freedom of will since it is unlikely to have the additional capacity for making informed preferences between competing sets of desires.⁷¹ Resultantly, a dog cannot be held responsible for biting a child or indeed having done anything else that by conventional standards may be viewed as wrong or inappropriate for the sole reason that it never formed (or could form) a prior intention to align its will with the one rather than the other of its desires. Conversely, a person who *can* form a prior decision with respect to which of his first-order desires he wants to be representative of his will may be held accountable, since his decision stands as a reliable indicator of the kind of will he wishes he had even though in certain cases, it also

certain reason for denying that the agent lacks freedom of the will when acting against his volitions. Specifically, why should it not be assumed that when an addict continues to drink his will plays the same crucial role, just as it does when he succeeds in curbing his addiction? To clarify, I would maintain that in both situations the addict acts of his own free will, though in the one which sees him succumbing to his disease he may more properly be said to exhibit a *weak* rather than a strong will, but even a weak will I contend may be presumed to stem from either a temporary or permanent *freely* chosen decision to obey the one rather than the other competing desire. My suggestion is that even a weak will ought to be seen as expressing the agent's freedom because its weakness can fundamentally be understood as something that was *allowed* or permitted to happen, rather than something that just took over as though it were some inexplicable opposing force that mercilessly carries a person to do things against their (*free*) will. It may further be assumed then that one's weaknesses of will are still expressions of a freely acting general will, which at some particular moment chooses to withdraw from whatever volition it previously made and identifies itself with its contrary. Unfavourably however, this now implies that one's second-order volitions could lack permanence which may serve to undermine their significance as being the primary determinants of personhood. That is, if a given volition favouring one rather than another first-order desire can easily change, and go on fluctuating, one may begin to wonder what exactly makes volitions so appropriate for defining the essence of personhood, especially when they do not seem capable of establishing any kind of definitive choice with respect to one's competing desires. At best then, they appear to function similarly to the familiar second-order desires during those phases when a person acts against his original volition. But then if volitions can revert back into second-order desires at will, why trouble oneself to draw such a sharp distinction between them as Frankfurt has done? My suspicion is that the central motivation for this lies in being able to demonstrate the disparity of attitude between a 'person' and 'wanton', where the latter does not even utilize his will in order regulate or change the content of his desires.

⁶⁹ Carrying that ability alone though is still insufficient for having a free will. Only when a person actually *forms* the necessary preference does something with the potential to be called a 'will' begin to assume expression (See p.90).

⁷⁰ In other words, what matters most according to Frankfurt is that a person has formed a certain intention, whether that intention also has the further opportunity to be realised is of secondary importance when it comes to making moral assessments. It is the intention here which expresses a person's will and this in turn permits us to make subsequent judgements concerning his character.

⁷¹ I am assuming quite liberally here that animals such as dogs may be considered to have something very similar to what we call desires, granted that these in all likelihood stem from an instinctual rather than rational source.

enables us to mitigate our judgement of him. This helps to account for why we might be less inclined to judge the unwilling addict harshly following a relapse knowing that his behaviour does not express the will he wants to have but instead, only confirms the gravity of his disease which makes it challenging for him to get his will to accord with his second-order volitions.⁷²

Having given the basic elements of Frankfurt's argument, it should now be more apparent why in the earlier example of Charming and Magdalene blame was the appropriate reaction to use. In fact, Charming's case is not that dissimilar to the unwilling addict's predicament; both persons have the opportunity to make a certain choice and both arguably are under the influence of a stronger force.⁷³ Charming's final decision results in his will coinciding with Magdalene's will unbeknownst to him, and so she never even has to lift a finger to aid in the murder. The addict's final decision opposes the will he has in that it signifies his desire for his existing will to be different. Whether he succeeds in this is another matter, but what shouldn't escape notice is that though the addict and Charming may not ultimately be free to have the will they want, they are still at liberty to either *identify* with their situation or to *withdraw* from it. Charming's decision turns out to be such that his preference is identical with what was meant to happen anyway, rendering this into an acceptance of the will he was bound to express.⁷⁴ The addict's decision conversely, is a reflection of his desire to withdraw from the will he has been given, which allows us to see him as a reluctant sufferer and to sympathise with his situation. As observers and casual judges we appear justified then on the basis of what is known, to place blame in the one instance though not necessarily in the other.^{75/76}

In support of the above explanation, Frankfurt himself states the following in the concluding section of his paper:

*"It is not true that a person is morally responsible for what he has done only if his will was free when he did it. He may be morally responsible for having done it even though his will was not free at all"*⁷⁷

Since Frankfurt's priority in this and the previous paper rests in defending the applicability of moral responsibility in the absence of AP, he does not trouble himself with formulating an actual theory of freedom. As long as responsibility can be established it is not so important that the person in question may never have been in a position to have the will he wanted, that aspect of one's freedom could be negligible and yet need not preclude the expression of another kind of freedom of the will which sees one aligning with or digressing from an already existing effective desire or state

⁷² Ibid. see pp.90-1.

⁷³ In Charming's situation that force is the witch, in the addict's it is his underlying disease.

⁷⁴ The identification being the result of coincidence of course since Charming has no idea he ultimately has no choice but to commit the murder.

⁷⁵ At least according to Frankfurt.

⁷⁶ That is, we blame Charming but may think twice before blaming the addict.

⁷⁷ Ibid, p.93.

of affairs, viz., that which exists because it had to exist given that determinism obtains. An instance of this can easily be shown by replacing our original example of the unwilling addict with one who is *willing*.⁷⁸ Once again, we are faced with someone whose will is not free since his desire for drink will be effective regardless of which of his competing desires he may want to constitute his will. In this particular instance however, the fact that drinking becomes the addict's effective desire is rooted in two causes; one is strictly physiological and serves as an explanation of his disease, the other though is psychological and reflects his own *willingness* to accept the harmful desire.⁷⁹ In accepting it then, the addict accordingly makes the desire in question become *his* effective desire, in other words, he identifies it as his own rather than 'other' – a process which makes the will he was given also *his will* thereby permitting us to draw the reasonable assumption that his actions are free, intentional, and voluntary when it comes to his continuing addiction.

Before proceeding to the final part of this chapter, two areas of concern need to be mentioned regarding the view presently outlined. First, there seems to be an obvious assumption driving Frankfurt's discussion of the will and that is that despite the possibility of our wills being determined, there is always the capacity for agents to express another kind of freedom which comes from making second-order volitions, and that this fact highlights an important aspect of our personhood which remains independent of the determinism constituting the will we were born with. It is argued we are capable of wanting to be different from the way we actually are, but this *wanting*, how does it come about and what is its composition? The answer I believe can be found in seeing Frankfurt's argument to imply that the will of a human being carries a certain duality, and that it is this duality which underpins our desire to be different. One part of the will is fixed from birth and is the cause of many of our first-order desires, the other part however is capable of self-reflection of a kind that may lead to permanent change perhaps, or at least a permanent refusal to accept ourselves and our situation as these stand. As I see it, of central relevance here is that our 'wanting' to be different, to have different desires from the ones we presently have, likewise has its origin in the will and is made possible *only* in virtue of the will's suggested duality.

Why should any of this be problematic though you may ask? One reason is that Frankfurt provides no argument to justify the underlying duality. Even if the will's dual tendency could be accepted at face value matters would still remain complicated, since our understanding of this will would continue to be fuzzy at best owing to the layer of mystery surrounding its nature and origin. A particular area of concern might concentrate on the issue of how the will manages to obtain this division in light of determinism? More specifically, if it is being maintained that the will we have is a

⁷⁸ Ibid, pp.94-5.

⁷⁹ By this Frankfurt means that he forms the necessary second-order desire that his first-order desire for alcohol should be effective, p.95.

fact that remains outside of our control, how then does it become possible to produce a division within it that results in a rebellion of the self-reflective part against the determined part? Frankfurt's response I suspect would be to refer to our distinctively human capacity for introspection, a capacity which also enables us to eventually make the kind of second-order volitions he discusses. But such a response would be question-begging on a two-tiered level; first, we are being fed a mere assumption about a supposedly independent capacity for introspection that remains impervious to the kinds of deterministic influence that ordinarily invades a good portion of the rest of our will. Secondly, the status of the accompanying tendency to make second-order volitions will be questionable too because of the uncertainty governing our introspective capacities. That is, it remains to be established whether the volitions we end up with really did develop spontaneously rather than just having occurred as necessary effects much like the will we received.

The first area of concern then raises two related issues: (i) what by way of argument do we have for accepting that a division within the will is possible? (ii) if we do take the duality of the will as our starting point, then a further argument needs to be given to show how this duality characterises a genuine tendency towards making freely willed decisions in opposition to the determined part of the will, instead of just producing more deterministic outcomes. More importantly though, we want to know how out of the will's initial determinism these independent acts of opposition succeeded in establishing themselves, or how the will can develop an indeterminate other self to mark a person's autonomous capacity for wanting to change.⁸⁰

A second area of concern which I wish to emphasize, concentrates again on the fictional example of Charming that was used to establish agent culpability. The specifics of that example do not require repetition aside from two core premises. First, regardless of the type of fictional scene envisaged, it is always the case that there exists some fixed or determined outcome which must obtain irrespective of what the agent himself chooses. Second, we assume that in all these scenarios agents always have the ability to make a choice and depending on what they decide, that choice will either coincide or conflict with the fixity of the outcome. Should a conflict occur, an external source of power be it a cunning sorcerer, evil scientist, or even God will intervene to see to it that the agent changes his original preference. Clearly, despite these conditions, agents still possess the *illusion* of

⁸⁰ It is possible that I may be accused of gratuitous pedantry in asking for an explanation for how the will can have both determinism and indeterminism constituting it and I appreciate the difficulty of fulfilling this requirement. However, I do not think the challenge unjustified when we consider how much explanatory exertion the libertarian is forced to undertake in his effort to ground the possibility of genuine acts of free will. The same degree of thoroughness I believe ought to be expected from the compatibilist, and if the latter's reasons for avoiding it stem from a fear of being presented with the Strawsonian "obscure metaphysics" objection that haunts libertarians, then perhaps compatibilism as a solution has even less to offer than its libertarian rivals or effectively, offers no more clarification than they have in the pursuit of turning our assumptions concerning free will into an intelligible reality.

being able to act as they wish and this is why AP is not considered crucial for establishing moral assessment in the example Frankfurt presents. For his argument holds that as long as the agent chooses the same morally questionable outcome that would have been realised anyway, then he/she continues to be responsible regardless of that outcome's overriding determinism.

Though the above explanation seems intuitively correct, there does however arguably remain a residual nagging feeling that somehow despite the agent having done 'x' as a result of independent choice rather than coercion, the fact that the outcome would have been just the same regardless is worrying. Looked at from a different perspective, all we have is a situation where there happens to have resulted a coincidental concurrence between the agent's choice and the externally determined outcome. Why should it not concern us then that had the agent chosen otherwise the same pernicious outcome would still have obtained? Well, here I think lies another potential difficulty with Frankfurt's approach. We feel entitled to blame Charming for his decision to murder his bride because that decision reveals something about his character or the kind of person he is. And yet if we distance ourselves from these "character judgements" and look with an impartial eye at the situation in its totality, what could become apparent is that we have a moral agent who doesn't 'really' have a choice to make and the question to ask ourselves, is whether this actuality should not be introduced as a reason for mitigating a substantial portion of the responsibility that would otherwise have to be incurred?⁸¹

I believe the above question is an interesting one with a capacity to yield a variety of fruitful responses though space does not permit an extensive examination of them here. However, two things I will mention before withdrawing from the present discussion is first, that the agent in question either maximally or minimally nevertheless continues to be accountable, so in this particular regard I think Frankfurt secures his victory in showing that AP is not an essential component of moral responsibility. The other feature worthy of attention is that Frankfurt himself seems to be aware of a similar objection, to which he responds sufficiently well by introducing a distinction between being *fully* responsible and *solely* responsible.⁸² To demonstrate this he asks us to imagine a light which can be turned on or off by a simple flick of *either* of two switches, and a situation where the light has been turned on by the simultaneous flicking of both switches from two different agents neither of whom are conscious of the other's presence. In this situation no one can be said to hold sole responsibility for the light turning on, and it cannot be claimed that they shared

⁸¹ The potential for a dichotomy between *judgements concerning character* and *judgements concerning responsibility* was suggested to me by Sarah Broadie following a close discussion of Frankfurt's paper. My own inclination is to say that such a distinction makes it both difficult and contentious to determine (i) which term ought to carry the most relevance or moral weight when assessing a person's actions and (ii) whether this would ultimately make any difference to the degree of blame placed upon him.

⁸² See p.95, footnote: *Freedom of the Will and the Concept of a person*.

responsibility if by sharing we mean to say that each was partially responsible; rather, both of them individually seem to have been fully responsible.⁸³ Likewise, a similar explanation can be used for holding Charming fully responsible for his decision, for we know the murder would have taken place *either* because the witch intervened *or* as a result of his own resolution. The presence of this fact suggests that it would be wrong to regard him as being the *sole* possible cause of the murder, but quite right to interpret his actions as bearing *full* moral weight since on this particular occasion, they did play a pivotal role in helping to achieve the necessary outcome.⁸⁴

III -Troubled by Compatibilism

In this final section I wish to address the objection that compatibilist accounts of freedom are too limiting, which will form the basis for two further objections in the subsequent chapters. I begin with two related examples which seek to frame my reservations proceeding then towards a development of their implications.

