



## ORBIT - Online Repository of Birkbeck Institutional Theses

---

Enabling Open Access to Birkbeck's Research Degree output

The problems of defective desires, dead sea apples, and intrinsically quirky desires for unrestricted non-mental state actualist desire theories of welfare

<https://eprints.bbk.ac.uk/id/eprint/40425/>

Version: Full Version

**Citation: Mariqueo-Russell, Atus (2019) The problems of defective desires, dead sea apples, and intrinsically quirky desires for unrestricted non-mental state actualist desire theories of welfare. [Thesis] (Unpublished)**

© 2020 The Author(s)

---

All material available through ORBIT is protected by intellectual property law, including copyright law.

Any use made of the contents should comply with the relevant law.

---

[Deposit Guide](#)  
Contact: [email](#)

THE PROBLEMS OF DEFECTIVE DESIRES,  
DEAD SEA APPLES, AND INTRINSICALLY  
QUIRKY DESIRES FOR UNRESTRICTED  
NON-MENTAL STATE ACTUALIST DESIRE  
THEORIES OF WELFARE.

Atus Mariqueo-Russell

Birkbeck, University of London

Master of Philosophical Studies (MPhilStud) thesis

October 2018

## Contents

Abstract.....	2
Introduction .....	3
Chapter 1: Welfare and desire theories.....	6
Chapter 2: The problem of defective desires .....	14
Chapter 3: The problem of Dead Sea apples .....	25
Chapter 4: The problem of intrinsically quirky desires .....	42
Conclusion.....	46
Bibliography: .....	47

## Abstract

Unrestricted non-mental state actualist desire theories of welfare claim that it is the fulfilment and frustration of our actual desires that determines how well our life goes for us. This paper defends this theory against a set of arguments that are often taken to reduce it to absurdity. It is sometimes claimed that unrestricted non-mental state actualist desire theories are unviable because some of our actual desires seem to be an intuitively inadequate, repugnant or bizarre basis for welfare determination. In response to this problem, some desire theorists have abandoned the actualist theory in favour of an idealisation theory of welfare. Other desire theorists have preserved the premise that the fulfilment and frustration of actual desires determines welfare and have augmented the theory with a 'restricted' desire theory in response to these problematic desires. The desires that serve as counterexamples to the claim that actual desires determine welfare have been referred to by different names in the literature on this topic. However, I have opted to go for the umbrella terms defective desires, Dead Sea apples, and intrinsically quirky desires to categorise the different arguments, based on the identification of intuitively inadequate, repugnant or bizarre desires, leveraged by critics as undermining the unrestricted non-mental state actualist desire theory. This paper claims that upon inspection none of these desires need undermine the unrestricted non-mental state actualist desire theory. Therefore, while there may be reasons to adopt an idealisation over an actualist account, or a restricted over an unrestricted account, these reasons are not to be found in the counterexamples presented by defective desires, Dead Sea apples, and intrinsically quirky desires.

## Introduction

Unrestricted non-mental state actualist desire theories of welfare claim that it is the fulfilment and frustration of our actual desires that determines how well our life goes for us. According to this view, desire fulfilments are the only source of non-instrumental welfare generation, and desire frustrations are the only source of non-instrumental welfare detracting<sup>1</sup>. This paper defends the claim that unrestricted non-mental state actualist desire theories are not undermined by examples of desire fulfilments and frustrations that seem intuitively absurd, repugnant or bizarre to accept determine welfare. The examples that I survey fall into the categories of: 'defective desires', 'Dead Sea apples' and 'intrinsically quirky desires'. Defective desires are those desires where fulfilment intuitively seems to effect welfare negatively rather than positively. Dead Sea apples are those desires where fulfilment seems *completely* devoid of any positive welfare enhancement whatsoever. And intrinsically quirky desires are those desires where it seems counterintuitive to claim that fulfilment effects welfare positively due to certain intrinsic features present in these desires. There is some overlap between these categories. For example, Dead Sea apples and intrinsically quirky desires can be subsets of defective desires if they lead to negative welfare generation. Nevertheless, despite this overlap, I have found this tripartite distinction to be useful primarily because it serves as a way of framing the different solutions that are required to explain the problems posed by each of these different sets of desire fulfilments. It should be noted from the outset that there is no guarantee that idealisation or restricted desire theories can convincingly solve the problems posed by this collection of problematic desires. Rather, my claim is that if idealisation or restricted desire theories are motivated by a need to account for these counter-intuitive desire fulfilments, then this motivation is unnecessary because unrestricted non-mental state actualist desire theories have the available resources to explain these desire fulfilments convincingly. My thesis is divided into four separate chapters, which are preceded and followed by an introduction and conclusion.

My first chapter outlines some of the basic concepts and terminology used in discussions of philosophy of welfare. I begin by first considering what we mean when we discuss a person's welfare, and the sorts of arguments that writers have used to defend their preferred conceptions of welfare over rival candidates. I then outline the features of unrestricted non-mental state actualist desire theories, and some of the considerations that often motivate their adoption. The purpose of this discussion is not to show conclusively that unrestricted non-mental state actualist desire theories are obviously correct or more representative of our pre-theoretical intuitions about welfare than rival theories. Rather, I use this chapter to frame the discussion that follows, so that it is clear what unrestricted non-mental state actualist desire theories are, and why some people find that these theories appeal to their intuitions.

My second chapter considers the problem of defective desires. More specifically, it examines how unrestricted non-mental state actualist desire theories can handle cases of ill-informed and irrational desire fulfilments that seem to lead to consequences that detract from welfare rather than

---

<sup>1</sup> I have opted to use the term 'non-instrumental' value rather than 'intrinsic' value. This is because the term intrinsic suggests that valuable phenomena contain their value independently of their relationships to other phenomena. Conversely, the term non-instrumental allows for the possibility that things of value derive their value partly from the relationships they have with other things. In Carson's words, 'There is no reason to think that everything that has non-causal or non-instrumental value possesses value in isolation, and there is no reason to think that things that possess value in isolation always possess non-causal value when they occur as parts of broader states of affairs' (Carson, 2000: 156). Writers on welfare tend to use intrinsic and non-instrumental interchangeably, so I have opted to regularise terminology to non-instrumental.

enhance it. I begin by outlining the problem of defective desires and the general form of restricted and idealisation accounts that have been given in response to it. The chapter then argues that the problem of defective desires can be answered without recourse to idealisation or restricted desire theories by claiming that ill-informed and irrational desire fulfilments can be explained as being all-things-considered bad for welfare, rather than completely devoid of welfare. The all-things-considered bad for welfare argument claims that these examples show that welfare is decreased overall by a desire's fulfilment, but that these desire fulfilments nevertheless contain some compensatory welfare generation so as not to undermine the premise of unrestricted non-mental state actualist desire theories that *every* instance of desire fulfilment non-instrumentally enhances welfare.

Chapter Three considers cases of Dead Sea apples. These are the types of desire fulfilment that intuitively appear to be completely devoid of any positive welfare generation whatsoever. Consequently, Dead Sea apples cannot be explained by the all-things-considered bad for welfare argument. I consider the response offered by Chris Heathwood that we can answer the problem of Dead Sea apples by adopting the 'concurrency view' of welfare. The concurrency view claims that we must be desiring a state of affairs at the same time as that state of affairs occurs in order for any welfare to be generated by a desire fulfilment. Consequently, Heathwood claims that Dead Sea apples do not exist because the examples given of them fail to describe instances of actual desire fulfilments. While the concurrency view does solve the problem of Dead Sea apples, I argue that adopting it entails assenting to a whole host of other theoretical commitments that we may have reasons to want to avoid. Instead, I propose that the problem of Dead Sea apples can be ameliorated by adopting a 'composite desire theory'. This theory claims that examples of Dead Sea apples often fail to describe instances of actual desire fulfilments because typically actual desires specify several clauses as prerequisites for fulfilment. I argue that examples of Dead Sea apples seem intuitive because often they appeal to the fulfilment of singular clauses of composite desires. In these cases, we mistakenly believe that a desire has been fulfilled when instead only a singular clause of a composite desire has been fulfilled. I argue that if we adopt this understanding of typical actual desires, then we can explain many examples of Dead Sea apples as simply failing to describe actual desire fulfilments. Moreover, I claim that those examples of Dead Sea apples that do describe actual instances of desire fulfilment can be unproblematically explained by unrestricted non-mental state actualist desire theories. I do this by pointing out that unrestricted non-mental state actualist desire theories do not require subjective feelings of desire satisfaction or pleasure to be generated in every instance of welfare enhancement. Consequently, to explain some examples of seemingly Dead Sea apple desire fulfilments we can make the claim that some welfare is non-instrumentally generated even in the absence of subjective satisfaction. Therefore, we ought not be troubled by the claim that some experiences that we may intuitively call Dead Sea apples nevertheless do contribute to our welfare.

Chapter Four considers the problem presented by intrinsically quirky desires. Specifically, I consider the issues presented by base desires, poorly cultivated desires, pointless desires, and artificially aroused desires. All these desires have all been said to contain intrinsic features that disqualify their fulfilment from effecting welfare positively. The examples that I have grouped together under this category are all desires that seem to be deeply held by individuals. Moreover, the fulfilment of many of these types of desires often generates subjective mental states of pleasure or desire satisfaction. Nevertheless, critics point towards intrinsic features of these desires that seem to undermine the claim that they enhance welfare when fulfilled. I examine some of the responses given to alleviate these concerns, and claim that we have good reasons to think that none of these desires contain intrinsic features that undermine the claim that their fulfilment non-instrumentally enhances

welfare. This chapter is considerably shorter than the preceding two, as the problems posed by these desires have already been thoroughly examined by the literature.

Overall, this thesis seeks to show that many of those desires often said to undermine the viability of unrestricted non-mental state actualist desire theories fail to do so. The thesis does not claim that there are no good reasons to adopt an idealisation or restricted non-mental state desire theory. It simply claims that the ability to solve the problems of defective desires, Dead Sea apples, and intrinsically quirky desires is not a reason to opt for these theories over unrestricted non-mental state actualist desire theories.

## Chapter 1: Welfare and desire theories

### Introduction

Philosophy of welfare considers the question of what makes a life go non-instrumentally well for the person living it. The term welfare is widely used to capture the value that tracks the feature(s) that determine how non-instrumentally well a life goes for the person living it (Feldman, 2004: 9)<sup>2</sup>. Moreover, it also makes sense to speak of a person's welfare at certain points of their life, and the welfare effects of certain individualised events (Dorsey, 2013: 152). Philosophy of welfare seeks to distil this concept down to its most basic features to explain how and to what degree events affect a person's welfare. In Mark Lukas' words, 'the main task of giving an account of well-being comes down to specifying how we ought to assign values to a life and its parts' (Lukas, 2010: 3). When considering welfare, it is important to note that welfare theories *in and of themselves* do not make claims about how welfare should be integrated into our wider normative framework. There are multiple values that we care about and which are often taken to have normative significance when assessing the merits of someone's life and welfare is simply one among these (Feldman, 2004: 8; Griffin, 1986: 23; Sumner, 1996: 20). Some of these values may be in opposition to each other; for example, it may be possible for a person to live a profoundly unethical life while maintaining high welfare (Lukas, 2010: 2). Indeed, we do intuitively seem to face choices between actions that enhance our welfare, and actions that contribute to living a more ethical life. The decisions we make about how much of our income we will give to charity, and how much we spend on things we desire for ourselves may be one such choice. Moreover, some of these values may be interrelated. It may be, as Simon Keller claims, that the ethical value of our life non-instrumentally contributes to the welfare value of our life (Keller, 2004: 31). Aristotle, on some readings of him at least, seems to go further than Keller by claiming that welfare and ethical values are essentially reducible to one another. On this interpretation of Aristotle, the correct prudential choice of action is also always the correct ethical choice of action (Barnes & Kenny, 2004: 1097b13-17; Sumner, 1996: 79)<sup>3</sup>. Therefore, on this account, the value that we assign to the welfare of a person's life always correlates with the ethical value that we assign to that person's life. Nevertheless, despite the plurality of values that can be used to assess someone's life, and despite the possible interrelatedness of some of these values, it seems to make sense that *one* of those values is to do with how well a person's life is going for them. It is certainly a concept that exists within our pre-theoretical intuitions and is widely employed by us to assess and guide actions. Therefore, an investigation into the principles that determine what non-instrumentally effects welfare seems warranted.

### Rival candidate theories of welfare

A plethora of different theories have been proposed as candidates to track the welfare value of lives. Notably, Parfit makes a tripartite distinction between hedonistic theories, desire theories and objective list theories (Parfit, 1984: 493). Hedonistic theories claim that only pleasure non-

<sup>2</sup> Sometimes the term 'wellbeing' is used synonymously with welfare in the literature. I have opted to regularise my terminology to welfare primarily because that is the term that the more recent literature has coalesced around. Moreover, for me, the term wellbeing carries connotations of subjective mental states in way that the term welfare does not. If wellbeing does carry these connotations, then we should avoid it for it seems to subtly prejudge the question of whether welfare/wellbeing is a mental state or not.

<sup>3</sup> I use the term prudential to mean what self-interest dictates. The concept of self-interest tracks a person's own welfare.

instrumentally enhances a person's welfare and that only pain non-instrumentally detracts from their welfare. Desire theories claim that it is only the fulfilment of desires that non-instrumentally enhances a person's welfare and only the frustration of desires that non-instrumentally detracts from their welfare<sup>4</sup>. Objective list theories claim that it is only the attainment of certain things on a specified list that non-instrumentally enhances a person's welfare and only the absence of those things on the list that non-instrumentally detracts from their welfare. For example, an Aristotelean account would qualify as an objective list theory on this schema, for it specifies a list of virtues, the possession of which are components of a person's welfare (Barnes & Kenny, 2004: 1098a13-16). Parfit's tripartite distinction is not a descriptively comprehensive methodology for distinguishing between the different welfare theories on offer. Firstly, it assigns a whole host of vastly different welfare theories to a single residual category of 'objective list' theories (Woodard, 2013: 790). More importantly, from the perspective of this paper, it fails track some crucial differences between welfare theories, such as whether a theory is a 'mental state' or a 'non-mental state' theory (Woodard, 2013: 792). However, for the purposes of this discussion, further elaboration of the distinctions and divisions between different welfare theories is unnecessary. It is enough to acknowledge that there is a polyphony of different theories vying to be candidates for welfare. Consequently, it appears that we need some way of evaluating the relative strength of rival welfare theories.

There is no consensus on how we ought to assess the criteria that ought to be used for welfare theory evaluation. Indeed, there is little systematic engagement in the literature on what criteria ought to be used to assess the plausibility of rival welfare theories. Plausible appeals to intuition abound and claims that theories ought to meet one requirement or another are plentiful. Yet there is a distinct lack of attempts to systematise a criterion for welfare theory evaluation. Such an attempt would be beyond the scope of this thesis. Nevertheless, an examination of some of the motivations that guide people to adopt one theory over its rivals is useful to understand why some desire theories do appeal to many people, and why unrestricted non-mental state actualist desire theories may be better placed than rivals to meet some of these motivations. There are several standards that have emerged within the literature that seek to make comparisons about the relative strength of differing welfare theories. I will briefly survey four standards that *may* affect whether we are or ought to be inclined to adopt one theory over its rivals. These are: extensional adequacy, normative adequacy, descriptive adequacy, and conceptual frugality. By no means is this list intended to be exhaustive, and I make no claims about the relative weight that ought to be assigned to each of these metrics of evaluating the viability of rival welfare theories. I simply offer them as a way of understanding and categorising some of the arguments that are often appealed to in order to support one theory over another. It is beyond the scope of this investigation to claim that unrestricted non-mental state actualist desire theories fare better on these metrics than rival theories. This thesis does not seek to show that this theory is the correct theory of welfare, or even that there is such a thing as a correct theory of welfare. Indeed, we need not be normative realists about welfare in order to find value in establishing shared conceptual understandings about what welfare could be, or to find value in systematising our own intuitions about welfare. Therefore, it is worth briefly outlining why some people hold the view that it is the fulfilment and frustrations of desires that determine a person's welfare, and why the unrestricted, actualist and non-mental state versions of the desire theory may

---

<sup>4</sup> The exception to this definition is antifrustrationism, which claims that the fulfilment of desires does not non-instrumentally enhance welfare. Instead, on this view, desire fulfilment simply alleviates the non-instrumental disvalue that comes from desire frustration and anticipation (Fehige, 1998: 508). If we classify antifrustrationism as a desire theory, then we must drop the requirement that desire fulfilments non-instrumentally contribute to a person's welfare. For simplicity's sake I shall refer to desire theories as if fulfilments do non-instrumentally contribute to welfare and set aside the case of antifrustrationism.

contain features that fare well on these metrics. This discussion is intended to show that unrestricted non-mental state actualist desire theories fare well on some of these standards of evaluation, which in turn may therefore generate reasons to adopt this type of theory. Chapters Two, Three and Four are all concerned with preserving the extensional adequacy of these theories, as the criticisms from defective desires, Dead Sea apples, and intrinsically quirky desires all appeal to intuitions that may undermine their extensional adequacy.

One approach to evaluating welfare theories is to assess their *extensional adequacy*. The extensional adequacy of a theory is determined by whether it produces conclusions that align satisfactorily with our pre-theoretical intuitions about what outcomes we expect a theory of welfare to label as non-instrumentally welfare enhancing and what outcomes we expect a theory of welfare to label as non-instrumentally welfare detracting. A theory of welfare is extensionally adequate when its judgements about the types of things that effect welfare generally align with our pre-theoretical intuitions, or at least do not clash with them severely enough to undermine the theory (Enoch, 2005: 766). That is not to say that we should not be prepared for surprises: it may be that some things that we pre-theoretically intuitively believe to be beneficial to welfare instead turn out to have negative consequences for welfare. However, in general, we may expect stronger candidate welfare theories to more-or-less align with most of the evaluative judgements made by our pre-theoretical intuitions. The term extensional inadequacy captures the idea that the implications of a theory can be so intuitively inadequate, repugnant or bizarre as to undermine the viability of the theory itself. As such, extensional inadequacy is a threshold concept, for it requires that a certain threshold of intuitive implausibility is identified in the conclusions of a theory before the theory is dismissed for being extensionally inadequate. Conversely, the term extensional adequacy describes a phenomenon that comes in degrees. A theory can be more or less extensionally adequate depending upon the extent to which its conclusions cohere with our pre-theoretical intuitions. Some writers have suggested that finding a welfare theory that aligns closely with the bulk of our pre-theoretical intuitions about what types of things determine welfare ought to be a large consideration when considering the viability of a welfare theory (Fehige & Wessels, 1998: xxviii). Regardless of the relative weight we assign extensional adequacy in our criteria for assessing welfare theories, it strikes me as intuitively plausible that it ought to play some role in evaluating rival theories. Many writers who argue against different theories of welfare do so from the standpoint of finding extensionally inadequate conclusions entailed by it. Moreover, many defences of welfare theories involve attempts to defend their extensional adequacy. This paper continues in that tradition and is primarily concerned with rebuking the claims that examples of defective desires, Dead Sea apples, and intrinsically quirky desires undermine the extensional adequacy of unrestricted non-mental state actualist desire theories.

L. W. Sumner outlines a further two ways of assessing rival theories of welfare (Sumner, 1996: 8). The first is to do with the normative adequacy of a theory. Normative adequacy is determined by what extent the welfare theory under consideration fits into our wider normative framework of values. An example of a concern about the normative adequacy of a welfare theory may be that it is expected to generate relatively intuitive assessments when combined with our ethical theory. If a welfare theory generates profoundly counterintuitive implications when combined with our ethical theory, and we are willing to accept that ethical theory has priority over or precedes welfare theory, then we may find that our welfare theory is normatively inadequate and in need of revision. Concerns about normative adequacy pick up on the observation that welfare theories have implications beyond simply generating prudential considerations, and the intuition that the effects that these theories have on other normative judgements ought to be considered when assessing their viability. However, Sumner correctly points out that this approach by itself is not enough – for there may be several welfare theories that cohere equally well with our wider normative framework, and some of those

theories are likely to be deeply intuitively implausible. Therefore, a second method of adjudication, descriptive adequacy, ought to play at least some role in the determination of the viability of different welfare theories (Sumner, 1996: 10).

Descriptive adequacy is based on the extent to which a welfare theory captures our pre-theoretical intuitions about what *principles* ought to determine welfare. A theory lacking in descriptive adequacy is likely to be extensionally inadequate, although the two concepts are subtly different. Descriptive adequacy examines whether the principles of a welfare theory cohere adequately with our pre-theoretical intuitions about what sorts of principles ought to determine welfare; extensional adequacy is to do with whether the *implications* of a theory align sufficiently with our pre-theoretical intuitions about what sorts of implications a theory ought to label as affecting welfare. A common procedural method evoked to identify and defend a welfare theory involves appealing to the process of reflective equilibrium (Feldman, 2004: 206). This process involves oscillating between adapting pre-theoretical principles about welfare and pre-theoretical beliefs about the sorts of things that are taken to determine welfare until a point is reached when principles and beliefs are aligned. One way of understanding reflective equilibrium is to claim that it involves an appeal to both the descriptive and extensional adequacy of a theory in order to generate a welfare criterion. A fourth way of assessing the viability of theories of welfare is to analyse their conceptual frugality. A theory of welfare may be superior to its rivals if it utilises a less complex set of interrelated principles to generate conclusions about the welfare effects of events. This is simply because by utilising fewer principles, there are fewer components that an objector can find fault with. Properly speaking, conceptual frugality is a component of descriptive adequacy for many people. However, it is worth mentioning this criterion separately, as it is one that many people appeal to when assessing a welfare theory's descriptive adequacy. I have outlined these conceptual approaches to evaluating the merits of welfare theories at the start of this thesis as often these considerations are referred to in the discussion that follows.