Example 1: A child is presented with an activity book and amongst the various engaging things on offer she selects a draw-by-numbers exercise. This exercise consists of a series of numbers being dispersed methodically all over the page, the numbers for this particular exercise range from 1 to 32. The child in turn is required to carefully follow each of these numbers starting with 1 through to 32 with her pencil or crayon and depending on the spatial location of 1, 2, 3, etc...she will have to guide her pencil accordingly in the direction of the next succeeding number. When she has finished the numbers should reveal a picture, in this instance it is of a cat.

Example 2: 12 contestants take part in a live television show called *Big Brother* which requires them to all live together in a large house where they will be watched 24/7 by internal cameras and asked to perform certain tasks determined by an undisclosed 'Housemaster'. Everything in this house has been prearranged right through to the set of house rules the contestants must follow and to make matters more interesting for the audience, housemates are often placed in trying psychological or

⁸³ Ibid.

⁸⁴ To clarify, there are of course some discrepancies between the two sets of examples but our fictional case maintains the key elements needed to establish similarity between the two, namely, we have a situation where the outcome can be achieved by either of two ways, viz., through Charming's deliberate choice or by way of external manipulation. The outcome is not dependent then on the obtainment of one particular circumstance nor on the partial sharing of responsibilities. Further to this, we have an agent who willingly desires to produce the outcome and produces it independently of the second available alternative. These factors arguably make a strong case for seeing him as being fully responsible, though not necessarily solely responsible in light of the fact that there was always the possibility for the witch to exert her influence if needs be. I am assuming here that for an agent to have sole responsibility for something there must be no other way of achieving the desired outcome except through the agent's own efforts, and this condition fails to obtain with respect to our fictional scenario.

physical conditions which the Housemaster knows will be likely to provoke some anticipated reaction from each contestant depending on their relative character traits.

So what exactly is the relevance of the above two examples and what do they share in common? Both examples involve agents executing their actions within the limits imposed on them by a given fixed structure that determines the scope and nature of the choices available to them. In other words, it guides their decision-making. To take the first example; if we had to comment on the child's freedom we would say that she is free in at least the following four ways: (i) she can choose to finish the exercise thereby revealing the drawing she wants to see (ii) she can also decide not to finish it, in which case she will be left with an incomplete image (iii) she could be especially obstinate or creative perhaps and may attempt to produce a different drawing from the numbers given on the page, though she is very unlikely to succeed in this endeavour (iv) finally, she may just decide that the activity requires too much effort and so forgo doing it altogether. At least some of these alternatives I think are plausibly within the child's sphere of possibilities and express her unique portion of freedom at that given moment in time and for that particular situation.⁸⁵

Now consider the second example which is a little more advanced; there is an external body in the form of a Housemaster who controls what tasks the housemates will be expected to perform, the rules they will need to obey, and even has the capacity to sometimes know what will act as an effective trigger to produce a certain form of behaviour he wishes to encourage.⁸⁶ This master could easily fulfil our idea of a malevolent demon. Once again we can examine what it would mean to comment on the freedom component – in this instance the housemates' freedom – and the following I think would be applicable, (i) they can choose to obey the house rules and perform all the tasks diligently but with *indifference*, meaning they are not concerned with the outcome whatever it might be (ii) they can choose to *rebel* and decide to leave the house (iii) they may choose to comply because they feel they must though inwardly *resent* doing so (iv) finally, they may wish to *accept* the situation because they are happy with it.

⁸⁵ Strictly speaking though (iii) may not be a 'real' alternative since this is something she is unlikely to accomplish, and as I argue for an alternative to be genuine it must not only be hypothetically possible to realise it but practically too.

⁸⁶ It is perhaps worth noting that the Big Brother programme in spite of its somewhat tarnished reputation is predicated on the psychologically respectable notion that people are predictable in that, given the right circumstances they will likely fulfil the expected behavioural output. Taking this a step further and applying it to the present discussion, one could even suggest that the behavioural output exhibited reflects a degree of deterministic causality.

So the above scenario describes the housemates' freedom in the following variety of ways: Indifference, Rebellion, Private Resentment, or Willing Acceptance.⁸⁷ All four categories represent possible alternatives available to the housemates, and these alternatives in turn define the extent of their freedom in this given situation. The alternatives arguably also shape the content of their future decisions and given what we know about the Housemaster's ability to manipulate their character traits when he desires to obtain a certain reaction, it is also reasonable to assume that often though not necessarily always, the housemates' decisions are a product of this manipulation. Differently put, we can assume that some of their decisions are periodically determined.

The examples I have given were only meant to capture a basic outline of an idea not yet complete, but the intention may now hopefully be stated with greater clarity. Both illustrations aimed to provide an analogy with compatibilism; specifically, they sought to reveal instances of the manner of existence we can expect to attain under a compatibilist notion of freedom. To many I suspect the resulting image will seem satisfactory, after all, if we consider the first example again only this time substitute the activity book with a deterministic world structure, there clearly still remains potential for agents like the child to exercise a relevant degree of autonomy concerning the choices available to her, even if these have been produced by a deterministic framework which she is unlikely to escape. In the spirit of Taylor and Dennett we can say that though the content and range of realisable choices present is determined, it is still up to the agent to decide if they will pursue them or not.⁸⁸

I acknowledge that this kind of freedom may be significant, since upon completion of a given task persons are arguably entitled to feel a certain sense of achievement from knowing that it was *because of one's effort* that the task was accomplished. However, I do not think that this interpretation gives enough meaning to the idea of freedom, and to see this we only need to imagine a case where the agent may wish to 'break out' of their situation in the pursuit of personal growth or reform, much like the unwilling addict in Frankfurt's example, and to be told that it cannot be done. Admittedly, some compatibilists like Dennett do try and overcome this difficulty by postulating theoretical possibilities in an alternative reality, but I have already argued that this is just not good enough as a solution. The fact remains I believe that the compatibilist's theoretical groundwork struggles to provide genuine possibility for change to take place, the kind that implies escaping from the deterministic whirlpool of cause and effect or the force of necessitation in

⁸⁷ It is possible to say of Indifference and Rebellion that these are not necessarily rational choices to make, especially Indifference which suggests a type of animalistic detachment; nevertheless, even these options reflect a palpable degree of autonomy.

⁸⁸ That is, if we revert to the original example, we can say of the child's situation that despite circumstances being sufficient for the completion of the exercise, it remains *necessary* that she exerts the requisite effort in order to see it completed.

general. We see this in our example of the child who cannot create a different drawing but may only realise the one already provided, or the unwilling addict who finds he cannot overcome his nature by permanently changing the content of his first-order desires which means he will always experience a compulsion to drink.

For the majority of compatibilists such cases do not present much concern. As long as it can still be shown that persons are responsible, the fact that they are not free to satisfy those desires which fall outside the range of their available alternatives is not generally considered crucial. The trouble however is that the range of alternatives themselves is very restrictive in comparison to the needs and desires a human being is capable of having, and this makes the resulting theory of freedom a grim one. It is for this reason that I compare the compatibilist's notion of freedom to what a child using a draw-by-numbers activity book may obtain, or the kind that contestants living in a controlled "Big Brother" environment may encounter. There is of course some room to *manoeuvre* which largely consists in the capacity for forming second-order volitions, but these the compatibilist cannot and does not guarantee the possibility of realising. What is needed though is a theory that can justify the presence of (i) an ability for persons to introduce original decisions *and* (ii) for those decisions to be capable of fulfilment. Without both components being present we either end up with a vacuous form of freedom of the will, or what is equally objectionable, a freedom of action where the preference towards a particular action does not stem from a will that is autonomous.⁸⁹

To sum up, I have argued that if we accept the compatibilists' world view then the freedom we can expect to have will be extremely limited, too limited I believe to satisfy the ordinary human expectations placed on this idea. In particular, I have questioned whether this account of freedom can permit persons to make 'out-of-character' decisions and eventually withdraw from the givens of their situation. The answer remains a contentious one of course, largely depending on which variant of compatibilist theory is being applied, but I also think that this is an issue the compatibilist either struggles with or casually overlooks. At best, we appear to have grounds for accepting a volitional capacity for making certain decisions that may conflict with one's fixed nature, but this capacity does not necessarily translate into a freedom of action which sees those decisions being put into practice. Such a result is disconcerting because at this point all we have is an inane kind of freedom of the will which *may* be enough to warrant a superficial form of moral assessment, but certainly not enough to maintain that persons are free.

⁸⁹ I contend that having freedom of the will without a corresponding freedom of action present is rather meaningless. The full purpose of having a free will fundamentally consists in the capacity to transform one's desires and subsequent choices into action, or at least it ought to be. And though it is not essential that one *always* does this, it is important that there exists the opportunity of doing this, otherwise our will's freedom can only be said to function as a psychologically conciliatory element and nothing more which cannot suffice as a basis for a compelling theory of freedom unless we are content to have that theory be incomplete.

A more serious worry however, is that the alleged volitional capacity might not even be present if something like psychological determinism is assumed. In this case the unwilling addict of Frankfurt's examples would strictly speaking cease being a proper candidate for moral judgement, for now not only is his need to drink the necessary effect of determined factors, but his stubborn insistence to beat his addiction is also paradoxically a consequence of necessitation. In this situation it would be unreasonable to blame such a person for his disease, but it would also be equally unreasonable now to praise him for his efforts in trying to conquer it.⁹⁰

In light of the above difficulties I see the present account of freedom as deficient and insist that we can do better by trying to construct our theory of freedom independently of determinism's corrosive presence. In thinking that not enough has been offered and that far more remains to be said I believe I am not alone, and so to this task attention will now be turned in the remaining two chapters.

⁹⁰ Having said this, many compatibilist would disagree with this austere conclusion on the grounds that it would result in the complete suspension of moral discourse. Though hard determinists like Pereboom argue this may not be as socially catastrophic as it initially appears, see his *Living without Free Will: The Case for Hard Incompatibilism*.

Kane's Labyrinth

I - Neglected Pathways

The preceding chapter dealt with some prominent compatibilist solutions to the free will problem and found them collectively to be unsatisfactory. The theory of freedom offered by them arguably doesn't quite meet some of the deep intuitions we have about ourselves, providing us with mere "elbow room" from which responsibility for actions may be extracted rather than presenting a positive account of a human being's freedom that hasn't been subjected to modifications or dilution. More importantly though, and something that will become clearer in the next chapter, responsibility is not the *only* consideration to worry about when it comes to discussing the free will issue, but it is nevertheless an integral feature of having a free will. For this reason, I would like to concentrate the present discussion on this topic by examining Robert Kane's extensive analysis which may help convince the reader that compatibilism *does not* provide the right theory of freedom. So for those who have always experienced a certain uneasiness when it comes to this doctrine, judging that what it offers surely cannot be the freedom we had in mind, that there is more to this idea and indeed must be more if it is to be sincere, then this starting chapter should prove useful in helping to echo some of these latent anxieties.

Ultimate Responsibility

The muffled sound of disgruntlement audible from the libertarians' corner is not unjustified, and much work has been done by Kane on the free will question reflecting an understandable desire I think, to provide what compatibilism has failed to ascertain when it comes to our intuitions about freedom. His method has involved re-examining the various parts of the free will debate he feels have not received the attention they are due, with the hope of being able to offer a better definition and defence of the will than has standardly been given by compatibilists as well as libertarians alike.

So what can be said or rather, ought to be said concerning the free will issue? According to Kane we begin to approach this question by examining the compatibility itself between free will and determinism which reveals the first neglected pathway, *Ultimate Responsibility* or UR for short.⁹¹ Traditionally the focus has primarily centred on the notion of alternative possibilities (AP) which is often used as the most convenient card amongst incompatibilists who wish to undermine the suggested coexistence between determinism and free will. Kane however feels the move provides

⁹¹ Kane, p.407: *Some Neglected Pathways in the Free Will Labyrinth*.

an insufficient basis for resolving the issue while also denying incompatibilists the argumentative force needed to really challenge compatibilism. A more fundamental reason to question the compatibility he thinks should be UR, which is the idea that persons hold ultimate responsibility for something when they have been responsible for anything that was a sufficient condition, cause, motive, or reason for that thing.⁹² So for instance, if a given choice issues from, and can be made explicable by reference to one's character and motives, then in order to bear ultimate responsibility for it, one needs to have been at least in *part* responsible for the specific character and motives in question through the voluntary making and enactment of various choices and actions done in the past.⁹³

Two things however require brief clarification at this point to avoid any future confusion. First, Kane is not suggesting that we abandon AP, only that we consider the prospect of it lacking the centrality originally ascribed to it. Secondly, the fact that AP continues to be important is supported by the nature of UR, since in order for persons to be able to claim ultimate responsibility for some action they need to have been responsible for that which was a sufficient cause of it, and to have that they needed to have AP. In other words, the agent needs to be responsible for at least *some* of those past actions which went towards forming his present character and motives, even if not for *every* action performed. But in order for this to happen AP needs to have existed during those past moments when the agent was in the process of deliberation, so that the resultant action can genuinely be claimed as his own because its obtainment necessarily implies that there was always the possibility that he could have acted *otherwise than he did*. AP then is a vital component in helping us to reach UR, its presence suggesting that the agent's character and motives which proved sufficient for the execution of a given present action resulted from basic choice rather than being formed by unknown causes.