We can briefly apply some of these evaluative metrics to desire theories to outline why some people find these theories convincing. Desire theories claim that it is the fulfilment and frustration of our desires that are the sole determinants of welfare. John Stuart Mill claimed that, upon inspection, all our desires appear to aim at the pursuit of pleasure and the avoidance of pain, whether that be for ourselves or for other people (Mill, 1993: 39). This led him to adopt a hedonistic theory of welfare. Hedonistic welfare theories claim that pleasure and pain are the sole features that determine welfare. If Mill is correct, then hedonism may be a descriptively and extensionally adequate theory of welfare, as through introspection people may be able to assent to his position that analysis of the aims of our desires supports hedonism. Mill's position can be interpreted as defending the descriptive adequacy of hedonism through appeal to claims about desires and their aims. This argument places the role of desires as lending credibility and structural support for his argument for hedonism. In this way, the intuition that desires play a role in welfare provides a measure of descriptive adequacy for hedonism. Mill appears to be drawing support from the philosophical position of what has become known as internalism. Internalism about welfare claims that, 'something cannot be good for a particular individual unless it can motivate her' (Rosati, 1996: 300). The idea that the things that are good for our welfare must somehow be related to our desires, or a subset of them, has widespread pre-theoretical appeal across many writers. Hedonistic theories of welfare meet the criterion of internalism, as we are often motivated by pleasure and the avoidance of pain. Mill's argument appeals to this intuition that a descriptively adequate welfare theory must adhere to internalism, and makes the additional claim that hedonism is the only theory that can meet this because all desires point towards hedonistic concerns anyway. Yet, as Richard Brandt points out, 'Mill was not even remotely correct when he said that 'desiring a thing and finding it pleasant' are 'two modes of naming the same psychological fact'. He would have been closer had he said that desiring a thing and its conception

being motivating are the same thing' (Brandt, 1979: 43). Indeed, the counterexamples to Mill's position are replete. For a start, Carson points out that there seem to be people who seek their own suffering (Carson, 2000: 19). More generally, it seems obvious to state that people desire all sorts of things for all sorts of different reasons – some of which do not point towards the pursuit of pleasure and the avoidance of pain (Griffin, 1986: 9).

It seems difficult to extrapolate any singular feature behind the aims of all our desires, except perhaps one: the fact that all desires aim at fulfilment. Therefore, if we are to draw reasons for preferring one welfare theory to its rivals from the contents of our desires, then a good place to start may be with the observation that all desires, across all persons, aim at fulfilment and the avoidance of frustration. From this observation, it appears that Mill's appeal to the underlying aims of desires is more supportive of a desire theory than a hedonistic theory. While both hedonism and desire theories abide by internalism, desire theories seem better placed than hedonism to capitalise on internalism, as they claim that it is desires, rather than one of their numerous aims, that determine welfare. Moreover, desire theories in their basic forms are as conceptually frugal as hedonism in their reliance upon only two principles to find fault with. The simplicity of desire theories can be considered a virtue. This is especially so of non-mental state, unrestricted and actualist desire accounts, which do not include caveats restricting which types of desires determine welfare. Consequently, we have reasons for thinking that desire theories may fare well on descriptive adequacy and conceptual frugality scales of assessment. Finally, Sumner has pointed to the normative adequacy of desire theories in sitting well with the dominant political philosophy of liberalism (Sumner, 1996: 123). Desire theories have similarities with liberalism as a political philosophy because they enable people to shape their own conditions for welfare generation in a way that rival theories do not always do. None of these reasons may be sufficient to accept the claim that it is the fulfilment and frustration of desires that determines welfare. Nevertheless, between them, these considerations provide a brief survey some of the concerns motivating the adoption of desire theories.

### **The basic unit of welfare: Mental state and non-mental state desire theories**

Desire theories claim that it is the fulfilment and frustration of desires that are the sole non-instrumental determinants of a person's welfare<sup>5</sup>. The basic unit of welfare for the desire theory emerges from the relationship between desires for states of affairs to obtain, and, depending on the type of account, either our subjective perceptions about whether those states of affairs have obtained, or, alternatively the objective fact of whether those states of affairs have obtained<sup>6</sup>. According to *non-mental state desire theories*, welfare is determined by the relationship between a person's desires and the objective state of the world. If the objective state of the world aligns with the state of affairs desired, then welfare is affected positively; and if the desire is misaligned with the objective state of the world, then welfare is affected negatively. In Mark Lukas' words, non-mental state desire theories claim that, 'the basic building blocks of well-being are states of affairs that supervene in part on mental states (desires) and in part on the state of the world (objects of desire)' (Lukas, 2010: 4). Consequently, non-mental state desire theories are labelled 'non-mental state' theories because while they do typically require the presence of a mental state to be a component of

---

<sup>5</sup> These accounts are sometimes called desire-satisfactionist accounts.

<sup>6</sup> The term 'state of affairs' is being employed here in a broad manner that can cover desires for other mental states, objective states of the world, or phenomenological experiences.

welfare generation, it is not solely mental state components that determine welfare<sup>7</sup>. To label a theory a non-mental state theory requires that welfare is determined, *at least in part*, by components that are not mental states. The distinction between non-mental state and mental state desire theories of welfare is based upon the different role that the state of the world plays in these two theories<sup>8</sup>.

According to *mental state desire theories*, welfare is produced by the relationship between a person's dispositional state of desiring and the *subjective perception* that they experience about whether their desire has been fulfilled. The positive welfare component is termed subjective desire satisfaction and is produced by possession of a subjective perception that a desire has been met (Heathwood, 2006: 548). Whereas the negative welfare component, subjective desire frustration, emerges when the subjective perception that a desire has not been met is present. Mental state desire theories conform to what Griffin labels 'The Experience Requirement'. This requirement stipulates that we must in some way be aware of the satisfaction/frustration of our desires in order for them to affect our welfare (Griffin, 1986: 13). Many people's intuitions claim that The Experience Requirement is necessary for descriptive adequacy. Their intuitions do not point to anything outside of mental states that could determine welfare and are untroubled by the problems, such as the experience machine, that will be examined shortly. However, for other people, The Experience Requirement is adopted largely in response to problems of extensional inadequacy that non-mental state theories seem to face. The later chapters of this paper seek to show ways of avoiding these extensionally inadequate conclusions without the abandonment of the non-mental state desire theory.

While the mental state desire theory conforms with The Experience Requirement. The non-mental state desire theory is a type of 'correspondence theory'. Correspondence theories claim that, 'how well things go for a person is determined by (i) the person's attitudes towards states of affairs or propositions (desiring them, believing them, taking pleasure in them), and (ii) whether those states of affairs are true' (Bradley, 2007: 47). Part of the attraction of correspondence theories is that they are able to discount pleasures or subjective desire satisfactions generated from false beliefs in a way that most mental state theories lack the resources to do<sup>9</sup>. This allows for correspondence theories to claim

---

<sup>7</sup> I say typically here because according to some idealisation non-mental state accounts of welfare, we need not actually hold a mental state of desire in order for our welfare to be affected by objective properties of the world. All that is required is that we *would* hold the desire under some specified set of ideal circumstances.

<sup>8</sup> These theories fit into the broader categories of being state-of-mind theories and state-of-the-world theories respectively. State-of-mind theories claim that welfare is determined solely by mental states. State-of-the-world theories claim that welfare is determined at least partly by objective properties of the world (Sumner, 1996: 82).

<sup>9</sup> A counterexample of a self-described mental state theory that does attempt to do this is Fred Feldman's 'veridical intrinsic attitudinal hedonism', which is itself a correspondence theory (Feldman, 2002: 616). Veridical intrinsic attitudinal hedonism claims that in order for pleasure to lead to welfare it must be based upon true beliefs. Feldman claims that this is a mental state theory because every instance of welfare generation is dependent upon the mental state of pleasure; it just that every mental state of pleasure by itself does not guarantee welfare generation. Wendy Donner has made a similar argument in her book *The Liberal Self* (1991), where she argues that the genesis of pleasures can contribute to their welfare value (Donner, 1991: 71). Where Donner and Feldman's veridical theories differ from typical hedonistic theories is that every mental state of pleasure *by itself* does not cause welfare generation. It is contestable about whether such theories qualify as forms of mental state hedonism or not. According to Carson, one of the four features of hedonistic theories of value is that, 'all pleasant experiences are non-instrumentally good' (Carson, 2000: 12). In this way, Carson's formulation does not allow for hedonists to base welfare upon state of the world considerations such as whether the pleasure is based upon true belief. The issues of whether Feldman and Donner's theories qualify as mental state theories, and whether correspondence theories can ever be mental state theories, I shall set aside. I set these issues aside because Feldman and Donner's theories only allow for the first benefit of correspondence theories, which is that they enable us to discount welfare based upon false

that cases of radical deception or misperception do have deleterious effects upon welfare. A prime example of the type of experience that we may want to discount as not enhancing welfare is illustrated by Robert Nozick's 'experience machine' thought experiment. In this scenario, Nozick asks us to envisage a life in which, unbeknownst to us, all our experiences, including our interpersonal relationships, are simulated by a machine that ensures that our life is pleasurable one (Nozick, 1974: 42-3). Such a simulated life is likely to rank high in welfare on mental state theories of welfare, for the machine is designed to produce positive mental states<sup>10</sup>. Yet despite this, many people hold strong intuitions that such a life would not involve as much welfare as had those experiences been authentic and not simulated. Correspondence theories are able to track intuitions that the authenticity of experiences affects the welfare generated them by claiming that, in cases such as Nozick's experience machine, our desires for authentic experiences and interpersonal relationships are not actually being fulfilled and therefore positive welfare is not being generated. Griffin captures this intuition in support of correspondence theories in the following way:

'My truly having close and authentic personal relations is not the kind of thing that can enter my experience; all that can enter is what is common to both my truly having such relations and my merely believing that I do. And this seems to distort the nature of these values. If I want to accomplish something with my life, it is not that I want to have a sense of accomplishment' (Griffin, 1986: 19).

Another attraction of correspondence theories is that they can claim that desires that we hold for states of affairs to be fulfilled do generate welfare, even in cases where their fulfilment or frustration does not generate any mental states. For example, we may care deeply that the Cambridge botanical gardens are preserved throughout our lifetime and think that their preservation affects our welfare irrespective of whether we come to gain beliefs about whether they have been preserved. We can imagine a scenario where we are permanently cut off from knowledge about whether the botanical gardens have been preserved, and yet still maintain a strong desire that they are preserved. Correspondence theories can allow for the outcome of the botanical gardens' fate to affect our welfare irrespective of the beliefs that we gain about their fate. Consequently, according to correspondence theories, whether the gardens are preserved or destroyed can indeed affect our welfare even in the absence of us gaining beliefs about whether our desire has been fulfilled or frustrated. At least in some cases, people do seem to value the objective state of the world itself as being a component of their welfare, and not simply the beliefs that they have about the state of the world. In Griffin's words, 'we do seem to desire things other than states of mind, even independently of the states of mind they produce' (Griffin, 1986: 9). Therefore, correspondence theories have two distinct advantages over most mental state theories: they are able to discount welfare generated through deception or misperception, while also being able to claim that welfare is generated by desires that we hold where fulfilment or frustration does not affect mental states. The drawbacks of correspondence theories come in the form of a unique set of challenges that they face in meeting the criteria of extensional adequacy. The question of how we achieve this is explored throughout the rest of this thesis.