The role of AP then is to provide agents with the opportunity for character-defining moments, or to enable what Kane calls "self-forming actions" (SFAs)⁹⁴ which are the building blocks when it comes to finding raw materials for deepening our understanding of the free will issue. Thinking about this for a minute, there is arguably nothing fantastical in the idea that for a person to be seen as having acted freely and responsibly, it is not enough that he acted voluntarily moments before the action took place, but that in addition to this his intentions and motives which proved sufficient for the action, stemmed from a will which was shaped by him and him alone. When agents perform SFAs, the resultant will they end up with is one that may truly be seen as "their own free will" because its development has been marked by the presence of AP and so involves the understanding

⁹² Ibid.

⁹³ Ibid.

⁹⁴ Or "Self-forming willings" (SFWs), p.408.

that every choice and action previously undertaken, solely reflected the agent's volitional effort to pursue *that* particular choice and action rather than another.⁹⁵

For Kane, the origin of our characters, motives, and purposes, together with the question of who produced them, and who is responsible for them, continues to be a basic mystery that lies at the heart of the "free will issue".⁹⁶ Here then is the real enigma fuelling past and present theoretical rifts, and if we are honest we should be able to admit that the primary tools at our disposal continue to be intuition and conjecture rather than irrefutable proof, where the latter represents unambiguous scientific validation.⁹⁷ So we continue to wonder who or what is really responsible for the way we are; is it God or heredity, nature or the environment, society or culture, maybe it is fate or possibly even hidden controllers? But what if it is we ourselves? Kane's implicit response here is to say that we just don't know for certain, but that for anyone earnestly wishing to defend our intuitions about free will, a very good place to start would be with the notion of the agent himself as the will-maker.⁹⁸

Plurality Conditions

The idea that persons must be authors of their own will in order to possess final responsibility or UR is an integral notion within any respectable free will debate according to Kane. Despite then the prevailing tendency amongst scholars to over accentuate the importance of AP and underplay or simply ignore the relevance of UR together with its accompanying condition that the agent be determinant of his own will, UR arguably stands as the more fundamental reason for rejecting the compatibility between free will and determinism. To fully appreciate the reasoning behind this however, we need to examine yet another abandoned pathway in the labyrinth, which brings us to *Plurality Conditions*.

⁹⁵ Ibid, p. 408.

⁹⁶ Ibid, p.409.

⁹⁷ Hard determinists may be tempted to intervene at this point on the grounds that there is overwhelmingly strong scientific backing for the view that events result from the presence of a determinate causal sequence rather than occur randomly. But this issue has not been settled, and questions concerning the role of human consciousness as well as the nature of the relationship between agent causation and event causation, continue being heavily disputed. Indeed, the fact that the role of consciousness is still not fully understood and that physics has been unable to adequately account for it has meant that the free will issue remains being an open-ended one. In light of such reasons then, I maintain that whichever side of the debate one is on, we are at best working with intuition and plausible theoretical conjecture, or blind obstinacy at worst. (For more on this see Robert C. Bishop, *Chaos, Indeterminism, and Free Will* as well as David Hodgson, *Quantum Physics, Consciousness and Free Will*; Hodgson especially argues that the relationship between the brain and self remains mysterious, leaving us in the dark as to how exactly the two interact with each other, and more centrally what the self happens to be? see p.108).

⁹⁸ That is, one who bears ultimate responsibility for his past choices and actions which went towards making his will the kind of will it is now.

When considering the freedom of a person, what matters more is not simply whether there was an opportunity at some past stage for things to have turned out differently than they did, but whether the contingent alternative outcome (or set of outcomes) would have been produced *voluntarily, rationally, and intentionally*, rather than by accident, luck, or some form of external coercion. The ability of the acting agent to exercise various choice options with respect to some particular event, and to do so rationally, intentionally and voluntarily, is what it means to have “plurality conditions” be present and that which partially determines whether a given action was done *freely*.⁹⁹

The importance of having plurality conditions is stressed because it is possible to have AP together with a certain amount of indeterminism, and yet – contrary to popular opinion – for this to be insufficient for an incompatibilist to establish free will. To see this Kane presents a possible hypothetical example in which he asks us to imagine a world where God exists but one in which there is also a certain amount of indeterminism, so that many things which we intend to accomplish will occur but some of them won’t as a result of chance, accident, or luck. As well as this, we are to further imagine that in this world all of one’s reasons, motives, and purposes for choosing to act in a given way have been pre-established by God, such that for instance, my wanting to go for a run because sitting idly bores me, or deciding to have pineapple ice-cream instead of chocolate because I find the latter too rich, are actions which all result from the presence of predetermined reasons and preferences imbued in me by a higher power.¹⁰⁰

In our hypothetical world persons would have indeterminism and AP, such that even if my desire to go for a run is not realised on every occasion due to involuntary spells of muscle cramps or heavy thunderstorms, the *initial* possibility of me taking one is something that always exists because of the prevailing indeterminism and on account of other relevant factors such as me being a keen sportswoman, which collectively indicate that the *capacity* and *opportunity* for it continues to be there.¹⁰¹ Conversely, if I do take that run after all, the possibility of me not taking it was always there too, hence whether I take it or not remains to be something that is often subject to indeterministic influences which leave things largely open to fortune or chance.

⁹⁹ I state that having plurality conditions only “partially” determines whether an action is free because it is possible to argue under a logical version of compatibilism such as the one offered by Dennett & Taylor, for the availability of AP while also maintaining that actions are being done voluntarily, rationally, and intentionally despite the background presence of determinism. Like Kane however I am of the opinion, as rigid as this may seem, that for an action to be truly free, it needs to stem from a nature that was self-created, rather than predetermined by past circumstances. Only then would the course chosen from a particular set of alternatives, genuinely reflect the *will* of the agent rather than some given tendency of his nature the development of which he had little control over, and presumably has even less control now to try and moderate.

¹⁰⁰ *Ibid*, pp.410-11.

¹⁰¹ *Ibid*.

In this kind of world then, Kane believes persons cannot be said to exercise real freedom of the will because (i) exercising AP or *doing otherwise* becomes a matter of chance or accident, rather than being indicative of the agent's deliberate use of choice¹⁰² and (ii) there is no real opportunity to *will* to do otherwise anyway because one's will has already been pre-shaped by God to be a certain way. In light of (i) and (ii) it is concluded that the "will" cannot be taken out of the free will issue, since it is in virtue of being able to exercise various choice options using one's will in conjunction with the presence of the aforementioned plurality conditions, that a person's actions genuinely become free.¹⁰³

Will-setting

In order to complete the necessary framework of Kane's position, a final pathway needs to be examined which is given through the notion of will-setting. This notion is not something that can be viewed in isolation from UR and Plurality Conditions, rather, once we understand the function it plays in relation to these two concepts the foundation for establishing a theory of freedom becomes completed.¹⁰⁴ The preceding section has shown that in a world where one's will results from pre-setting, genuine "doing otherwise" where a person is seen to voluntarily choose the alternative course of action, could never exist. That person's decisions would always mirror the motives and reasons given to them by God, where these motives and reasons would be seen as determining the kinds of choices one ends up making, as well as precluding the possibility of intentionally selecting a different one.

The will-setting agent on the other hand is someone who sets their own will, that is, the motives and purposes they eventually end up having, *during* the performance of a given action rather than having an already pre-set will *before* the action even takes place. Specifically, Kane draws a distinction between "will-settling" and "will-settled" to mark the difference between on the one hand, actions which are done from decisions reached *at* the moment of deliberation when the agent is seen to prefer a particular set of reasons in favour of one option over another by *choosing* or deciding on them in that moment, and those actions which merely *reveal* during the deliberative stage that option which the agent has been in favour of all along.¹⁰⁵

¹⁰² I.e. without which we only have a "limited kind of freedom at best" (Ibid).

¹⁰³ Ibid.

¹⁰⁴ Within the context of the present discussion it should be especially apparent that there can be no "theory of freedom" without a basis for a freedom of the will; the above trinity (UR, plurality conditions, and will-setting) acts as a means towards establishing this basis from which it is hoped a theory of freedom may then be allowed to grow.

¹⁰⁵ Ibid, p.412.

The first category of actions are appropriately believed to be will-settling, since having no known prior preference concerning a given option, agents 'set' their wills at the moment of choice by *making* a choice. When however there is already some pre-existing choice available which merely becomes apparent at the time of deliberation, the subsequent action is taken as belonging to a will-settled disposition.

With will-setting explained, Kane is now able to offer a sequence of notions which according to him jointly give us freedom of the will together with freedom of action, as well as, *indirectly* producing an argument for the presence and so the sustained relevance of AP.¹⁰⁶ This sequence is as follows:

{From (1) acting "of one's own free will" to (2) being ultimately responsible for the will one has (UR), to (3) "will-setting," to (4) the "plurality conditions," which in turn imply (5) "could have done otherwise" (i.e. AP), – since if one can do otherwise intentionally, voluntarily, or rationally, rather than merely accidentally, or through some form of external control, then there is even greater reason to think that such a person really does have AP.}¹⁰⁷

It is necessary that all 5 notions are included in order for free will to be present since they function reciprocally. So for example, (2) which stipulates the need for UR, only works as long as persons are their own will-setters (i.e. (3)), since one cannot be held ultimately responsibility for a given choice if that choice stemmed from reasons and motives in whose production one had no influence. But one cannot have the opportunity for performing will-setting actions, unless plurality conditions also obtain (i.e. (4)) to ensure a person's actions are not random, but rather "more-than-one-way voluntary", "plural rational", and "plural intentional."¹⁰⁸ If we work our way backwards, we can now see how each notion supports the one preceding it by providing it with what it needs in order to be realised,¹⁰⁹ and how each concept in itself gives relevance to the next.

¹⁰⁶ This is important because as already explained Kane does not wish to remove AP from the free will debate, but only to accord it a more appropriate place within it.

¹⁰⁷ Ibid, p.413.

¹⁰⁸ Since will-setting involves persons having to shape their will by selecting from a number of possible options, with no fixed preferences at this stage one way or the other, for this reason their actions will be seen to be "more-than-one-way voluntary" rather than "one way" voluntary as in the case of persons with already set wills. A similar type of reasoning applies to the actions of will-setting agents when they are "plural-rational" and "plural intentional" for whichever way one chooses, one acts rationally when selecting one set of reasons over another to realise a preferred option and any genuine selection of options arguably ought to entail that the agent was also acting deliberately, or with full acknowledgement of one's actions (See Kane, p.412).

¹⁰⁹ That is, the presence of AP suggests the existence of plurality conditions, these enable will-setting, and will-setting in turn allows for the possibility of UR.

At this point, Kane leaves the foundational elements of his discussion with two resulting requirements stated as necessary for free will; the first, asks that we be *ultimate sources of our actions*, the second, that we be *ultimate sources of our wills*.¹¹⁰

II -The Intelligibility Question & Kane's Ultimate Proposal

The last three sections dealt with explaining those areas of the free will debate which Kane feels have not received proper attention. In this penultimate part before proceeding towards an assessment of Kane's proposal, that proposal will now be introduced.

It is known that those incompatibilists who are also libertarians, that is, those who not only think that free will and determinism cannot co-exist, but who also believe that freedom of the will must exist, face at least two big obstacles. One, which we have already encountered, is to be able to present a convincing argument for *why* free will and determinism are seemingly incompatible; the other, which we will look at now, is to be able to present a coherent account which specifies *what* a libertarian free will actually is. The second of these obstacles is considerably more burdensome, and the problem lies in trying to account for how persons could become ultimate sources of their wills if the libertarian assumes indeterminism as a prerequisite for grounding the will's freedom.

By now of course this problem is very familiar, but I will briefly restate it nonetheless. Namely, if we allow some events to be undetermined in order to create the possibility for genuinely spontaneous acts of the will, the window of opportunity this opens also inadvertently creates a doorway through which our newly found freedom can slip out. That is, if actual undetermined events exist, the implication is that choices and actions would become *arbitrary* and *inexplicable* since by definition, an event that is indeterminate may or may not occur given the exact same past which in turn suggests that something like *chance* rather than human agency comes to play the dominant role in producing it. The presence of chance however as a determinant in the eventual outcomes of our doings, initially at least, appears to negate any notion of control we may be said to have as moral beings. For if it is possible that a different choice could have resulted given the same past, that is, the same beliefs, motives, thought patterns etc, which went towards making it, and yet,

¹¹⁰ Clearly the compatibilist cannot meet these two requirements in order for persons to hold freedom of the will. And though it is possible to be sympathetic to compatibilism by charging Kane with having stated conditions that are perhaps (unnecessarily) too demanding, nevertheless I maintain that if one reflects on this more earnestly, they may realise that unless we state them thusly free will could always be challenged, and the agent's claim to responsibility suspended in theoretical limbo.

by chance these could just as easily have lead to a different one, it arguably becomes suspect then whether in such circumstances agents would still be regarded as free and responsible.¹¹¹

The above quandary is also defined as the 'Intelligibility Question' which is really a potent objection asking the incompatibilist to explain, how a libertarian theory of freedom is possible given that indeterminism seems to undermine rather than develop it? The solution to this predicament is something that numerous past thinkers have tried to develop,¹¹² with the consequences of some of these endeavours producing hypotheses which stated the *will* to be a mysterious substance that resides within the agent, as well as those which accounted for human control in action by asserting the presence of *sui generis* acts of volition. In being about to examine Kane's own contribution then, it should be observed that it was developed in response to what he feels have been on the whole, unsatisfactory explanations.