---

beliefs and pleasures. These theories do not have the resources available to enable the second benefit of that correspondence theories, which is discussed below.

<sup>10</sup> While Nozick's example involves the machine simulating pleasurable experiences, his example of pleasure is substitutable for any other mental state that we may want to value, and the thought experiment will illustrate the same limitations to that theory as it does with one based on pleasure. With the exception of Feldman and Donner's veridical theories, if we take these positions to be mental state theories, then these are issues replicable across all mental state theories, including mental state desire theories.

Before continuing, it is useful to first briefly outline what desire is, and what constitutes a desire fulfilment. James Griffin describes the phenomenon of desire fulfilment as emerging from the following conditions: 'A desire is 'fulfilled' in the sense in which a clause in a contract is fulfilled: namely, what was agreed (desired) comes about' (Griffin, 1986: 14). This definition allows for desires to specify intricate details that can only be realised under a very narrow set of conditions, such as meeting Atticus at the Cambridge botanical gardens at 5:30pm on Tuesday; while also allowing for desires to be constructed as broad and realisable under a wide set of diverse circumstances, such as the desire for pleasurable phenomenological experiences generally. In this way, the contract analogy can account for the different types of desires that we experience. In addition to the contract analogy, we can also state that, 'desires are propositional attitudes' (Heathwood, 2005: 489). On this description, desires are mental states that make descriptive claims about states of affairs and involve a disposition in favour of or against those states of affairs occurring. They are fulfilled when those states of affairs are objectively brought about and frustrated when those states of affairs fail to be objectively brought about. According to desire theories, the amount of welfare generated from a fulfilled desire tracks the intensity of the desire felt – so that those desires that are experienced more strongly generate more welfare when fulfilled. Moreover, those desires experienced more intensely also contribute more negative welfare when frustrated than less intense desires.

## Conclusion

The theory that I am interested in is the unrestricted non-mental state actualist desire theory of welfare. This chapter has set the scene for why someone may hold such a theory. Before I consider the problem of defective desires, it is first worth briefly explaining the 'actualist' and 'unrestricted' parts of the unrestricted non-mental state actualist desire theory. The theory is actualist if it claims that it is the fulfilment of actual desires that determine welfare. And the theory is unrestricted if it claims that the fulfilment of all types of desires count towards determining welfare. In Mark Lukas' words, the unrestricted theory, 'places no restrictions on which desires are relevant to well-being. It says that *all* desire satisfactions are good for us and all frustrations are bad, no matter what the relevant desires happen to be about' (Lukas, 2010: 4). The alternatives to these positions are idealisation and restricted desire theories. I will examine what these positions are and why someone may hold these positions in the following chapter. My thesis is not that idealisation or restricted desire theories are incorrect or unviable in general, rather it simply claims that one need not adopt these positions in response to the problems of defective desires, Dead Sea apples or intrinsically quirky desires.

## Chapter 2: The problem of defective desires

### Introduction

In this chapter I consider the problem of defective desires that has been raised against unrestricted non-mental state actualist desire theories. Defective desires, alongside other types of desires deemed problematic for the extensional adequacy of unrestricted non-mental state actualist desire theories, have motivated many writers to adopt idealisation or restricted desire accounts. Arguments that claim that there are desire fulfilments that appear not to enhance welfare seek to undermine the *extensional adequacy* of unrestricted non-mental state actualist desire theories. These arguments claim that the implications of these theories are so intuitively inadequate, repugnant or bizarre that we ought to reject the theories outright. The identification of these sorts of desires has led writers to the conclusion that, ‘some of our [actual] responses are clearly bad, not reason supported, or not in accord with what really is of value, or with what is really good for us. Idealization is called for in order to save even just a possibility of extensional adequacy for response-dependence views’ (Enoch, 2006: 766). The concerns about problematic desires that are most prominently voiced against unrestricted non-mental state actualist desire theories fall under the rubric of arguments from ‘defective desires’ (Heathwood, 2005), ‘Dead Sea apples’ (Lauinger, 2016), and intrinsically ‘quirky desires’ (Bruckner, 2016). My later chapters examine Dead Sea apples and intrinsically quirky desires. This chapter is concerned with examining defective desires, as well as providing an outline of how restricted and idealisation accounts have been used to respond to them. I argue that defective desires need not undermine the extensional adequacy of unrestricted non-mental state actualist desire theories because Chris Heathwood’s ‘all-things-considered bad for welfare’ argument can account for them without appeal to idealisation or restriction.

### Three possible responses to the problem of defective desires

Many writers have argued that unrestricted non-mental state actualist desire theories are not viable because of the existence of, what has been termed in the literature, ‘defective desires’<sup>11</sup>. The problem of defective desires is said to undermine the extensional adequacy of unrestricted non-mental state actualist desire theories and thereby rule them out as viable candidates for occupying the welfare criterion. Desires can be defective for a number of reasons, but all defective desires pose the same basic problem for non-mental state actualist desire theories, which is presented by Heathwood as arising from the fact that, ‘sometimes, it is bad for a person to get what he wants’ (Heathwood, 2005: 487). Posed as a question for unrestricted non-mental state actualist desire theories, Heathwood’s observation is that: if desire fulfilment is what non-instrumentally enhances welfare, then how do we account for instances of desire fulfilment that are in fact detrimental to welfare? It is worth refining this problem provisionally to more precisely capture our intuitions about what conditions make a

---

<sup>11</sup> While this problem has been levelled against unrestricted non-mental state *idealisation* desire theories as well, the problem of defective desires is more threatening to unrestricted non-mental state *actualist* desire theories. Indeed, as has been noted, idealisation is often evoked as a response to the problem of defective desires because such accounts are said to have more available resources at their disposal to answer it. Because of this, and because the central argument of this thesis is to outline and defend unrestricted non-mental state actualist desire theories, I shall focus my attention on defending the actualist position from the argument from defective desires. However, it is worth noting that idealisation in and of itself does not necessarily solve the problem of defective desires. As Bruckner puts it, ‘It is not clear that such [defective] desires would not also survive exposure to the facts and unerring reasoning’ (Bruckner, 2016: 12).

desire defective. I propose the following: *defective desires are those desires for things that are bad for a person's welfare when fulfilled, but which are nevertheless taken by the unrestricted non-mental state actualist desire theory to non-instrumentally enhance a person's welfare if fulfilled*<sup>12</sup>. If defective desires exist, then on face value they pose a challenge to unrestricted non-mental state actualist desire theories. This is because, according to unrestricted non-mental state actualist desire theories, welfare is non-instrumentally enhanced through the fulfilment of our actual desires – yet, according to the defective desire problem, some desires are bad for a person's welfare when fulfilled. Griffin forcefully summarises the defective desire case against unrestricted non-mental state actualist desire theories in this way, 'Notoriously, we mistake our own interests. It is depressingly common that when even some of our strongest and most central desires are fulfilled, we are no better, even worse off' (Griffin, 1986: 10). For Griffin, it is this point that means that, 'the objection to the actual-desire account is overwhelming' (Griffin, 1986: 10). Indeed, Griffin's claim that we often misidentify our interests and subsequently desire those things that are not in our own interest to desire is one that seems to have considerable intuitive force. Therefore, any credible desire theory will need to account for this intuition in the form of answers to the problem of defective desires.

There are three popular positions that a defender of non-mental state desire theories can take when confronted with the problem of defective desires. Each of these positions can also be used as a response to the related problems of Dead Sea apples and intrinsically quirky desires. The first position is to opt for an idealisation over an actualist desire account. Once we abandon the claim that actual desire fulfilments and frustrations constitute the welfare criterion, then we can argue that the idealisation process would necessarily disqualify all defective desires from affecting welfare. Consequently, defective desires would not be claimed to non-instrumentally enhance welfare when fulfilled because they are not desires that would survive idealisation. The idealisation approach to answering the problem of defective desires involves claiming that we could not hold these types of desires if we occupied an 'ideal observer' position (Bruckner, 2016: 4). The sort of features that an ideal observer would possess are contested and vary between rival idealisation accounts. However, commonly these accounts stipulate the hypothetical desires of a version of ourselves that possesses all, or all the relevant, information deemed necessary to inform our desires, and is free of cognitive mistakes when processing that information and forming desires (Carson, 2000: 223). Many of the more influential idealisation accounts, such as that of Richard Brandt, also stipulate that the person must fully appreciate the implications of the desire, 'by repeatedly representing it, in an ideally vivid way, and at an appropriate time' (Brandt, 1979: 113). This is supposed to make the person undergoing idealisation fully appreciative of the affective implications of the desire's fulfilment as well as achieving a sufficient cognitive appreciation. According to many of these views, defective desires are a symptom of irrationality and incomplete knowledge – the removal of which through idealisation would remove them as problems for unrestricted non-mental state desire theories of welfare. The ability to take this

---

<sup>12</sup> Crucially, Heathwood's initial definition does not account for the possibility that not all desires that are bad for our welfare when fulfilled are likely to be labelled defective by unrestricted non-mental state actualist desire theorists. For example, we may for instance choose to pursue ethical considerations that require great personal sacrifice of our own personal welfare, and few people would call the desire that motivates that decision necessarily defective. Unless one is an egoist, or one holds the view that ethical and welfare concerns are extensionally equivalent or indeed reducible to one another – as Aristotle seems to have done (Sumner, 1996: 79) –, then we must define defective desires in such a way that is not merely all desires where fulfilment detracts from welfare. Therefore, my formulation of defective desires refines Heathwood's definition to encompass only those desires that it appears that non-mental state actualist desire theories seem forced to mistakenly classify as enhancing welfare when fulfilled. If readers are unhappy with the reformed definition, then they are free to substitute it with Heathwood's original. The substantive points that follow are not dependent upon any precise formation.

line of argument against the problem of defective desires would seem to be one of the primary motivations for adopting an idealisation desire theory (Griffin, 1986: 21). Indeed, without the added benefits of increased extensional adequacy, it is difficult to see why someone would convolute their welfare theory with the idealisation account. Idealisation makes the desire theory significantly less conceptually frugal.

Idealisation desire theories of welfare tend to be based on idealisation theories of rationality, which claim that we only have reasons to do that which we would be motivated to do from an ideal observer position. Idealisation desire theories of welfare go beyond idealisation theories of rationality by making the additional claim that one's welfare is determined only by desires based on good reasons. While there is division on how the idealisation theories ought to be constructed, Hubin points out that these processes tend to emphasise the importance of true belief in forming welfare-determining and reason-giving desires, 'There is at least near unanimity that false beliefs need to be purged – typically supplanted with true ones' (Hubin, 1996: 31). Therefore, the ideal observer position tends to be one from which the formation of true beliefs is consistent or constant. Idealisation accounts of welfare draw support from the claim that idealised desires still conform with the internalist requirement that our welfare is related to us by being present in our subjective desires (Rosati, 1996: 298). The account claims that these desires must then undergo the objective process of idealisation to work out exactly which desires determine our welfare<sup>13</sup>. I will not consider the merits of idealisation in being able to preserve the extensional adequacy of unrestricted non-mental state desire theories convincingly. It is enough to outline that this approach exists and is commonly motivated by problems of defective desires. However, it is worth noting that even if these theories are plausible and justifiable, it remains unclear whether idealisation would successfully remove all desires that are claimed to be defective (Bruckner, 2016: 12).

The second approach to solving the problem of defective desires involves adopting a 'restricted' non-mental state actualist desire theory. Restricted non-mental state actualist desire theories claim that, while it is the fulfilment and frustration of actual desires that determines a person's welfare, only a certain subset of actual desires count towards determining a person's welfare when fulfilled or frustrated. According to restricted non-mental state actualist desire theories, everything that counts towards determining a person's welfare is constituted by the fulfilment and frustration of that person's actual desires. However, not all of a person's actual desires count towards

---

<sup>13</sup> The literature on desire theories is replete with references to 'Full Information Theories' or 'Full Information Accounts' (Bruckner, 2016: 11; Carson, 2000: 224; Loeb, 1995; Rosati, 1995). These accounts make the claim that what constitutes the welfare criterion is the fulfilment and frustration of the desires that an individual would hold had they had full information. Moreover, there is also frequent reference in the literature on this topic to 'Ideal Observer' theories of determining axiological value (Brandt, 1955; Carson, 2000: 204; Enoch, 2005: 759). It is worth clarifying the relationship between these two theories and desire idealisation accounts. Full information accounts claim that the ideal observer for determining personal welfare is a version of ourselves with full information. Whereas, ideal observer accounts can specify more characteristics than simply full information, such as requiring cognitive changes in an individual (Brandt, 1979: 247). If an ideal observer account specifies a version of ourselves as being the genesis of the desires that determine our welfare, then it is a desire idealisation account. If the ideal observer account specifies an individual who is not a version of ourselves – such as a deity for example – as being the arbiter of the desires that determine our welfare, then it is *not* a desire idealisation account. This is because desire idealisation accounts are versions of ideal observer theories that maintain that welfare originates in subjective desires and not in external phenomenon. According to these definitions, all full information accounts of welfare are versions of desire idealisation accounts, and all desire idealisation accounts are versions of ideal observer accounts. Therefore, the term desire idealisation is used in this paper to refer to full information accounts *and* those ideal observer accounts that stipulate that it is idealised versions of our own desires that are determinative of our own welfare.

determining that person's welfare when fulfilled or frustrated. We can contrast restricted accounts with idealisation accounts. Idealisation accounts claim that it is the fulfilment and frustration of a set of hypothetical desires that an ideal observer version of ourselves would have which determines our welfare; whereas restricted accounts claim that the fulfilment and frustration of a subset of our actual desires determines our welfare. The differences between these two approaches are subtle but significant, and non-mental state desire theories can be both restricted and idealisation theories. This difference is not always clearly marked in the literature and often these approaches are elided. However, they do represent distinct and alternative ways of conceptualising welfare<sup>14</sup>. A common version of the restricted desire theory claims that only desires about our own lives count towards determining our welfare (Carson, 2000: 75)<sup>15</sup>. How we work out exactly which desires are about our own lives and which are not is deeply contested terrain (Portmore, 2007: 27; Sobel, 1997: 506). Restricted desire theories can also be used to help combat the problem of defective desires. We can apply this model to the case of defective desires by claiming that the fulfilment and frustration of defective desires does not belong to the subset of actual desires that determines an individual's welfare. A restricted actualist desire theory could claim that defective desires, Dead Sea apples and intrinsically quirky desires do not belong to the subset that determines welfare. Consequently, defective desire fulfilment and frustration need not contribute to welfare on the restricted desire account. Therefore, these theories can claim that the problem of defective desires can be navigated while preserving the non-mental state actualist desire theory framework.

The third approach, which is the approach that I take, is to argue that defective desires, Dead Sea apples and intrinsically quirky desires simply do not exist in way that is problematic for non-mental state actualist desire theories to accommodate. In relation to defective desires, I claim that defective desires can be explained as being examples of all-things-considered bad for welfare desire fulfilments, and that this claim is not inconsistent with the observation that they do nevertheless generate some countervailing positive welfare in virtue of their fulfilment. Consequently, while these defective desires do exist, they do not exist in a way that causes problems for the extensional adequacy of the unrestricted non-mental state actualist desire theory. Therefore, we do not need to make claims about the necessity of restricting the types of desires that determine welfare to a subset of our wider desires or make appeals to idealisation accounts in order to solve the problem posed by defective desires. To make sense of the claim that defective desires are not a problem for unrestricted non-mental state actualist desire theories we need to explore more thoroughly what people mean when they say that a desire is defective. My subsequent discussion about defective desires draws heavily on observations by Chris Heathwood, who has made a series of comprehensive arguments in his paper 'The Problem of Defective Desires' (2005) against the claim that defective desires undermine non-mental state actualist desire accounts<sup>16</sup>.

---

<sup>14</sup> Tom Carson's book *Value and the Good Life* (2000) contains a good example of a discussion that does make this distinction clear.

<sup>15</sup> Parfit terms restricted desire theories that count only those desires about our own lives 'Success Theories' (Parfit, 1984: 494). Success theories are only one possible way of restricting the desires that determine welfare.

<sup>16</sup> This is interesting given that Heathwood himself rejects non-mental state desire accounts primarily because of what he terms the problem of 'remote desires' (Heathwood, 2005: 500), and instead adopts a mental state desire theory / attitudinal hedonist theory, which he claims are extensionally equivalent (Heathwood, 2006). Nevertheless, Heathwood provides an important analysis and rebuttal of defective desire arguments against unrestricted non-mental state actualist desire theories.

### All-things-considered bad for welfare desire fulfilments

Heathwood outlines a list of defective desires that he divides into the following types: 'ill-informed desires, irrational desires, base desires, poorly cultivated desires, pointless desires, artificially aroused desires, and the desire to be badly off' (Heathwood, 2005: 487). These in turn he subdivides into three broad categories: '(i) arguments from 'all-things-considered defective' desires, (ii) arguments from 'intrinsically defective' desires, and (iii) a third category to be named later' (Heathwood, 2005: 490)<sup>17</sup>. I shall analyse some of Heathwood's arguments, which can provide the basis for some convincing rebuffs to the argument from defective desires. However, before I begin considering defective desires, I want to suggest a reworking of the categories that Heathwood labels defective. Defective desires are those desires that intuitively seem to make our life go worse for us when fulfilled. However, it seems too strong a statement from Heathwood to claim that all the desires that he lists are examples of defective desires. Instead, I want to suggest that pointless desires, poorly cultivated desires, and artificially aroused desires seem to fit better into the category of intrinsically quirky desires. I also propose to add base desires to this category, for while there is a strong intuitive case to be made that these desires are defective – in that fulfilment intuitively detracts from welfare –, the sort of argument that we can employ to explain them bears more similarity to intrinsically quirky desires. These desires all fit into the category that Heathwood labels arguments from 'intrinsically defective' desires.

The categorisation of a desire as quirky involves making a weaker claim than its categorisation as defective. Quirky desires are those where fulfilment is not *necessarily* bad for welfare but rather fulfilment is simply not good for welfare. In this way, quirky desires are an umbrella term that encompass defective desires alongside other problematic desires for the extensional adequacy of unrestricted non-mental state actualist desire theories. Bruckner writes that, 'A quirky desire is one that is difficult to understand or appears downright inscrutable, extremely strange, or unusual or maximally idiosyncratic' (Bruckner, 2016: 2). Quirky desires pose a similar problem for unrestricted non-mental state actualist desire accounts to defective desires. This is because unrestricted non-mental state actualist desire theories claim that the fulfilment of desires ought to be non-instrumentally beneficial to welfare, and not simply be not bad for welfare. I shall save consideration of pointless desires, poorly cultivated desires, artificially aroused desires, and base desires until Chapter Four, which specifically deals with those 'intrinsically' quirky desires that are not considered in Chapters Two and Three. This chapter will confine itself to considering ill-informed desires and irrational desires as prime examples of defective desires. Heathwood claims that these desires can be explained as being 'all-things-considered defective'. I think this is true of many ill-informed and irrational desires, but there are some troubling counterexamples that require additional explanation if we are to preserve the extensional adequacy of unrestricted non-mental state actualist desire theories. These counterexamples come in the form of Dead Sea apples and will be examined in Chapter Three.

Before progressing, let us recall our revised definition of what it means for a desire to be defective:

---

<sup>17</sup> I do not examine this third category because it involves a paradox that can emerge when a person desires to be badly off. This paradox seeks to undermine the conceptual coherency rather than the extensional adequacy of unrestricted non-mental state idealisation desire theories. Ben Bradley's paper 'A Paradox for Some Theories of Welfare' (2007) contains the most recent comprehensive discussion of this paradox that I am aware of.

*Defective desires are those desires for things that are bad for a person's welfare when fulfilled, but which are nevertheless taken by the unrestricted non-mental state actualist desire theory to non-instrumentally enhance a person's welfare if fulfilled.*

What is striking about this definition of defective desires is that upon inspection it is unclear why desires that are defective in this way undermine the unrestricted non-mental state actualist desire theory. The above problem trades on the implicit assumption that being bad for welfare and non-instrumentally enhancing welfare are mutually exclusive conditions, which when combined generate a paradoxical outcome that undermines the viability of unrestricted non-mental state actualist desire theories. This is the same implicit assumption embedded in Heathwood's formulation of the problem of defective desires for unrestricted non-mental state actualist desire theories that, 'sometimes, it is bad for a person to get what he wants' (Heathwood, 2005: 487). It would follow then that a defender of an unrestricted non-mental state actualist desire theory would have to either reject the overwhelmingly intuitive claim that desire fulfilments can in some cases be bad for a person's welfare or reject a central premise of the unrestricted non-mental state actualist desire theory: That the fulfilment of actual desires is what non-instrumentally enhances welfare.