Kane's Position

Those who hold that there is a special but inexplicable form of causation between the agent and his actions are better known as agent-cause theorists, or AC theorists for short. Kane however subscribes to a different kind of incompatibilist theory, namely, that of "teleological intelligibility" or simply, TI. TI theories, to state it in Kane's own words, "[...]try to make undetermined free actions intelligible in terms of reasons and motives, intentions and purposes, *without* invoking extra entities or special forms of causation."¹¹³ More precisely, Kane is a TI theorist of the *causalist* variety, or a "causal indeterminist" to be exact,¹¹⁴ meaning that he does not require reasons explanations of actions to be *noncausal*¹¹⁵ or that he does not subscribe to the view that reasons and motives cannot function as causes of action.

With this in mind, his proposal can be defined as arguing for the *presence of indeterminacy within our neural processes themselves during tense moments of practical and moral deliberation.*¹¹⁶ An idea that can be stated more informatively in the following way:

* Assumption: The indeterminism libertarians require need not present problems.

¹¹¹ Kane further adds; "if such a thing happened, it would seem a fluke or accident, like an uncontrolled quantum jump in the brain, not a rational, free, or responsible action" (p. 415).

¹¹² For example, Kant's postulation of noumenal selves or Descartes' immaterial egos, as well as "transempirical power centres" that intervene in the brain, as postulated by the physiologist Sir John Eccles. Kane references all of these by way of criticism in order to prepare the reader for what he sees as the advantages of his own method (See page 415).

¹¹³ Ibid,p.416 – italics added.

¹¹⁴ Sometimes also referred to as "event causalist."

¹¹⁵ Unlike Ginet, who was discussed in chapter 1.

¹¹⁶ Ibid, pp. 417-18.

*Premises: (1) We only need indeterminism during those moments in our lives when we engage in “will-setting” or “self-forming” actions (SFA’s).

(2) Actions of will-setting happen when a person is left to choose between competing visions of themselves regarding what they should do or become.

(3) During such demanding moments of a person’s life, there is both tension and uncertainty in the mind, because one may be torn between (for example) doing what is morally right and that which is opportunistically desirable.

(4) The inner conflict character-forming moments provoke, is reflected in appropriate regions of the brain by the presence of a kind of *chaos*, making it sensitive to micro-indeterminacies at the neural level. (Conversely, when we act from settled dispositions or motives, such indeterminacies remain muted).

*Conclusion: There is then indeterminacy in our neural processes themselves owing to the private conflict persons face during episodes of will-setting.^{117 118}

*Consequence: During the described phases of uncertainty, agents may be said to exercise the capacity for both *rationality* and *voluntariness* despite the prevailing indeterminism. They can will either way at these moments because during such acts of character-formation (or SFA’s), their will is divided rather than fixed by the presence of conflicting motives. The potential threat caused by indeterminism then is arguably neutralised, since when we view persons as “will-setters” the indeterminism assumed and the effort of will that performing SFA’s entails, shows that indeterminism to be a feature which potentially *resides* in the effort of will itself, rather than being something distinct from it or hostile to it.¹¹⁹

To better understand the significance this consequence has concerning the free will issue, the following example will serve to further illustrate Kane’s meaning:¹²⁰

Nathan, a promising young man with big dreams has saved up enough money to go to one of the best universities in the country which has offered him a place. His family, so as not to spoil his future prospects, never disclosed to him the recent financial troubles they’ve been experiencing, the kind that would certainly lead his father to declare bankruptcy unless he finds a way of producing the necessary amount to pay off his debts. Now, it just so happens that a few days before Nathan is due to leave, he notices an open letter sitting on the table from which he discovers that his father owes a substantial sum of money which unless paid promptly, will lead to legal action being taken against

¹¹⁷ Kane does not formulate his argument in this manner, but I do so as to simplify his position. Nonetheless, points 1-5 can be extracted from his paper, pp.416-9.

¹¹⁸ He also adds that “what is experienced personally as uncertainty corresponds physically to the opening of a window of opportunity that temporarily screens off complete determination by influences of the past” – the point is both interesting and of course contentious, and I will return to it in the final part of this chapter.

¹¹⁹ Specifically, Kane states that the indeterminism and the effort of will are “fused” – Ibid, p.417.

¹²⁰ The following example is loosely based off Kane’s own in this paper, see p.417.

him. Nathan now arguably faces a dilemma; should he (A) leave his father to deal with this matter on his own? Or (B) give up on his dream of getting a higher education and doing something remarkable with his life, and help his father instead by giving him his college money?

Nathan's situation is just like the one described earlier in premises (2) and (3), where the agent is said to experience mental tension and uncertainty while deliberating over which course of action he ought to pursue. In our particular case, Nathan will arguably be torn between acting for his own advantage and acting for the sake of his family, where the latter may be taken as the morally right thing to do.¹²¹ It is because he is torn though, with his will set at a crossroads up until the moment he actually *decides* that the indeterminism is said to reside in the effort of will; as soon as a decision has been reached however, the indeterminate effort of will that was used in helping him to reach it, transforms into a *determinate* choice.

To do the right thing, (assuming that option (B) is the morally correct one) Nathan must try to overcome his desire for a bright future by sheer effort of will. If he succeeds, this will have been achieved owing to his effort alone, and if he fails, it suggests that he did not *allow* his effort to succeed. An explanation for the occurrence of the latter outcome would simply be that despite wanting to help his father, he also wanted to fulfil his ambitions and so wanted to fail in his first intention in virtue of having a different set of reasons supporting the alternative wish. Whichever choice he opts for in the end however, it will become his *determinate* one through having at last made a decision that allows one set of his competing reasons or motives to trump the other set.¹²²

The accusation then that indeterminacy somehow challenges or undermines an agent's capacity to make rational and voluntary choices, seems to be diffused once we acknowledge it to be a feature that resides *within* the will and terminates *by* sheer effort of will at the point of decisional resolve. And the same kind of explanation may also be used to respond to the charge that indeterminism equally threatens an agent's *control* and *responsibility*. Again, Kane maintains there is no reason to suppose it would have this effect "in and of itself" and offers several possible cases where such indeterminism does not appear to significantly impair the applicability of these two features of an agent's freedom. I will state one of them now:

Suppose we are trying to solve a difficult mathematical problem and that there is indeterminacy within our neural processes making this task all the more arduous. This indeterminacy can be best understood as a sort of "background noise" and it is this "noise" which makes the outcome of our

¹²¹ I will not defend this claim here, though I acknowledge that in choosing to keep the money instead, he need not be taken as having made the morally 'wrong' decision. All that is being suggested here is that one of his choices *could* be seen to carry greater moral worth than the other, and that this may be used as a reason to see that particular choice as the *morally right* one. Regardless of what he ends up choosing though, his action will serve to define his future character or *form* his will, according to Kane.

¹²² *Ibid.*

endeavours uncertain. Clearly, it is a distracting feature, but must it be a permanent impediment? Kane does not think so. It is possible to conceive he argues that we may still solve the problem by concentrating despite the background distraction, then, if we do in fact succeed, there is good reason to think that this was made possible as a result of the effort and that one is therefore responsible for having succeeded even though it was always uncertain whether one in fact would. In other words, in this instance, control and responsibility can be explained as having obtained through the agent's *effort to overcome a particular obstacle*.¹²³

Though the example is not perfect as Kane himself admits, it is still a "step in the right direction" as long as agents are assumed to have the capacity for making conscious efforts of will. Therefore, the main issue is not whether a particular action such as banging one's fist against a mirror with the intention of breaking it would actually result in that intention being realised, rather what matters is that *if* the mirror does break, we can and ought to be held responsible.

Having given the basic premises of Kane's argument, his proposal can now be stated more formally:

When an agent deliberates, two competing neural networks are at work and an act of self-formation (SFA) reflects the *voluntary* succeeding of the one choice over the other. This struggle to resolve two competing cognitive tasks is a reflection or consequence of our human *complexity*, a complexity that is necessary for self-formation and ultimately free will.¹²⁴ Applying this now to our earlier example of Nathan, we can say of him that he is a man torn between two possible visions of himself; in one of them he leaves to pursue his dreams, in the other he stays to help his family. It is because he happens to be a complex human being, with desires (both selfish and selfless) hopes, caprices, likes and dislikes, as well as strengths and weaknesses, that he comes to experience this mental angst. If the discovery of his father's dire situation caused virtually no change in him, no real decisional conflict, then it would suggest either that his character or will has already been pre-shaped, or that the requisite complexity is absent in him for whatever reasons.¹²⁵

Let's suppose for argument's sake that he resolves to help his father. Let us also suppose that this decisional outcome was reached after much deliberation and inner effort to overcome his corresponding desire for self-fulfilment, and that ultimately it was indeterminate throughout this whole process up until the moment of actual choosing which option he would settle on. We are further assuming of course, that his will has not been pre-set by any external agency. Then, in the instance of his deciding to help, this choice (i) proceeds from his will – implying that it was voluntary, (ii) contributes towards shaping his will, or shaping his character since it represents an act of self-

¹²³ Ibid, pp.417-18.

¹²⁴ Ibid, pp.419-21.

¹²⁵ At least, I believe this would be Kane's response to such an individual.

formation (SFA), (iii) through this capacity for making such deliberate choices, he thereby demonstrates that his will is free, or that he has freedom.

III - Out of the Labyrinth?

In this concluding section I would like to consider the effectiveness of Kane's strategy in dealing with the free will problem. It should not be surprising that I am sympathetic to his approach, though not exclusively owing to the fact that it seeks to challenge compatibilism. If we recall, the chapter began by questioning the compatibility between free will and determinism¹²⁶ which was challenged on the basis that compatibilists cannot provide "ultimate" responsibility. Consequently, the concepts of "plurality conditions" and "will-setting" were introduced which jointly facilitated in constructing the desired image of the agent as one who is *ultimately* free and responsible.

My reasons for supporting Kane's argument rest on both intuitive and rational grounds. Intuitively, I believe the proposal is accommodating to our potential view of the self; rationally, I think that if one can get past the initial cynicism at being confronted with a picture of reality that is said to contain a certain amount of indeterminacy with respect to those pivotal moments which serve to define our nature, the argument stands to present an informative and plausible account of the agent as determinant of his own will and architect of his freedom. That is, in overcoming the tension experienced when two equally desirable outcomes present themselves to us, by rationally, voluntarily, and intentionally resolving to select one of them and ignore the other, we effectively shape our nature and determine the type of person we are to be, in other words, in one sense we set the limits of our freedom for which we then bear ultimate responsibility.¹²⁷

One question to consider though, is whether identifying neglected pathways in the free will labyrinth helps us to escape it, or leads us closer to an effective solution? A prevalent objection waged against Kane at this point, centres around his claim that the proposal offered does not evoke mysterious causal factors concerning the relationship between the agent and events in its account of freedom.¹²⁸ However, we know that Kane's account receives its justification on the basis of the

¹²⁶ Also known as the "compatibility question."

¹²⁷ To be clear, "setting the limit" implies that while the individual was free to shape his own will, the resultant 'shaped' or constituted will adds another dimension to his freedom by determining the type and range of future doings he will be capable of. As a consequence, one person's range of doings may be quite different from another's and comparatively, it may be possible to judge the one as bearing greater freedom than the other. Kane does not mention this possible consequence, but I think it adds a further interesting layer to his proposal, viz., the idea that we express our freedom through self-forming acts of volition, *and* that these in turn resultantly specify the nature of our freedom including its limit or extent.

¹²⁸ For example, Randolph Clarke extensively criticises Kane's TI "causal indeterminist" approach. Amongst other things, he argues that it succeeds no better in dissolving the metaphysical vagueness surrounding the

assumed indeterminacy which he argues functions as a, “temporary screen that blocks determination by influences of the past” thereby leaving the agent with a window of opportunity to make a genuinely free choice.¹²⁹ The concern here then is that this type of explanation *does* seem to evoke a mysterious causal factor, only in this case it stems from the suggested indeterminacy of our neural pathways in virtue of which persons become capable of exercising the necessary kind of *effort* of will needed for their actions to be seen as free.

The opponents of Kane therefore maintain that a proposal which asserts the presence of indeterminism as a key factor in allowing to explain the relationship between an agent and events, is no less mysterious than those of his libertarian predecessors. Kane however arguably anticipates this type of response since he includes the following comment in a footnote:

*“Whether the requisite indeterminacy is there in the brain is an empirical question of course. The question at issue here is what we could do with it to make sense of incompatibilist freedom, if it were present.”*¹³⁰

My suspicion is that the real reason for the charge against him arises primarily because his account relies so heavily on neuroscientific findings, which do seem to lend themselves positively to a libertarian theory.¹³¹ The suspicion this type of dependency engenders is understandable of course, the findings themselves are more speculative than definitive in nature with every detail not being accounted for in the description of the agent’s causal power. We are simply told that *if* there can be indeterminacy in the brain *then* something suggestive of the will’s “causal power” can be exercised, expressed through the agent’s effort.