And yet, there is no reason why these two conditions of non-instrumentally enhancing and being bad for welfare need be mutually exclusive. If we treat 'bad for' and 'non-instrumentally enhancing' as constituent parts of the welfare effects of a desire fulfilment and not as claims about the aggregate or overall effect of a desire fulfilment, then it is perfectly possible for a desire fulfilment to be both bad for our welfare in some ways and non-instrumentally enhancing of it in others. Therefore, unrestricted non-mental state actualist desire theories can claim that an instance of desire fulfilment is *overall* bad for our welfare, while also maintaining that it *in some respects* non-instrumentally enhances a person's welfare. Making this claim allows for the preservation of the premise that desire fulfilment always non-instrumentally enhances welfare *to an extent* and adds to it the caveat that desire fulfilment is not always an *overall* good for welfare. This is necessary to explain our intuition that some desire fulfilments are bad for welfare. Heathwood claims that many desires that may seem intuitively defective can be explained in a way that is coherent with the non-mental state actualist desire theory as being *all-things-considered* bad for our welfare when fulfilled<sup>18</sup>. The fulfilment of these desires is all-things-considered bad for our welfare because, while some positive welfare is generated non-instrumentally by their fulfilment, this gain is outweighed by the instrumental negative effects on welfare of their fulfilment (Heathwood, 2005: 492). These instrumental negative welfare effects are generated by the frustration of other desires that are thwarted as a consequence of fulfilling an all-things-considered bad for welfare desire (Heathwood, 2005: 493). Accordingly, the benefits of desire fulfilment are outweighed by the harms of desire frustration in cases of defective desire fulfilment. In Carson's words, '[fulfilment of] ill-informed desires often frustrates *other desires* that are of greater importance to the agent' (Carson, 2000: 73).

---

<sup>18</sup> Sumner foreshadows Heathwood's longer treatment of this in his book *Welfare, Happiness and Ethics* (1996), where he writes that, in cases where getting what we want is bad for us, 'satisfying the desire made me *to that extent* better off, but it also frustrated other, more important desires, so that on balance I ended up worse off' (Sumner, 1996: 131). This idea is also present in Carson's book *Value and the Good Life* (2000), where he talks of desire fulfilment only being good for welfare 'other things equal' (Carson, 2000: 72) and not always good for welfare on aggregate. Going back even further, Henry Sidgwick gives a brief treatment of this issue when he writes of some desire fulfilments that, 'in such cases the desired result is accompanied or followed by other effects which when they come excite aversion stronger than the desire for the desired effect' (Sidgwick, 1907: 110). However, Heathwood's consideration of this argument is the most extensive and systematic treatment of it in the literature, so my theoretical engagement is largely based on his formulation of the argument.

This explanation highlights that all that a defender of an unrestricted non-mental state actualist desire theory needs to claim for their theory to remain viable is that the fulfilment of defective desires does generate a degree of positive welfare. These theories are not committed to the deeply implausible claim that desire fulfilment is always all-things-considered good for a person's welfare.

The all-things-considered bad for welfare account also provides an intuitively plausible explanation of what happens when we act in accordance with ethical and other-regarding values over prudence. In these cases, it would seem that we are fulfilling a desire and thus, according to unrestricted non-mental state actualist desire theories, generating some welfare by acting in accordance with the pursuit of these values. At the same time, acting in accordance with ethical and other-regarding values often frustrates other self-directed desires that may, if they had been fulfilled, have generated more welfare and which would therefore be the prudentially prescribed course of action. The all-things-considered bad for welfare account can represent the intuition that often ethical actions do require a degree of self-sacrifice of a person's welfare, while at the same time representing the intuition that often people do generate some welfare for themselves when acting ethically. Theories of welfare do not claim to have a monopoly on what constitutes rational action and are therefore compatible with the claim that other values may clash with and override our motivation to pursue our own welfare at least some of the time. Heathwood puts it like this, 'If theories of welfare have any implications about behaviour, they imply only what we *prudentially ought* to do. But since sometimes I prudentially ought to do what I morally ought not to do, and sometimes morally ought to do what I prudentially ought not to do, theories of welfare on their own imply nothing about what we morally ought to do' (Heathwood, 2005: 496).

Utilising the all-things-considered bad for welfare position to explain unprudential desire fulfilments accounts for our intuitions about what happens in cases of ethical and other-regarding desire fulfilments, as well as providing an argument against the extensional inadequacy said to be generated by defective desire fulfilment. Without this explanation the non-mental state actualist desire theorist is left in the uncomfortable position of claiming that ethical action is as prudential as self-directed action, as both involve the fulfilment of desires. To avoid this conclusion, it is sometimes claimed that adopting a restricted desire theory is a necessary conceptual commitment of anyone trying to defend the non-mental state actualist desire theory (Portmore, 2007: 27). One of the reasons that people give for the necessity of adopting the restricted desire theory is to explain ethical action in a way that does not mean it is reduced to prudential action. According to this version of the restricted desire theory, only self-directed desires determine welfare. This allows for ethical and other-regarding desires to not be considered welfare-enhancing when fulfilled in the same way as self-directed desires are, providing one way of accounting for these unprudential desire fulfilments. However, the all-things-considered bad for welfare explanation is able to explain unprudential desire fulfilments for unrestricted non-mental state actualist desire theories in a conceptually frugal way and without recourse to restrictions on the sorts of the desires that have standing when determining welfare. While the all-things-considered bad for welfare account of defective desire fulfilment would, if viable, answer the problem of defective desires, it remains to be seen whether this account can be convincingly applied to the whole range of desires that have been labelled defective.

There is a litany of examples that abound within the literature on defective desires that the all-things-considered bad for welfare argument can be employed to answer. To use Heathwood's schema, we can begin with the case of ill-informed desires, which are a subset of defective desires. Ill-informed desires are those desires that we hold on the basis of incomplete or false information. Clearly our desires are determined partly by our beliefs, and so if those beliefs are false or not adequately representative of all the relevant information, then they may generate desires that are ill-informed

and defective. An example of an ill-informed desire that Heathwood provides is the desire of a person to quench their thirst with water that unbeknownst to them contains poison (Heathwood, 2005: 491). They hold this desire on the basis of incomplete information and its fulfilment is clearly intuitively bad for them. Heathwood explains this intuition by referencing the all-things-considered badness of the desire for welfare. He points out that drinking the poison water will inevitably frustrate other stronger desires – the desire to not be in pain for example, as well as any desires that will be thwarted during the recovery period (providing that the poison is not fatal, and that the person is lucky enough to have a recovery period). Nevertheless, Heathwood contends that a small amount of positive welfare is generated from drinking the poison water, it is simply *outweighed*. In order to support his claim that a small amount of positive welfare is generated by consuming the poison water, he asks us to imagine that the poison takes a few minutes to take effect. In this time, we can imagine two hypothetical scenarios: One in which the person drinks the water and the world ends before the poison can take effect, and one in which the person does not drink the water and the world ends just before the poison would have taken effect. Heathwood claims that in the scenario in which the person did consume the poison water, their welfare is marginally higher than in the scenario in which they did not consume the poison water. This is because the scenario in which they did consume the poison water has all the benefits of the scenario in which they did not, plus the added bonuses of the experience of quenched thirst and the absence of the frustration of thirst (Heathwood, 2005: 492). This thought-experiment is designed to show that the all-things-considered bad for welfare account fits our intuitions about what happens when we satisfy an ill-informed desire.

This example of Heathwood's may seem trivial, and some people's intuitions may reject the idea that the fulfilment of desires which have a comparatively small effect on a person's overall life, such as quenching one's thirst for a few minutes, can play any role in determining how well that life has gone for them. To represent this view within the non-mental state actualist desire framework, we could adopt a version of the restricted desire theory, which claims that only the fulfilment and frustration of a certain subset of our actual desires determines our welfare – such as, for example, a requirement that desires have a certain level of intensity or duration in order to qualify as welfare determining when fulfilled or frustrated. An example of a restricted desire theory that takes roughly this position is a goal theory. These theories claim that only those desires that are elevated to the status of actual goals are determinative of welfare (Keller, 2004: 28)<sup>19</sup>. However, to hold such a view does not eliminate the problem of ill-informed desires or the validity of Heathwood's response to it – it simply confines the problem and its response to affecting only those desires that fall into the required subset of our desires that are considered determinative of welfare.

---

<sup>19</sup> This statement is somewhat complicated by the fact that Keller defends in his paper, what he terms, an 'unrestricted goal theory', which claims that the achievement or frustration of any goal, 'even when [the] goals are crazy, self-destructive, irrational or immoral' contributes to welfare (Keller, 2004: 27). Nevertheless, I think there is a good case to be made that goals are themselves a subset of our wider desires. In which case, Keller's theory is a restricted desire theory, but an unrestricted goal theory. Indeed, Griffin and Carson both suggest that it is 'global desires' that determine welfare (Carson, 2000: 73; Griffin, 1986: 13). These 'global desires' are those desires that we have about our life as a whole and seem similar to the notion of goals. Goal theories themselves face the problem of defective goals, which is analogous to the problem of defective desires, with writers on both problems often drawing on the same examples (Keller, 2004: 30; Bruckner, 2016: 3). Goal theories / global desire theories are sometimes adopted to represent the intuition that minute desire fulfilments do not affect welfare; they have also been said to be in a stronger position to meet the challenges posed by the problem of changing desires, especially when combined with idealisation accounts (Carson, 2000: 85-86).

For those who take this approach of restricting the desires that count to include only those deemed sufficiently important, Heathwood's example can be easily reformulated so that something more significant is at stake than simply a few minutes of quenched thirst. Take the following example:

you are offered an experimental pharmaceutical drug that you believe has the reliable benefit of enabling you function optimally without sleep in perpetuity. Unbeknownst to you the drug will cause sudden death after precisely one year of first consumption. You desire to take the drug because of the incomplete information that you possess about the harms that would come with consumption, combined with your correct beliefs about how optimal functioning without sleep could assist you in the fulfilment of other desires and therefore in the enhancement of your welfare. Had you known the full consequences of consuming the drug you would certainly reject it as being very bad for your welfare.

As with the poison water case, we can appeal to the all-things-considered badness of this desire for welfare to explain why intuitively we reject the idea that its fulfilment would benefit your welfare. Nevertheless, we can plausibly claim that a degree of positive welfare is generated from the consumption of the drug, it is simply outweighed. This plausibility stems from the observations that without the need for sleep you could fulfil your desires to produce more things of artistic and intellectual merit, to have more leisure time, and to dedicate more time to building deep and meaningful personal relationships – all things that seem to our intuitions to be unproblematically beneficial desires to satisfy, and the achievement of some of which would surely constitute a goal. We can apply Heathwood's hypothetical case of the two scenarios again to highlight this point: one in which you take the drug, experience the benefits, and the world ends before the fatal consequences of its consumption can take place, and the other in which you do not take the drug, do not experience these benefits, and the world ends just before the fatal consequences of its consumption would have taken place. When applying our intuitions to which scenario contains more welfare, it should be clear that the first scenario would undoubtedly contain significantly higher welfare than the second. This is because the first scenario contains all the benefits of the second in addition to the extra benefits that come with the fulfilment of other desires, while containing no additional harms. If the claim that desire fulfilment can be all-things-considered bad for welfare but nevertheless contains a degree of positive welfare works in this scenario, then opponents of the theory ought to supply reasons outlining why it ought not apply to other situations of ill-informed desire satisfaction. The difference between our two cases appears to be simply one of scale, not one of type, and therefore, so far, it looks like ill-informed desires can be reliably explained by non-mental state actualist desire accounts as being all-things-considered bad for welfare.

The all-things-considered bad for welfare explanation of what happens when we fulfil an ill-informed desire can also be applied relatively straightforwardly to some other cases of defective desire fulfilment. In Heathwood's words, what unites all these cases is that, 'they are bad not in themselves but for what they lead to, or what they prevent' (Heathwood, 2005: 491). He applies this same logic to irrational desires, where he claims that, 'were the desire to be satisfied, a less favourable balance of desire satisfaction over frustration would result than would result if the desire were not satisfied' (Heathwood, 2005: 493). To claim that a desire is irrational presupposes a criterion of rationality for desires. Whether desires are the sorts of things that it makes sense to call rational or irrational is fraught debate, as many writers in the Humean tradition take desire to be fundamentally arational in character, while others in different traditions argue that desires themselves can be rational or irrational (Hubin, 1991). However, there are some desires that we can label as straightforwardly irrational desires, such as for example the desire for both P and not-P to obtain simultaneously. In a case like this, it may be fair to say that the desire is irrational because it is both

logically unconceptualizable in a coherent way and subsequently logically unattainable<sup>20</sup>. While these desires are perhaps relatively uncontroversial to label irrational, their existence does not bear directly on the phenomenon of defective desires. This is because straightforwardly irrational desires are completely impossible to conceptualise in a coherent way, let alone fulfil, and therefore their fulfilment can never be bad for our welfare. The case of straightforwardly irrational desires is not what Heathwood has in mind when he discusses irrational desires. Instead, Heathwood's position is that an irrational desire is an informed yet unprudential self-directed desire<sup>21</sup>.

The example that Heathwood gives of irrational desires clarifies that he is discussing informed yet unprudential self-directed desires, 'Suppose I need to see a dentist for a procedure that will require drilling. And suppose this time lack of information isn't my problem: I know it is in my long-term interest to see the dentist. But I still want no part of it. Clearly, it would be bad for me not to go' (Heathwood, 2005: 487-488). In this case, we have the example of a person with full information who chooses an unprudential self-directed desire fulfilment for presumably no rational reason. Sidgwick offers a hypothesis to explain what happens psychologically when we can act upon an informed yet unprudential self-directed desire. Sidgwick suggests that in these cases 'bad effects are *fore-seen* but not *fore-felt*' (Sidgwick, 1907: 110). If we accept this, then it would seem that, in these cases, we have an intellectual appreciation that the desire's fulfilment will be bad for our welfare but that we fail to emotionally internalise that intellectual appreciation. Heathwood's position is that, other things being equal, it is irrational to act unprudentially when pursuing a self-directed desire fulfilment. In cases like Heathwood's dentist example, we have a prime example of someone failing to make full use of their beliefs and desires in order to choose prudential action, and instead opting for a harmful course of action to their welfare. I will not attempt to arbitrate on whether the pursuit of fulfilling informed yet unprudential desires is, other things being equal, irrational when it comes to our self-directed actions. The problem of 'irrational desires' in Heathwood's sense exists despite the terminology we use to describe it. Whether or not we deem the label of irrational appropriate, we still have an example of a seemingly defective desire that a non-mental state actualist desire theory will need to account for.

As with cases of ill-informed desires, many examples of irrational desires can easily be accommodated by the all-things-considered bad for welfare argument in order to explain how we can maintain that desire fulfilment enhances welfare in some respects while nevertheless detracting from welfare in other respects. The temporal gap in Heathwood's example is useful to illustrate this. By not going to the dentist now I appreciate the benefits of avoiding an unpleasant situation in the present

---

<sup>20</sup> Note that the example of the desire for P and not-P to obtain should not be confused with cases of *countervailing desires*. Countervailing desires emerge when we have two or more desires that weigh in on either side of us wanting/being aversive to a proposition being fulfilled. For example, I may have a desire to eat an éclair in order to experience some sensory pleasure and a countervailing desire not to eat the éclair in order to avoid the deleterious health consequences of excessive sugar consumption. This is a perfectly rational case of countervailing desires, where we do in some sense want both P and not-P to obtain. In these cases, we have two logical but opposing desires, for P and not-P. While only one can be fulfilled, both are independently conceptualizable and independently logically possible. Countervailing desires are present in everyone to some extent (Carson, 2000: 233). However, a straightforwardly irrational desire would be to want P and not-P to obtain as a single integrated desire. The fulfilment of this type of desire is simply logically unconceptualizable and subsequently logically unattainable, and therefore is a good candidate for the title of a straightforwardly irrational desire.

<sup>21</sup> The notion of an informed yet unprudential self-directed desire can be dismissed by an idealisation account as being impossible. Such accounts can claim that an informed yet unprudential self-directed desire is a logical impossibility. Either the desire is not informed, or it is not self-directed, or it is not unprudential; a desire could not be all three simultaneously, or so the argument goes.

in exchange for far worse consequences in the long term<sup>22</sup>. Yet as with our previous examples of ill-informed desires, we can contrast hypothetical scenarios to illustrate that the fulfilment of the irrational desire itself does generate some positive welfare. If the world were to end at the end of the day of my dentist appointment, then my life would have gone better for me for having avoided the unpleasant experience and instead doing something else. The fulfilment of the irrational desire itself then seems to benefit welfare, it is only the wider consequences that make it unprudential. Therefore, if a person wishes to pursue an irrational desire, then, while the desire fulfilment does generate some positive welfare, it also necessarily obstructs other desires that they hold or could cultivate, and which would, if fulfilled, have led to higher welfare generation than the irrational desire would have. At the risk of giving the case of irrational desires a perfunctory treatment, I shall not examine these types of cases further. If we accept the explanation of fulfilment being all-things-considered bad for welfare in the case of ill-informed desires, then irrational desires are analogous enough to also be accepted without further investigation. However, while this explanation does work well for many irrational and ill-informed desires, we do seem to face a set of counterexamples where it is much harder to claim that fulfilment generates any welfare whatsoever. This problem has been termed the problem of Dead Sea apples in the literature.

## Conclusion

This chapter began by considering the problem that many of our desires would intuitively appear not to enhance our welfare when fulfilled. I then outlined three possible responses to this observation: Adopting an idealisation theory, adopting a restricted desire theory, or attempting to explain this phenomenon in a way that does not undermine the unrestricted non-mental state actualist desire theory. I then turned my attention to defective desires and claimed that Heathwood's all-things-considered bad for welfare explanation can provide an adequate account of why defective desires do not undermine the unrestricted non-mental state actualist desire theory. I observed that this explanation is also necessary if we are to make sense of ethical and other-regarding desire fulfilments without resorting to a restricted desire account. I then focused on two particular types of defective desire fulfilments, those that are ill-informed and those that are irrational. The examples I considered of these types were found to be adequately explained by the all-things-considered bad for welfare account. I now move onto consider a set of counterexamples that the all-things-considered bad for welfare account finds more challenging to intuitively answer. These are found in the existence of what have been termed Dead Sea apples.

---

<sup>22</sup> I leave aside the issue of whether such temporal discounting is in itself irrational and for the sake of argument assume with Heathwood that it is. Although I think there are good reasons for thinking that temporal discounting is often unjustifiable on the social level (Cowen & Parfit, 1992).

## Chapter 3: The problem of Dead Sea apples

### Dead Sea Apples

The examples that Heathwood provides of defective desire fulfilments that can be answered by the all-things-considered bad for welfare argument involve cases of ill-informed and irrational desire fulfilments where there are some clear intuitively perceptible beneficial consequences of the desire fulfilment for welfare. Moreover, these beneficial consequences are built into the desire for the object itself and are not simply incidental or unforeseen consequences of the desire fulfilment. I benefit from drinking the poison water *because* it quenches my thirst, I benefit from consumption of the sleep conquering pharmaceutical drug *because* it enables me to fulfil other desires, and I benefit from avoiding the dentist *because* I avoid a painful experience in favour of less onerous activities. These benefits are of course all outweighed by the accompanying deleterious consequences for welfare, but their existence is nevertheless clearly intuitively perceptible. This intuitive perceptibility supports the unrestricted non-mental state actualist desire theory's premise that actual desire fulfilments enhance welfare. The temporal gap between the immediate benefits to welfare of desire fulfilment and the long-term harmful consequences in Heathwood's examples makes the all-things-considered bad for welfare argument eminently plausible when used to explain them. This is useful for illustrative purposes because it shows how the all-things-considered bad for welfare argument *can* work. However, objectors may claim when presented with this argument that the fulfilment of *all* ill-informed and irrational desires need not necessarily involve any such clear intuitively perceptible beneficial consequences. These objectors can point to compelling counterexamples of ill-informed and irrational desire fulfilments that do not involve a temporal lag between immediately clearly intuitively perceivable beneficial consequences for welfare and longer-term harmful consequences. Indeed, they can point to examples where desire fulfilments contain no clearly intuitively perceivable beneficial consequences for welfare at all. These counterexamples make the claim that the all-things-considered bad for welfare argument solves the problems of ill-informed and irrational desires for unrestricted non-mental state actualist desire theories much harder to defend.

It would be disingenuous of defenders of the unrestricted non-mental state actualist desire theory to point to cases where there are clear intuitively perceptible beneficial consequences of ill-informed and irrational desire fulfilments without considering these counterexamples. Consequently, all that Heathwood and I have succeeded in showing thus far is that there are such things as all-things-considered bad desire fulfilments for welfare, and that some cases of ill-informed and irrational desire fulfilments can be explained through appeal to the all-things-considered bad for welfare argument. We have not shown that *every* instance of desire fulfilment contains at least an amount, however small, of positive welfare generation, and which therefore makes every instance of ill-informed and irrational desire fulfilment at least all-things-considered bad for welfare and never completely devoid of any positive welfare generation. Therefore, we will need to provide additional reasons to support the claim that all ill-informed and irrational desires do indeed make some positive contribution to welfare when fulfilled. I think that this claim can be persuasively made, but I accept that the challenge to show this is greater than the examples of the poison water drinker, sleep conquering pharmaceutical drug taker, and dentist absconder suggest. In order to find the most compelling counterexamples to the all-things-considered bad for welfare explanation of ill-informed and irrational desires, we will have to consider cases of Dead Sea apple desire fulfilments.