However, I do not take this criticism to be especially strong. Here at least we have an attempt to substantiate conjecture with promising scientific backing, already making Kane’s account more plausible and less evasive in comparison to the standard tactics of his libertarian predecessors who presumably only from want of adequate explanation, chose to reduce the will together with agent causality to a specimen of largely inexplicable *sui generis* phenomena. Admittedly, the relative infancy of Kane’s external source of support¹³² does potentially make his account vulnerable as well as somewhat theory-laden perhaps, but vulnerability does not amount to implausibility and if the theory can help produce a more coherent explanation, then it ought to be given a chance together

notion of *agent causality* which nearly all libertarian accounts find problematic (See; *Libertarian Views: Critical Survey of Noncausal and Event-Causal Accounts of Free Agency*).

¹²⁹ This is from an earlier footnote which I promised to address.

¹³⁰ *Ibid*, p.434, Note 16.

¹³¹ For instance, in the same footnote Kane adds, “there is growing evidence that chaotic activity plays a significant role in cognitive processing in the brain” as well as, “but chaotic amplification is not the only possibility. It is conceivable that novel indeterminacy could arise at macrolevels, as suggested by Senchuk”. This shows clear indication I believe that Kane’s theory draws its confidence from this kind of research.

¹³² I am assuming that the field of neuroscience is not yet at that stage of maturity where it can be appealed to by libertarians for indubitable support, as a superior arbiter during theoretical disputes of the present variety.

with the explanation. To reject an explanation however largely because of theoretical prejudice, without actually finding specific fault within the argument's reasoning, is an ignoble way to confront one's opponent.

One other objection though which is believed to have found significant fault within Kane's reasoning, is given by Galen Strawson's *Regress Argument* in which he endeavours to show that it is never possible for someone to be 'ultimately' responsible without this leading to absurdity.¹³³ I will state an abbreviated version of this objection before suggesting possible responses to it.

The objection runs as follows: 1. (assumption) to be ultimately responsible for our actions, we need to be ultimately responsible for our nature which produced them; our nature however, is not wholly a product of our own doing but developed from various genetic, social, political, and psychological factors to name a few. 2. Let's suppose we *can* be ultimately responsible for our nature – call it 'N' – because there was a way in which we created it at some stage in our lives.¹³⁴ 3. But creating N implies there was something before N, our other nature 'M' which existed at an earlier time prior to N's existence. 4. Then, if we are to be ultimately responsible for N, we must *also* have been ultimately responsible for M, (otherwise it could be argued that M *caused* me to choose N, and then I am only 'responsible' for N, but not ultimately responsible). 5. The "ultimacy" requirement leads to an infinite regress whereby in order to justify one's ultimate responsibility for N, it is necessary to do the same for whatever came before it (M), and then whatever came before that *ad infinitum*. This regress suggests it is never possible to establish a starting point of justification to show that *you* are the sole originator of N. 6. Therefore, it seems one cannot after all be ultimately responsible for their nature, and if one cannot have ultimate responsibility over that, then there cannot be ultimate responsibility concerning one's actions either; so, there is no *ultimate* responsibility to speak of when it comes to what we actually do.

At first glance the objection looks quite effective, but does it in fact trivialise Kane's proposal? Well, this depends I think on how convinced one is by that proposal's argument. If Kane's assumption that neural indeterminacy "screens off" determination by past influences allowing agents to perform genuine acts of "will-making" seems persuasive, then there is good reason to maintain that Strawson's objection can be overcome on the grounds that, if there are moments in our lives when we are able to make undetermined decisions which go towards developing our character, then arguably the resultant character is one I am fully responsible for together with the actions stemming from it. Strawson however, may treat the assumed possibility of a mental *tabula*

¹³³ I refer here directly to Strawson's argument in *The Bounds of Freedom* pp.441-458. The argument itself is not new I believe, Nietzsche amongst others has proposed something of a similar kind, but Strawson gives it a modern interpretation and directs it in particularly towards Kane's account.

¹³⁴ Perhaps by doing self-forming actions or SFA's that Kane discusses.

rasa with partial scepticism, suggesting instead that the supposed indeterminacy even if it should exist, could never *completely* block all influences from the past. Some influences would still persist albeit in a milder form perhaps, and the agent's decisions could still be partly traced back to them even if not on every occasion, at least on some of them. But this would then make the resultant character or nature something which was only partially formed volitionally.¹³⁵ Its possessor, just as the objection maintains would of course continue to be responsible for his subsequent actions, but never in the ultimate sense.

The question of this final section asked whether Kane offers a potential way out of the free will labyrinth, specifically, whether his proposal is better than the compatibilists' solution. Then, having discussed Strawson's objection it was further questioned whether it trivialises his argument. I believe that the answer largely depends on which side of the debate one is on. For Strawson, there is no free will however you look at it, and especially none of the kind that a libertarian envisions.¹³⁶ Whereas compatibilists who already see very little need for having 'ultimate' prefixed to the notion of responsibility, find the objection useful in showing that even if we did need it we cannot obtain it, and so incompatibilism of the libertarian variety is effectively nonsensical. To discuss this further would take us too far off course, so I will close this section with saying why I think Kane's argument deserves to be defended and offer two examples of how ultimate responsibility could be maintained despite Strawson's objection, leaving the reader to decide on their own whether they find this convincing.

Dealing with the first task of saying why Kane's argument is important, my answer to this is a fairly simple one. It rests on a very basic assumption shared by others before me, namely, that compatibilism when honestly scrutinized, is no more than a "wretched subterfuge..., [and] petty word-jugglery" which needs to be treated with the outmost caution.¹³⁷ My reasons for agreeing with this characterisation lie primarily in what I think a philosopher ought to strive for. As a discipline, philosophy is arguably the kind of subject that pursues truth through careful reasoning and analysis of a wide range of available sources; the philosopher then, is also someone who wishes to reach the highest tier of knowledge and understanding available to him in the hope of discovering that truth. Applying this description to the free will issue, it seems preposterous that compatibilists feel they can defend their arguments without embarrassment when what they offer is in fact an avoidance of the truth. By insisting that we can do just as well if not better in the absence of 'ultimate'

¹³⁵ Here, 'volition' is taken to mean something which derives from the agent's rational and conscious choice, though also presupposes the absence of determination by past influences resulting from something like a pre-set will.

¹³⁶ Richard Double in 'Metaethics, Metaphilosophy and *Free Will Subjectivism*' also places Strawson into the "No-free-will-either-way theory" category, p.513.

¹³⁷ The words were originally expressed by Kant, though I take this quote from Strawson's paper to show his own distaste for compatibilism (See '*The Bounds of Freedom*' p.450).

responsibility, they are refusing to admit the logical priority and superiority of this concept, viz., they are refusing to acknowledge that this *is* what freedom ought to produce if it is to really be freedom, and if it does not, then that freedom is mistaken. And yet, they doggedly insist that this is good enough, perhaps the same way a charlatan may insist that living under synthetic light will do just as well since there is no natural sunlight.¹³⁸

Now turning attention to the second task, I will present two possibilities through which Strawson's objection may be diffused, with the second of these being the one I consider the most promising.

First Response: The whole dilemma rests on there being no evident starting point from which to claim ultimate responsibility for one's nature. It could be argued though, that such a starting point does in fact exist, and it begins its existence from the moment that persons reach sufficient mental maturity which allows them to make intentional decisions whose consequences or possible consequences they are rationally aware of. This would then suggest that *before* one has reached such a stage, it ought not to be assumed that anything like a concrete 'nature' formerly existed, rather, we may loosely speak of there being a certain set of disposition or tendencies as well as inclinations, all haphazardly placed in no definite form. From the moment we enter that stage of maturity though, so begins our nature-formation and the possibility of being ultimately responsible for actions proceeding from it.

The basic premise of this type of response assumes that there was never a nature 'M' preceding the creation of 'N'; it also assumes that N was created authentically from the agent's choices, with all former inclinations and tendencies having only an indeterminate influence.¹³⁹

Second Response: The other possibility is that we can agree that there was a prior nature M for which we were not ultimately responsible, though disagree that it *necessarily* has to influence N indefinitely. That is, we assume that despite M being given, from the time we make the decision to change our nature into N, a gradual wearing away or erosion of M's influence over us will begin so that finally a point will be reached when it is no more than an ineffectual fact of our past history. To see this more schematically imagine that there is a timeline, ranging from t_1 to t_{10} during which the agent is engaged in altering her nature. Imagine further that at t_1 that agent is still being influenced significantly by aspects of M, and N is under construction, by t_2 though that influence has already

¹³⁸ I make this comparison since I do not think many compatibilists genuinely seek to defend free will, rather they argue that we can create something akin to this idea by stretching the parameters of determinism, thereby gaining some "elbow room" as Dennett once said.

¹³⁹ For reasons of space I cannot discuss possible objections here. An interesting one which comes to mind however could argue that the stage of maturity needs to be specified, but that this may be difficult to do. Furthermore, what about those who never manage to reach this requisite stage, does it mean they can never be ultimately responsible? Does it also imply they do not and cannot possess a nature? These questions are interesting, but will have to be answered at another time.

perceptibly lessened, and by t_5 even more so. Finally, the agent reaches t_{10} and not a trace of M can be said to have survived in the resultant product N; then despite N's development *initially* being intermingled with remnants of M, its *finished* nature is arguably one whose identity is wholly the result of the agent's own arduous doing, and for whose future actions one is responsible for in the ultimate sense.

This particular reply borrows its idea from the well-known "Ship of Thessous" paradox concerning identity, and argues that if we gradually replace the parts of the original ship with new ones, we would not qualitatively be left with the same ship but rather an entirely different one.¹⁴⁰ Similarly, if we apply this reasoning to the idea that aspects of M are being slowly discarded and replaced by new ones in the creation of N, it becomes possible to argue that there is an eventual loss of identity which increases with time until a sufficient stage is reached when N can no longer be identified with M. Or simpler, we have a new nature which is not obviously 'mN'.

Both responses are controversial of course and require much more work before it may be claimed that they are effective. The last one in particular needs to develop an account of how exactly the agent would be able to discard influences from his past nature, and whether it would indeed be possible as suggested, to discard all of them. However, for the present, if the responses at least succeed in showing that there is potential to thwart Strawson's objection and preserve Kane's account of UR together with its accompanying presuppositions, then I would argue time has not been wasted and in the next and final chapter I hope amongst other things, to explain a further reason for why I believe the basic principles of Kane's position continue to be essential to our understanding and construction of a theory of freedom.

¹⁴⁰ Some may also argue we would not quantitatively have the same ship either, since the parts from the original ship can be put together again giving us two ships instead of one then.

The Myriad

I - Retracing Steps

This work began by introducing the determinist's challenge to the idea of free will. Two of the most prevalent responses in the form of compatibilist and libertarian views were then introduced and assessed¹⁴¹ where it was argued that neither approach, for differing reasons, manages to give us quite what is wanted from a theory of freedom. Standard libertarian solutions primarily fall prey to either constructing metaphysically dubious accounts of freedom, or offering virtually no account at all based on the assumption that the "will" is a unique and inexplicable phenomenon bearing limited capacity for evidential justification. Compatibilism, with its almost linear focus on preserving agent-responsibility, appears to have abandoned as a result of this some of the most intuitive assumptions that the original idea of freedom contained, producing instead what may only be described as a Frankensteinian imitation of that original because the imitation happens to be miserably grotesque in comparison to what was once conceived.

The preceding chapter concentrated on Kane's TI "causal-indeterminism" which despite incurring its own share of objections, at least exposed exactly what portion of that idea of freedom was being sacrificed for the compatibilist's cause, namely, a conception of the agent as his own 'will-maker' or 'will-setter' and with that the adjoining notion of UR. Now in this concluding chapter, I wish to give my second reason for why I believe something like Kane's position is requisite for the construction of a plausible theory of freedom, incorporating and developing one specific aspect of it into the 'alternative' that was introduced at the start of this work in order to flesh-out the latter and provide a further assumption which I consider fundamental to any respectable notion of human freedom.¹⁴² Before I do this however, I think it will be prudent to begin the discussion by taking a broad look at the myriad of successor views to hard determinism that have evolved over the years since its introduction, for I will argue they suggest something telling about the free will problem which could be of benefit to my own account.

¹⁴¹ Though I have chosen to place greater emphasis on examining the compatibilist strategies, since these pose the most risk I find. Traditional libertarian accounts on the other hand, primarily suffer from the tendency of presenting either vague or unhelpfully technical explanations of freedom neither which serve to develop this idea's full implications.

¹⁴² To clarify, though Kane's position is libertarian, I do not take it to represent that category of 'standard' approaches which I see as problematical.

II - The Breakers of Freedom

An aspect that is perhaps overlooked is that some of the strongest objections to traditional libertarian and compatibilist solutions come from determinists themselves. Their reasoning for rejecting libertarian theories ought to be self-explanatory at this point, for at bottom there is a fundamental clash of opposing viewpoints resulting from scientifically and metaphysically antagonistic belief systems.¹⁴³ It is therefore more interesting to try and understand why (hard) determinists equally dismiss compatibilist modifications, or rather, what specific element they find lacking in them collectively.¹⁴⁴

Ted Honderich's explanation is both frank and aptly succinct in this case. He maintains that something important is indeed lost once determinism is accepted, and that it is a feature which cannot be re-introduced through a compatibilist ideology though it continues to be of significance when it comes to understanding the basic hopes people have concerning their lives and the concepts these hopes entail, which in turn furnish our ordinary ideas about freedom. His argument can be put thusly: 1. 'life hopes' come in the manner of two distinct types of desires,¹⁴⁵ first; that actions be voluntary, second; that they *not* be fixed products of one's nature or environment, viz., that they are in some sense 'originated'. 2. Our social institutions provide ample evidence that we do in fact hold such life hopes, since the dual practice of preventative and retributive punishment is such that the former notion is modelled on the idea that actions are done voluntarily, while the latter assumes they were undertaken independently of one's psychosocial circumstances, or that they were direct products of one's will. 3. The "real problem" then comes down to this; both voluntariness *and* origination shape our ideas of freedom, but if determinism is true, origination is threatened and arguably lost.¹⁴⁶

Effectively, Honderich expresses the same worry that Kane's proposal attempts to solve; the idea that if actions are to really be 'free' they must not be necessitated by factors outside of our control. This is something both thinkers regard as integral to the idea of freedom and a key element that

¹⁴³ A discussion of them will not be taken up here. It suffices to say that a basic point of departure for many libertarian and determinist theorists begins with the dichotomy between deterministic causality and random or indeterminate causation (though of course, not all libertarians assume indeterminism as one of their theoretical premises).

¹⁴⁴ Though there is no unequivocal definition of "hard determinism" a possible way to appreciate its distinctiveness from standard deterministic theory, is to know that it does not support the possibility of determinism being compatible with free will, or accommodating to this notion. That is, it rejects compatibilism as much if not more as it does libertarianism. In view of this, compatibilists have often been labelled "soft determinists" precisely because they *do* believe that some respectable form of human free will can still be constructed despite the presence of a deterministic framework.

¹⁴⁵ Or as Honderich states, we are all "more than capable" of having these desires (See, p.473: *Determinism As True, Both Compatibilism And Incompatibilism As False, And The Real Problem*).

¹⁴⁶ *Ibid*, pp.473-4.

compatibilists cannot successfully deliver. As previously observed, the compatibilist is primarily interested in the responsibility component of freedom, and this to be fair he manages to partially secure albeit with questionable validity.¹⁴⁷ For the majority of compatibilists then, I suspect the loss of “origination” from the idea of freedom is seen as both inevitable and surmountable as long as it can be shown that actions could still be voluntary and so suitable for moral assessment. For Kane and Honderich though (due to differing reasons of course) this loss is not inconsequential, and reveals a fundamental crater in the compatibilist’s reasoning which gives us cause to mistrust it.

The fact that Honderich believes determinism to have fixed consequences is something I will not be addressing here per se; instead, I wish to move to a slightly different line of argumentation from Derek Pereboom who in nutshell form insists that the worry about consequences is virtually irrelevant, because free will is largely irrelevant from a social or external viewpoint.

Being a “hard incompatibilist” Pereboom not only denies the possible coexistence of free will with determinism, but also denies the existence and relevance of free will altogether. According to his particular reasoning, both compatibilism and libertarianism are necessarily refutable on the grounds that (i) “causal history” rather than AP is what determines moral responsibility, (ii) only “agent-causation” can endow persons with the requisite degree of responsibility needed for pronouncing moral judgements, *but* (iii) the notion of the agent as the source or origin of her actions is, though “coherent” not really “credible” so, (iv) we do not have free will of the kind needed for moral responsibility to obtain.¹⁴⁸

An outcome of this sort which Honderich would find troubling, does not really faze Pereboom whose general tenor suggests he takes *free will* to be a largely romanticised and overvalued notion. The loss of it need not lead to personal or public catastrophe; on the contrary, its absence is substitutable and even desirable in many respects once we open ourselves up to an alternative mode of existence as envisioned by hard incompatibilism.¹⁴⁹

¹⁴⁷ I take the matter to be debatable with the verdict being largely of a kind that depends on one’s private theoretical demands with respect to the concept under question.

¹⁴⁸ There are several aspects of this argument which I find have been insufficiently defended, the most notable being the claim that agent-causation is not a viable option since “evidence from our best scientific theories” show little support for the idea (p.478). The fact that science does not hold all the cards, and ought not to be relied on in cases where the primary aim is to disguise the absence of convincing philosophical reasoning, is a practical condition that Pereboom clearly fails to observe. However, since my present aim is to discuss several contemporary determinist accounts for the purposes of proceeding to my own argument, I will say no more about the soundness of this or any other types of determinist position introduced. (All references to Pereboom are taken from, *Living Without Free Will: The Case For hard Incompatibilism*).

¹⁴⁹ For example, Pereboom believes that the acceptance there is no free will serves to undermine the legitimacy of “moral anger” which he sees as having a predominantly destructive influence, since its expression often aims at causing “physical or emotional pain” and “as a result, moral anger tends to damage or destroy relationships...” It is because of such consequences that he argues the loss of free will to be desirable (pp.487-8).

In Pereboom's 'new world' ordinary ideas and conceptions about *wrongdoing*, *love*, or even the *meaning in life*, can all be effectively dealt with and sustained – albeit in an altered state – without the need for free will. So for instance, 'praise' and 'blame' the customary attendants of the notion of 'wrongdoing' more appropriately become replaced with "moral admonishment" and "encouragement" where the former only presupposes that a wrong has been committed.¹⁵⁰ Likewise, concerning our attitudes and hopes about 'meaning in life' or what Honderich defines as *life-hopes*, which he believes imply the psychological need to see ourselves as praiseworthy for our achievements, these too according to Pereboom can survive even if the practice of giving and receiving praise does not. *Achievement* for example can continue being a relevant concept that we use, even though its sense may be "diminished" now that we no longer view our actions as having been freely performed. However, fearing that the ability in general to entertain certain life-hopes may be compromised due to a feeling of resignation as a result of this awareness, is not really warranted either given that the future for most people remains unknown – a fact that should reasonably enable one to continue to hold on to various hopes concerning it.¹⁵¹

The essence of Pereboom's account then aims to offset the kind of worry Honderich evinces about determinism having consequences once the idea of origination is recognized as invalid, and it does this by arguing free will to be little more than a superfluous fancy of our imagination whose presence is something that humankind can easily learn to live without. Before concluding this section then, I will introduce one further perspective offered by Saul Smilansky (despite its rejection by Pereboom) which discusses the need to uphold the *illusion* of free will.

Like Honderich, Smilansky believes that the truth of determinism undermines something important both socially and privately, and he responds to this problem by arguing that the illusion of free will should be sustained, indeed must be sustained, if we are to avoid the unsettling consequences determinism implies. This "illusionism" has two sources as he sees it; first, it derives from our tendency to hold *both* compatibilist and incompatibilist intuitions concerning free will and the compatibility question itself, which more properly renders us "fundamental dualists" rather than

¹⁵⁰ A consequence of this then being that Retributivism as a measure of punishment, on the assumption that the offender *deserves* blame or pain becomes inapplicable from a hard incompatibilist's perspective (See, p.479).

¹⁵¹ Without wishing to descend into a discussion about this, I will only say that the revised image of life presented here is simply not good enough. Pereboom ignores the particularities that many life-hopes contain, amongst them arguably resides not just the 'hope' of achieving our dreams, but often the real *conviction* that we will do so, which is something I take as being sustained largely by faith or confidence concerning one's abilities and potential to succeed through effort. But if determinism is true, then this kind of faith in oneself turns out to be something one cannot really be entitled to in the traditional sense of the word. The insistence then that the worry about consequences is defeasible, does not seem especially convincing when the implications of this suggestion are considered more carefully.

monists with respect to the free will problem; secondly, we are motivated towards illusionism following the failure of libertarianism.^{152, 153}

Smilansky believes that compatibilism should be rejected because we need responsibility and control to be *ultimate*, while hard determinism is rejected because people desire to be members of a “community of responsibility” which is something that compatibilism does manage to secure by allowing persons to receive credit for actions that were clearly up to them.¹⁵⁴ The illusionism then builds on these intuitions while also reflecting the vacuum generated by the absence of libertarian free will, and so creates a form of quasi-reality about freedom which is practically necessary to help sustain our commonsense beliefs about agency and the social practice of conferring moral judgement.¹⁵⁵

The illusionism then is something that is *already* there in our psyche without the need to be advocated, functioning similarly I presume to a defence mechanism which blocks the realization that what is in fact being believed, happens to be false.^{156, 157} Nonetheless, its role is mainly “positive” and “crucial” according to Smilansky in helping to avoid the difficulties which would otherwise surface. Amongst these difficulties, is the danger of developing a compromised self-awareness of one’s moral responsibility as an acting agent upon the discovery that the majority of our actions are nothing but the unavoidable outcomes of the way things were. For this reason, Smilansky urges that it is preferable for people to never fully grasp that their doings are ultimately inevitable lest it distorts their sense of responsibility.¹⁵⁸

Some other difficulties which Smilansky identifies include, the “*Danger of Worthlessness*” and the “*Danger of Retrospective Dissociation*” which though being consequences of psychological origin, also have the potential to manifest physically resulting in likely societal disruption.¹⁵⁹ The sense of “worthlessness” is a response to the withering of our moral self-respect once knowledge of the

¹⁵² Smilansky, *Fundamental Dualism and the Centrality of Illusion* pp.491, 500.

¹⁵³ Again, I wish to refrain from objecting to the specifics of Smilansky’s position, particularly the latter claim that libertarian free will has been proven untenable as a result of Strawson’s regress argument. I have already suggested in the third chapter that the problem can be overcome, but more importantly, this kind of objection does not really threaten my alternative.

¹⁵⁴ *Ibid*, pp.492-5.

¹⁵⁵ Specifically, Smilansky suggests that those elements of our self-understanding that are merely illusory could still prove important. Illusion, he argues, “not only helps to sustain independent reality, but also is *in itself* a sort of “reality,” simply by virtue of its existence” since the “falseness of beliefs does not negate the fact that they exist for the believer ” and this he thinks is how libertarian assumptions survive despite their invalidity (*Ibid*, p.502).

¹⁵⁶ False according to the determinist, of course.

¹⁵⁷ This illusionism perhaps also undermines the very capacity to question the belief in free will, though Smilansky says nothing of this.

¹⁵⁸ *Ibid*, pp.498-9.

¹⁵⁹ *Ibid*, p.499.

“ultimate perspective” is gained and internalised.¹⁶⁰ Through this perspective we learn that the kind of person we are, including one’s desires and beliefs as well as the necessary progression of one’s life, is really a result of one’s particular “facticity” – viz., something which couldn’t be helped, being predominantly a matter of luck.¹⁶¹ Equally, “dissociation” could develop when having accepted the ultimate perspective, a person no longer feels himself able to experience genuine *remorse* for his actions, or that state of “compunction” which subconsciously assumes that things could have been otherwise if only one actually tried.¹⁶²

Now that we have arrived towards the end of this brief discussion of various determinist positions, I will try to explain what I think they all potentially reveal. We have seen on the one hand, that compatibilism as an outlet for our libertarian hopes is fundamentally rejected, yet there is also evidence it seems, that the determinist himself is uncomfortable with the image of reality produced. Why else I ask would so much energy be spent on highlighting consequences and trying to justify the position by attempting to give it a more appeasing interpretation. Of course, neither Honderich nor any of the other determinists discussed are willing to relinquish their beliefs, but neither are they prepared to state the ‘facts’ and leave it at that. Honderich is blunt about the effects, and Pereboom though obstinate, still endeavours to produce a more palatable alternative of conceiving the matter. Smilansky’s position is one I find of most interest, in that it not only acknowledges the persistence of the *free will belief* but also accepts our practical human need for it.

Is the tension indicative of anything then? My own view is that it is. Specifically, I suggest that the defensive strategies used to dissipate the tension, mask an implicit reluctance on behalf of the determinist to abandon the idea of free will for good. Before being savaged for assuming this however, let me try to clarify what I mean. There is something present in all of us to a greater or lesser degree which can loosely be described as *human nature* and though its exact meaning remains poorly understood, its influence can often be felt. The determinist’s disquietude is reflective of this influence I believe, suggesting at bottom a psychological inability to give up the ‘free will habit’ in its entirety. For this reason I maintain that from the practical standpoint of human nature, categorically relinquishing the belief in free will is impossible.

¹⁶⁰¹⁶⁰ The assumption here being that moral self-respect depends on a view of the *self* as the source of choices, actions, and achievements (see p.499).

¹⁶¹ The notion of something being a *facticity* here borrows from Sartre’s definition. According to it, it is simply a feature of our human constitution usually comprising of a set of ‘fixed’ particulars which characterise a portion of one’s identity, and in whose selection one had no freedom of choice. A person’s gender, their date of birth, the names of their parents, or their eye colour for instance, would all be included amongst such particulars. When this notion is applied to describe the nature of determinist theory however, it takes wider scope in that it now assumes that *everything* (or almost everything) about a person is a matter of facticity.

¹⁶² *Ibid*, p.500.

The attempts to provide substitutes in the hope of convincing others and oneself that very little needs to change if free will is left out, primarily speaks of a subconscious effort to maintain some aspect of its perceived value active and to appreciate the value is to continue to hold on to the source itself. What the situation reveals then is a fundamental tension between two existing standpoints – the *practical* and the *theoretical* – in virtue of which, the determinist like any other human being struggles to indefinitely sustain practically what he believes theoretically. That is, the abstract belief that free will is illusory conflicts with the private need for it not to be.¹⁶³

It is because the determinist himself cannot psychologically let go of free will that he seeks refuge in theoretical justifications which continue to indirectly utilise the concept. In Smilansky's account this tendency is clearly apparent, in Pereboom's a little less so, but the alternative he offers makes no secret of his intention to mainly preserve rather than replace the social function performed by this notion. But why try to preserve the effects of something whose existence one theoretically rules out? I argue; perhaps because a small nagging voice of inner sense tells you that just like the rest of humankind, you too cannot rid yourself completely of its influence – forever depending on its continued presence in some form.

III - Mending Freedom: The Alternative

I have argued somewhat controversially that even the determinist seems incapable of fully abandoning the idea of free will. But all this presently suggests is that there are certain conditions which make it difficult for humans to give up a spurious belief, not that the belief could be true. Therefore, in this concluding section I would like to concentrate on developing a different means of introducing the latter as a possibility, by initially focusing on a specific area of Kane's argument which will then enable me to resume discussion of the alternative that was casually outlined in the first chapter. This I will pursue in the following three stages: (i) Kane and libertarianism (ii) the alternative (iii) closing thoughts.

(i) – Kane's contribution to the free will debate is notable; not only does he offer a sufficiently intelligible account for the feasibility of libertarianism, but his examination of the relevant issues also happens to be more comprehensive than that of many of his contemporaries. Amongst the things he addresses, the notion of "will-setting" or "character-setting" through the performance of self-forming actions or SFA's is arguably a central one in that it enables persons to attain ultimate

¹⁶³ The existence of two possible standpoints is something that was usefully suggested to me during a discussion with Professor Jane Broadie. Originally, my explanation made heavy use of the notion of 'intuition' to account for the feeling of freedom, but this was felt to lack the precision necessary to make the idea clear. The concept of a *practical* perspective seems to work better however, in that it explains the probable cause of such intuitions.

responsibility over their actions. And since most libertarians hold only this form of responsibility applicable, one of the strengths of Kane's incompatibilism then is that it demonstrates precisely just where the compatibilist theory stumbles.

Presently however, responsibility is not what I wish to predominantly focus on, but rather a different outcome of Kane's approach which is less obvious perhaps, but one that introduces another equally important feature of free will, viz., the possibility for *personal self-reform*. Kane of course does not explicitly discuss this, since his main priority is to secure UR, but the idea that persons could be capable of permanently transforming their nature is one arguably realisable only under a libertarian doctrine of the will, or something similar in kind. After all, Kane's own hypothesis of neural indeterminacy already assumes that there is potential to shape one's nature, so why not go a step further and demand that a theory of freedom also allow us to 'reshape' it if we choose?

At this point however, I feel Kane's and my position beginning to drift apart; of course, it may be possible to argue quite reasonably that the neural indeterminacy would just continue even after the completion of will-setting, and that this *may* leave open the possibility for persons to reform an already existing nature if they wanted to. However, I think Kane's proposal would struggle with this kind of addition, its defence being likely to proliferate into numerous sub-explanations that eventually cause the original argument to become undesirably convoluted while the soundness of this extra feature continues being uncertain at best.

Kane's argument would indeed be stretched to its limits if he supplements it with the aforementioned notion; however, I do not think that the central challenge lies in the potential difficulty of accommodating the idea of personal self-reform, rather, it is the general indifference towards it which identifies the real obstacle. Let me explain. The basic principles of libertarianism are ones that my own position happens to echo, and I have already insisted that only this type of doctrine of the will could produce anything remotely satisfactory when it comes to formulating a theory of freedom. However, where the libertarian's and my opinion seem to diverge concerns the particular aspect of freedom one is interested in, or finds central. Like the compatibilist the majority of libertarian scholars primarily focus on securing agent-responsibility while giving no attention, or very sparse attention to other relevant features of freedom.

Responsibility then, is arguably an area most moral philosophers interested in free will choose to concentrate on as if it presented the only conceptually justifiable motive for valuing freedom or wanting to defend our idea of it. I certainly do not wish to give the impression here that I think the endeavour trivial, but I do wish to maintain that responsibility only partially explains why free will

matters. To concentrate principally on this feature then, is to treat freedom as a “mere means”¹⁶⁴ and to see it for its consequences alone, without ever trying to properly understand this concept in its entirety.

Effectively then, where the libertarian and I seem to differ is in our approaches to the free will question, for though I agree that defending agent-accountability is a necessary requirement, I do not believe that this is all we should be concerned about when faced with the threat of determinism. The determinist does threaten something important, but it is not *just* responsibility that is at stake, viz., not just the *social* significance of freedom that is being undermined, it is also its *private* value which suffers. Despite this however, most efforts continue to be channelled towards preserving the former at the expense of the latter, and yet the latter matters too – possibly more.

When speaking of the importance of defending the private value of freedom then, I am specifically referring to the way in which persons relate to their freedom, that is, how they might understand and use this concept. In response, I propose that the understanding of this aspect of one’s freedom assumes a simple ‘I can...’ mentality which the individual habitually presupposes in virtue of believing oneself to be free to embark upon any number of desirous alternatives. This mentality then arguably provides the basis for more complex ideas, such as the idea of personal self-reform, which here speaks of an individual’s implicit conviction that one is free to try and pursue a different course of action which is likely to result in the unfolding of a different life path.¹⁶⁵

Being at the end of this particular part of the discussion, I will conclude with two thoughts. First, I have argued that the libertarian perspective offers the best theoretical platform from which to pursue a theory of freedom, and within this perspective I have insisted Kane’s analysis to be the most rigorous as well as bearing the most potential.¹⁶⁶ However, I have also suggested that a possible limitation of this approach is that it primarily concentrates on the “social” signification of freedom without really trying to explore the “private” meaning this concept contains and that as a consequence we continue to have only fragmentary understanding of its value.

Secondly, to defend freedom from its private standpoint is logically more advantageous, since establishing its validity from this basis allows us to simultaneously establish the validity of responsibility. Namely, since responsibility is arguably just a natural by-product of having a free will, if it is shown plausible that persons are free in the way they think they are, then it is *necessarily* the

¹⁶⁴ I here adopt a Kantian turn of phrase in order to emphasise the general tendency to view freedom solely as a means towards something, in this case the obtainment of responsibility, rather than seeing it as an end in itself too.

¹⁶⁵ That is, a different course of action from that which one would standardly take, or is used to taking, or even continues to *want* to take perhaps.

¹⁶⁶ It is rigorous because its examination of the issues is thorough, and it bears the most potential because the explanation it offers is both lucid and plausible despite the lack of present confirmation for it.

case that they are also responsible agents¹⁶⁷. Whereas, if we merely approach freedom from its social standpoint, though we may be able to devise explanations that show we still bear responsibility for our actions, we may not always, on closer examination be able to identify anything about those actions that would suggest they were free in the private sense of freedom, that is, that they were *originated* or *self-willed* rather than just voluntary.¹⁶⁸

(II) – I come now to the penultimate part of the work where it is my intention to elaborate more fully on the alternative sketched at the beginning. This chapter started with a discussion of successor views to hard determinism and argued that they reveal the determinist's implicit reluctance to surrender the belief in free will. The second half then concentrated on defending Kane's variety of libertarianism where it was suggested that the idea of will-setting gave credence to a further one, namely that of personal self-reform. However, it was also maintained that Kane's theory would struggle to accommodate this additional possibility and the reason for it was found to reside in the nature of libertarianism itself which like the majority of free will theories, prefers to focus on the responsibility component of a person's freedom.

In this section then I wish to try and bring together two elements of the discussion so far, this being on the one hand the need to present an alternative account of freedom to show how our basic belief in this idea could be true or at least sustainable, while on the other hand the desire to defend it in terms of its *private* value for the individual which takes responsibility into account but does not defend it solely for the sake of this feature. Towards a possible attempt at this I now turn then.

To begin, I would like to clarify that my own position regarding the free will issue is not classically libertarian though it could be described as a species of it because it does hold that free will and determinism are incompatible, and that a theory of freedom needs to contain origination if it is to be genuine. However, unlike the libertarian I aim to avoid justifying the belief in free will by suggesting elaborate metaphysical postulations, or borrowing from the most recent finding of science, rather, I approach the question of whether human beings are free via a purely rational domain which offers neither 'proof' nor 'evidence' in the customary sense of these terms but instead, appeals to a person's own judgement and asks one to trust it. For this reason I have tentatively defined my

¹⁶⁷ In other words, if it is shown that we do have the capacity for character-reform together with the requisite ability to intentionally form alternative decisions and pursue alternative courses of action.

¹⁶⁸ Though this objection applies more to compatibilists, I think the libertarian is also not entirely immune to it either, for despite being concerned with establishing the agent's originative capacities, these are almost wholly esteemed *for* their ability to supplement responsibility with the 'ultimate' component and so receive very little analysis beyond that. When it comes to compatibilism however the general tendency towards producing quite distorted accounts of freedom is plausibly explained by two main factors, the first argues that the compatibilist is really a covert determinist and the second, that he regards the relevance of freedom solely from and for its social position.

position as a *Rational Conviction View* in accordance to which I will now try to present my argument for free will.

If one recalls, the first chapter presented a key premise for accepting the possibility that one could transcend their “organic freedom” in favour of their “acquired freedom” which essentially maintained that the belief in free will was incapable of ultimate proof because the free will question itself could never be definitively settled. For this reason it was concluded that everything really comes down to faith, which consists in the individual’s ability to make a practical choice and trust his personal judgement as regards the matter. He could ultimately be wrong, he could ultimately be right, yet which he will never know and it does not really matter either I argued because the faith in freedom is *basic* to the wellbeing of a person.

The dichotomy postulated earlier then between the practical and theoretical perspectives could now be used to formulate the above premise more coherently. My suggestion is that one is rationally entitled to maintain a lively conviction in free will not only because the determinist’s arguments remain largely inconclusive, but also because the attempt to abandon the practical perspective from which we see ourselves as free for the theoretical one which tells us that we’re not is both unrealistic and arguably impossible.

As human beings we are tied to our own subjective awareness of things and to the practical perspective which forces us to relinquish whatever theoretical musings we may entertain from time to time, or to at least confine them to just those quiet moments of reflection. The practical standpoint also arguably helps explain the psychological phenomenon of why we intuitively hold the belief in freedom and why subsequently, there exists an additional capacity for persons to rationally take a leap of faith and judge themselves to be free. In other words, the practical perspective provides a convincing foundation from which a person can obtain both the *intuition* and the *faith* in free will, and since my theoretical stance argues that the only certain ‘proof’ at our disposal is in fact a form of faith in the rationality of one’s own convictions, I suggest then that this perspective offers one way, perhaps the only way directly available to us through which these convictions could in one sense be legitimized and our faith subsequently reinforced.

In arguing that one’s conscious experience of occupying the practical perspective provides rational grounds for upholding faith in free will I am not simply claiming that this perspective carries its own truth, instead I am suggesting that whatever ‘truth’ may be obtained from it is *dependent* on our own phenomenological assessment as creatures who live in the world and respond to it in appropriate ways. The practical perspective then acts as a channel through which human beings “live out” their various beliefs, with some eventually being discarded while others kept based on the private nature of one’s experiences and the subsequent judgement of reason which followed them.

Our faith in freedom is something that arguably results from judging one of those beliefs to be valid and the only evidence directly available to us for the appropriateness of this belief comes from the individual's own rational assessment of his experiences, which are accessible precisely because he *exists* only within the practical perspective.

I conclude the present segment of the discussion by suggesting that the determinist has not done enough to preclude one from exercising a rational faith in free will.¹⁶⁹ Moreover, since ultimate settlement of this issue is beyond mortal means, I maintain that one of our best sources of evidence remains our own subjectivity which for all its imperfections, induces its own set of convictions with the capacity to suspend whatever doubts a more objective approach may foster. These convictions in turn arguably carry their own grains of truth which though rejected by theory, continue to be defended by practical experience which aids belief in the reasonableness of one's judgements.¹⁷⁰

Having described my position on the free will issue I will now end this section by turning my attention towards explaining the particular conception of freedom I support. Specifically, I will state my reasons for why I believe the *private* value of freedom to be so integral to our understanding of this concept by concentrating on the previously introduced notion of personal self-reform.

Recalling what was said about compatibilism, I argued that the image of freedom it presented was both distorted and impoverished, with the individual 'boxed' into a pre-set life path that carried very limited opportunities for making significant changes. Indeed, one was 'free' as long as one stayed within the preordained course and followed it voluntarily, but not if one wanted to change it and see a different set of opportunities become available and more importantly realisable. The compatibilist doctrine then at best provides a workable account of agent-responsibility¹⁷¹ in other words, it addresses the *social* aspect of freedom but that as I argued is only one side of the coin.

¹⁶⁹ Indeed, I will go further and state that he can never do enough in acknowledgement of what has been discussed.

¹⁷⁰ I regret that there is insufficient space left to discuss C.A. Campbell's and Richard Double's arguments since both I believe make claims of a similar nature to the ones just evinced. In particular, Campbell defends what he calls the "vulgar" sense of free will as maintained by the "plain man" in opposition to the "refined" sense which is sympathetic to deterministic psychology or deterministic metaphysics. Asserting free will in the vulgar sense to be a "fact" one of his key assumptions state that the basic error committed is confining ourselves to "external observation" while ignoring the standpoint of "living experience" that is, dismissing the internal standpoint. In light of this, he further suggests that though it is possible to disbelieve in free will when thinking about the situation "abstractly" the same cannot be done when we are actually "living" it. (See his, *In Defence of Free Will*, 1938). Double equally presents a very interesting analysis of the issue, at the heart of which lies the claim that the belief in human freedom is largely a subjective matter with both free will and responsibility said to reside "only in the eye of the beholder" and nowhere else. For this reason, Double insists that the whole free will debate is more properly a psychological dispute with no definitive answer or truth obtainable (See, *Metaethics, Metaphilosophy and Free Will Subjectivism*).

¹⁷¹ Though in all honesty I do not think it provides even that as long as what one does was ultimately necessitated by factors outside of one's sphere of control. Compatibilism may not lead to futilism, but it does render the person into a willing slave.

In the first chapter I offered an initial model of the kind of freedom it was felt a human being ought to be capable of, or the kind of freedom whose existence one could realistically assert. That model contained two core ideas, that of “organic” and “acquired” freedom where the former depicted a possible starting state of the human being characterised by the presence of certain ‘given’ psycho-physical features, while the latter reflected a definite rejection and overcoming of certain inherent character traits and dispositions. Lastly, it was argued that the will could be used as a medium between the two types of freedom for transcending the organic state in favour of the acquired one where this transcendence was seen as purely voluntary and therefore optional, consisting solely of the individual’s *effort* of will.

The above representation was created with the notion of personal self-reform in mind, and helps demonstrate the centrality of the private aspect of a person’s freedom which arguably renders a defence of it both necessary and desirable. For if one stops to reflect on why someone may wish to be in possession of a free will, I doubt that the first answer which comes to mind will say ‘because we want to be held accountable for our actions’ instead, I suggest that having freedom matters to a person *primarily* because one wants to believe in one’s own potential, perhaps even consciously needing to believe in it in order to lead a purposeful life, and only secondarily because one also desires that his conduct be capable of moral assessment. Furthermore, the ability to reform, to permanently change one’s ways so to speak, not only provides the individual with hope and purpose but also it seems defines their personhood. That is, being a person does not – as Frankfurt once insisted – singularly consist in the ability for having second-order desires, it consists jointly in the *capacity* to see those desires come to fruition and to know that one’s effort was both responsible and sufficient for their realisation.

The notion of self-reform then contained within the idea of *acquiring* a different tier of freedom by ascending to it through sheer effort of will, reflects aptly it seems a person’s inner need to believe in one’s potential,¹⁷² while the model for it offers a plausible structure explaining how this potential could be actualised. Before leaving the discussion though, I would like to elaborate a little on the model’s rationale in order to justify my motivation for the preceding argument.

The following extract from Nietzsche’s ‘On the Genealogy of Morals’ contains a set of passages where he is seen rejecting the “common” person’s conception of the origin of ‘good’ because it assumes that people are always at liberty to choose to act against their basic nature:¹⁷³

¹⁷² I am not suggesting that self-reform perfectly defines what it means to believe in one’s potential, the latter idea need not consist of a desire to change of course, it could just as easily be characterised by the individual’s conviction that one is capable of accomplishing certain set goals in virtue of one’s present character etc. Nevertheless, I do think that self-reform as a notion describes a substantial portion of what a belief in one’s potential implies.

¹⁷³ The common person is also typically described by Nietzsche as “the man of *ressentiment*” (p.29).

“That lambs bear ill-will towards large birds of prey is hardly strange: but is in itself no reason to blame large birds of prey for making off with little lambs.”

“To demand of strength that it should not express itself as strength [...] makes as little sense as to demand of weakness that it should express itself as strength”

“Just as the common person distinguish lightning from the flash of light and takes the latter as doing, as the effect of a subject which is called lightning, just so popular morality distinguishes strength from expressions of strength, as if behind the strong individual there were an indifferent substratum which was at liberty to express or not express strength. But no such substratum exists; there is no ‘being’ behind doing, acting, becoming: ‘the doer’ is merely a fiction imposed on the doing...”¹⁷⁴

On a first reading of the extract one may either be pleasingly convinced or secretly repulsed at its image of the human being, seemingly depicted to be virtually indistinguishable from a predictable automaton or some lower species of animal. The impression this left, set me thinking about the reasonableness of insisting on the presence of a capacity for overcoming some of one’s adverse inclinations and the accompanying possibility of changing one’s first-order desires with the aim of reforming one’s character for good.

Of course, the wish to want to defend one’s intuitive beliefs is arguably natural and the belief in free will for those who hold it, is no exception. But it is also important to be sensible about how much can be expected from this concept without driving it to absurdity. I have already presented my argument for why a belief in free will ought to be permitted, this being largely a matter of taking a rational leap of faith which the practical perspective arguably permits us to do, but now I wish to qualify the notion of self-reform itself which constitutes the private value of a person’s freedom, by supplementing it with one other idea which could potentially help delineate the appropriate boundaries of this freedom and thereby specify the meaning of that self-reform.

The idea I wish to introduce then is known as the doctrine of *causal essentialism* and it argues that for every individual substance be that a person or some other form of living being, that substance has certain ‘essential’ properties whose influence is causally primary with respect to the future history of that substance. This doctrine, though implying that the properties function as natural determinants of a substance’s possible trajectory, entails neither classical determinism nor fatalism but only suggests that there is an organic limit to the kinds of things each individual substance could reasonably accomplish.¹⁷⁵

In view of this, there is a given body of scholars who insist that Nietzsche himself subscribed to this type of doctrine, and so interpret passages such as the one previously given as evidence that he

¹⁷⁴ Ibid, underlining added.

¹⁷⁵ See, Brian Leiter, p.225: *The Paradox of Fatalism and Self-Creation in Nietzsche*.

simply believed the essential natural facts about a person to limit the possible range of 'life-paths' one could be seen realising. In other words, what one may ultimately become seems fundamentally determined by the scope and identity of one's unique trajectory, which in turn implies that a person's freedom could never be absolute if by that is meant unlimited, but rather remains a feature that will always be significantly circumscribed.¹⁷⁶

I take the doctrine of causal essentialism as something that *may* be useful to my account of the private value of freedom for the following reasons; firstly, as already suggested I believe it introduces a plausible type of limiting factor that is capable of describing the natural boundaries of a person's freedom, though without trivialising its meaning. Secondly, the doctrine could be employed to better illustrate the notion of self-reform contained within the idea of a person transcending one's organic freedom in favour of the acquired. Namely, the "organic" freedom can be better explained as that which constitutes a person's *essential* properties which restrict but do not necessarily control what may be achieved. In this case then, a person remains at liberty to try and reform oneself though the exact nature and extent of that reform will remain subject to that person's given trajectory and its relative scope. Thirdly, in applying this doctrine to my account of freedom, I am able to offer a more precise explanation of my reasons for finding both determinism and compatibilism so objectionable. The former is rejected because it realistically amounts to the claim that a person's *particular* life-path could be and is pre-chosen for them, being permanently set by circumstances outside one's control and irrespective of the original scope of one's given trajectory, while the latter is fundamentally denied for its inability to deal with this consequence.

So causal essentialism is a potentially useful supplement to my account because it imposes some practical limits on our idea of freedom in order to prevent it sounding flippant, and yet implies that within those limits one is perfectly free to transition from one possibility to the next and so free to engage in the kind of self-reform which permits one's aims to be both diverse and realisable. Having said this, the doctrine also arguably has its problems of which I will mention just one, since I take it as central, together with a possible response before proceeding to the conclusion.

It is apparent from the examples in Nietzsche's passages, that he takes the psycho-physical characteristics of a person as representative of one's *essential* properties, and therefore features which plausibly could not be altered. This explains why he asserts that one is never at liberty to "express or not express strength" for instance, and why the resulting image of the person may seem so unsatisfactory to someone like myself who has insisted throughout the discussion that we must be in possession of the kind of freedom that would allow control over this, viz., that would permit

¹⁷⁶ Ibid.

one to overcome certain natural tendencies through effort of will, thereby utilizing the capacity for self-reform.

It would appear then that the notion of causal essentialism is not as compatible with my account as initially conceived, perhaps even directly hostile to it, but I maintain this need not be so. Nietzsche's essentialism is arguably particularly restrictive, and for that reason it might be objected that it remains debatable concerning which of a person's properties are to be determined as essential. It could be all the physical ones, likewise it could also be all the psychological ones, but what if it's neither? Or perhaps, there is just one particular property which can fairly be assumed to play any causal role at all – that of being *human*, for instance. The point being, it seems to be an open question concerning which properties are to be seen as essential, and even if this could be settled, there will be a further need to explain how a given set of properties – for instance, the psychological ones such as fear – could actually limit a person's range of possibilities, that is, how they could *necessarily* rule something out to the extent that one would not be capable of transcending them in order to achieve character reform.

(III) – The free will issue has been drowned in a myriad of conflicting accounts and theories, and I do not think that the present discussion will have done enough to diffuse them. Its primary aim was to offer an alternative explanation of how it may be possible to continue believing sincerely in the idea of free will, and to present a distinctive conception of a person's freedom in virtue of that idea. If I can provide a short assessment of the primary theories examined within the course of the whole work, I would say the determinist is reckless, the compatibilist pusillanimous, and the libertarian compromisingly bold. Determinists fail to appreciate the private value of freedom, though arguably struggle themselves with the image of mankind they try to perpetuate, and compatibilists though with a slightly better awareness of what could be lost, disappoint with their marginal attempts which offer only deceptive palliatives and not what is really needed. The libertarian alone remains to shoulder the burden of keeping the idea of freedom afloat, but has little to sustain him apart from commonsense and so continues to be at risk of having it capsize.

As for my own contribution, though it is still largely underdeveloped its central claim is nonetheless distinct and argues that the validity of our belief in freedom primarily derives from taking a rational leap of faith which the human being is entitled to do, and indeed ought to do given that the matter is incapable of being settled by any other means. And though in taking this leap one will not be able to establish certainty or proof, I argue that the person who does take it will already have gained something valuable that makes his own life not just another form of existence but a purposeful model of *living*, the kind that is consistent with the notion of persons as creative beings

who are active rather than passive and who need to believe in their potential to be the architects of their own lives – otherwise, I fear we may be resigned to give up on ourselves on account of the determinist's doubt.

References

1. Bernstein, Mark; 'Fatalism' – *The Oxford Handbook of Free Will*: Ch, 2; 2002 Oxford University Press, Robert Kane.
2. Bishop, C. Robert; 'Chaos, Indeterminism and Free Will' – Ibid. Ch, 5.
3. Clarke, Randolph; 'Libertarian Views: Critical Survey of Noncausal and Event-Causal Accounts of Free Agency' – Ibid. Ch, 16.
4. Campbell, C.A, '*In Defence of Free Will*', 1988 Lecture.
5. Double, Richard, 'Metaethics, Metaphilosophy, and Free Will Subjectivism' – Ibid. Ch, 23.
6. Frankfurt, Harry; '*Alternate Possibilities and Moral Responsibility*' 1969, *The Journal of Philosophy*; pp. 829-839.
7. Frankfurt, Harry; '*Freedom Of The Will And The Concept Of A Person*' 1971, *The Journal of Philosophy*; pp. 5-20.
8. Ginet, Carl; 'Reasons Explanations of Acton: Causalist versus Noncausalist Accounts' – Ibid. Ch, 17.
9. Ginet, Carl; '*Can The Will Be Caused?*' 1962, *Philosophical Review*; pp. 49-55.
10. Hodgson, David; 'Quantum Physics, Consciousness, and Free Will' – Ibid. Ch, 4.
11. Honderich, Ted; 'Determinism as True, Compatibilism and Incompatibilism as False, And the Real Problem' – Ibid. Ch, 20.
12. Hugo, Victor; '*The Last Day Of A Condemned Man*' Trans. Christopher Moncrieff, 2009, Oneworld Classics.
13. Kane, Robert; 'The Contours of Contemporary Free Will Debates' – Ibid. Introduction.
14. Kane, Robert; 'Some Neglected Pathways in the Free Will Labyrinth' – Ibid. Ch, 18.
15. Leiter, Brian; 'The Paradox of Fatalism and Self-Creation in Nietzsche' – Janaway, Christopher; *Willing And Nothingness Schopenhauer As Nietzsche's Educator*: Ch, 8; 1998; Oxford University Press.
16. Nietzsche, Friedrich; '*On The Genealogy of Morals*' Trans. Douglas Smith, 1996; Oxford University Press.
17. Pereboom, Derek; 'Living Without Free Will: The Case for Hard Incompatibilism' – Ibid. Ch, 21.
18. Pascal, Blaise; '*Pensées*' Trans. A.J. Krailsheimer, 1995, Penguin Books.
19. Strawson, Galen; 'The Bounds of Freedom' – Ibid. Ch, 19.
20. Strawson, P.F. '*Freedom and Resentment And Other Essays*' 1973, Methuen & Co LTD, pp.1-25.
21. Smilansky, Saul; 'Free Will, Fundamental Dualism, and the Centrality of Illusion' – Ibid. Ch, 22.

22. Taylor, Christopher & Daniel Dennett; 'Who's Afraid of Determinism? Rethinking Causes and Possibilities' – Ibid. Ch,11.