The use of the term Dead Sea apples to describe desire fulfilments that do not benefit a person's welfare *in any way* at all appears to have its origins in Henry Sidgwick's book *The Methods of*

*Ethics* (1907). There he writes of some desire fulfilments that, 'It would still seem that what is desired at any time is, as such, merely apparent Good, which may not be found good when fruition comes, or at any rate not so good as it appeared. It may turn out a 'Dead Sea apple,' mere dust and ashes in the eating' (Sidgwick, 1907: 110). This rather poetic description of some desire fulfilments being completely devoid of any benefit to welfare and therefore being analogous to the eating of the mythical Dead Sea apples, said to turn to dust and ash in the mouth, captures the type of desire that is unexplainable by the all-things-considered bad for welfare explanation of ill-informed and irrational desire fulfilments<sup>23</sup>. If such Dead Sea apples do exist, then their existence poses a problem for unrestricted non-mental state actualist desire theories, which claim that *all* actual desire fulfilments enhance welfare at least by a small amount. The problem of Dead Sea apples led to Sidgwick suggesting the adoption of an idealisation theory for those who want to preserve a non-mental state desire theory (Sidgwick, 1907: 111)<sup>24</sup>. To date the most comprehensive discussion of Dead Sea apples in relation to desire theories of welfare that I am aware of is provided by William Lauinger. His paper 'Dead Sea Apples and Desire Fulfillment Welfare Theories' (2011) follows Sidgwick's observation in claiming that the existence of Dead Sea apples undermines unrestricted non-mental state actualist desire theories. Moreover, he goes further than Sidgwick and claims that the existence of Dead Sea apples also undermines non-mental state idealisation theories of welfare as well (Lauinger, 2011). Lauinger's concern is later echoed by Brucker, who points out that *even if* idealisation accounts of welfare were able to navigate other serious concerns people have raised about them, then they would still not necessarily remove the offending defective desires problem, of which Dead Sea apples are the most pertinent example (Bruckner, 2016: 12).

Heathwood gives a very brief consideration to the problem of Dead Sea apples in his paper 'The Problem of Defective Desires' (2015) but dismisses the problem quickly as unthreatening to those unrestricted non-mental state actualist desire theories that also adopt the concurrence view of welfare generation (Heathwood, 2015: 493). Perhaps because Heathwood himself does not endorse a non-mental state desire theory at all, he does not consider alternatives to adopting the concurrence view to answer the problem posed by Dead Sea apples. However, accepting the concurrence view has major ramifications on the sorts of things that determine welfare, and therefore it is a worthwhile investigation to see whether the adoption of such a view is truly *necessary* in order to preserve the extensional adequacy of the unrestricted non-mental state actualist desire theory. We may have independent reasons for accepting the concurrence view, but it remains to be seen whether its acceptance is the only viable response to the problem of Dead Sea apples for unrestricted non-mental state actualist desire theories. This is something that Heathwood fails to show, or indeed consider, in his paper. I will examine the concurrence view shortly, but to pre-empt that discussion slightly, it is not a position that I find convincing. More importantly, it is not a necessary theoretical commitment for those who wish to adopt an unrestricted non-mental state actualist desire theory. There are alternative and compelling solutions available to the problem of Dead Sea apples for unrestricted non-mental state actualist desire theories. Therefore, Heathwood's discussion warrants updating to consider how we can remain agnostic on the concurrence view and nevertheless preserve the extensional adequacy of the unrestricted non-mental state actualist desire theory.

I make two separate arguments in order to preserve the extensional adequacy of the unrestricted non-mental state actualist desire theory without appeal to the concurrence view. Firstly,

---

<sup>23</sup> According to some renditions of the Dead Sea apples myth, the apple turns to smoke or ashes when plucked from the tree, rather than dust and ashes when eaten (Lauinger, 2011: 324).

<sup>24</sup> Readers should note that the terminology I am using here did not exist when Sidgwick wrote. However, I am confident that these terms nevertheless are being employed accurately to capture the outline of his views.

I claim that many examples given of Dead Sea apples fail to describe instances of actual desire fulfilments at all. Therefore, these cases can be dismissed as unproblematic to the extensional adequacy of the unrestricted non-mental state actualist desire theory. Secondly, I claim that those examples that are put forward as Dead Sea apples, and that do describe actual desire fulfilments, are not actually Dead Sea apples. This second argument involves making the claim that there are good reasons to think that these cases do involve some degree of welfare enhancement, and thus fit into the mould of being all-things-considered bad for welfare and consequently unproblematic for unrestricted non-mental state actualist desire theories. If these arguments are successful, then we will have extinguished the problem of Dead Sea apples arising from ill-informed and irrational desire fulfilments for unrestricted non-mental state actualist desire theories. This is done by claiming that, despite some pre-theoretical intuitions to the contrary, Dead Sea apples simply do not exist. Indeed, the existence of Dead Sea apples is incompatible with unrestricted non-mental state actualist desire theories.

To illustrate the type of ill-informed desire fulfilment that is not so obviously all-things-considered bad for welfare, and would seem a strong candidate to qualify as a Dead Sea apple, we can construct the following example:

I desire to drink the coffee on the table in front of me. My desire is based on the belief that the coffee will taste pleasing and therefore I proceed to add sugar to the contents of the mug. I take a sip of the contents of the mug, only to experience an immediate and unmitigated disgust. My desire to drink the coffee was based upon the incorrect belief that I had been adding sugar, when I had in fact been mistakenly adding salt. I gain no pleasure from the experience and no feeling of subjective desire satisfaction is aroused within me.

The above example differs from the cases of the poison water drinker and that of the sleep conquering pharmaceutical drug taker because the case of the salted coffee does not contain any clear intuitively perceptible component of compensatory positive welfare generation in the same way as our other examples do. In the salted coffee case, the negative experience is instantaneous and not deferred, and this immediacy means that we cannot point to any obvious benefits of the desire's fulfilment through appeals to Heathwood's comparisons of two world ending scenarios. Recall that this thought experiment involves getting us to imagine two scenarios: One in which we experience the benefits without the harms, and one in which we do not experience either the benefits or the harms. In the example he gives of the poisoned water case, there is a temporal lag between the immediate benefits and the deferred harms. Because the benefits and harms are present at different temporal locations, it is easy to show that welfare is affected both positively and negatively in different measures by this desire's fulfilment. Yet this temporal lag is not present in the salted coffee case, and there are no clearly intuitively perceivable benefits for welfare that we can point to, making the all-things-considered bad for welfare argument intuitively difficult to apply. Indeed, an objector to the unrestricted non-mental state actualist desire theory can appeal to strong intuitions that the fulfilment of this desire is completely devoid of any positive welfare generation whatsoever. We have here a strong candidate for a Dead Sea apple in the case of the salted coffee, and one that, if left unanswered, seems to undermine a central premise of unrestricted non-mental state actualist desire theories: That welfare enhancement is present to an extent in all instances of desire fulfilment. It is this sort of Dead Sea apple desire fulfilment that poses the challenge that this chapter is concerned with. Without adopting a restricted or an idealisation version of the desire theory, the non-mental state actualist desire theorist has three alternative ways of responding to this challenge. The first approach is to accept the concurrence view. The second approach is to give an alternative account of why Dead Sea apple examples do not describe actual desire fulfilments. And the third approach is to

claim that Dead Sea apple examples do describe actual desire fulfilments, but that these fulfilments can be explained by the all-things-considered bad for welfare account and are therefore not actually Dead Sea apples. I think that our understanding of what happens when we encounter a Dead Sea apple example is best accounted for through a mixture of the second and third approaches. But firstly I will examine in more depth what the concurrence view claims and why we should be reluctant to appeal to it to solve this problem.

### **The concurrence view**

The concurrence view claims that a person must be desiring a state of affairs at the same time as the state of affairs occurs in order for welfare to be generated<sup>25</sup>. Heathwood defines concurrence views as stipulating that, 'A desire of mine is satisfied only if [I] get the thing while I still desire it, and continue to have the desire while I'm getting it. The theory above therefore does not take the duration of a desire to be as prudentially significant as some have taken it to be' (Heathwood, 2005: 490). It is the requirement that we continue to desire an experience when getting it that allows Heathwood to dismiss Dead Sea apples as not being problems for the unrestricted non-mental state actualist desire theory. Dead Sea apples are no longer problems for the unrestricted non-mental state actualist desire theory that adopts the concurrence view because under this view Dead Sea apple examples do not describe instances of actual desire fulfilments at all. Indeed, the problem of Dead Sea apples could be taken as a motivating reason to adopt the concurrence view as a way of preserving the extensional adequacy of the unrestricted non-mental state actualist desire theory. For Heathwood, his acceptance of the concurrence view means that compensatory welfare generation is built into the very definition of every instance of desire fulfilment. For if we were to lose the desire for a state of affairs' continuation at the moment that we perceive the desired state of affairs' occurrence, then we would not have achieved actual desire fulfilment at all on his definition. If the desire persists after fulfilment, even for a short time, then Heathwood can point to this as evidence of at least some positive welfare being generated from fulfilment. This evidence of some positive welfare being generated preserves the unrestricted non-mental state actualist desire theory's premise that desire fulfilments always enhance welfare to an extent, even if in some instances they are all-things-considered bad for welfare. Therefore, Heathwood's position navigates the problem posed by Dead Sea apples by claiming that these examples do not describe instances of actual desire fulfilment at all, let alone instances of ill-informed or irrational desire fulfilments devoid of any welfare generation.

However, adopting the concurrence view does come with its own unique set of limitations and theoretical commitments that may undermine some of the intuitive appeal of unrestricted non-mental state actualist desire theories. Firstly, it requires us to claim that the length of time that we hold a desire for does not non-instrumentally affect the welfare generated by its fulfilment. This means that prudentially we have no non-instrumental reasons to fulfil more long-standing desires than we do more recently acquired desires. To offer some amelioration of this concern, we can point to instrumental reasons to fulfil long-standing desires over newly acquired desires. For example, the persistence of long-standing desires is often evidence of their intensity and, according to unrestricted non-mental state actualist desire theories, the intensity of a desire does affect the welfare generated by its fulfilment. Moreover, we can claim that such intensity rarely exists in newly acquired desires,

---

<sup>25</sup> The theory also says that a person must have an aversion to a state of affairs at the same time that the state of affairs is experienced in order for negative welfare to be generated. For the sake of simplicity, I shall speak primarily of (positive) welfare generation and not of negative welfare generation and leave the reader to apply the same observations symmetrically to desire frustrations.

and that we may often be mistaken in our beliefs about our desires when we attribute high intensity to newly acquired desires. Therefore, prudence may consistently recommend the fulfilment of long-standing desires over newly acquired desires. Nevertheless, this explanation may not satisfy everyone's intuitions, for it still allows for some newly acquired desires of high intensity to displace long-standing life goals as being the prudential course of action. Committing to this view may be descriptively inadequate to some people's intuitions, which often seem to suggest that long-standing desires ought to count more to welfare when fulfilled for non-instrumental reasons than newly acquired desires<sup>26</sup>. That is not to say that we should presume that the fulfilment of long-standing desires matters non-instrumentally more than the fulfilment of newly acquired desires to welfare. Intuitions are divided by this issue, and it is by no means implausible to accept Heathwood's position that the duration of a desire does not have non-instrumental effects on the welfare generated by its fulfilment. The point of this discussion is simply to highlight that there is a debate to be had about this issue, but that adopting the concurrence view in response to the problem of Dead Sea apples necessarily commits one to the stance in this debate that the length of time that a desire is held does not non-instrumentally affect the welfare generated by its fulfilment.

A related issue for concurrence views occurs when we observe that often our present desires are 'future-directed'. Future-directed desires are those that stipulate concerns about what happens in the future. For example, my desire for my favoured sports team to win their respective league competition is a future-directed desire as it is something that I do not want to happen, and indeed could not happen, immediately. Whereas my desire to eat an éclair may be one that I desire to happen immediately and is therefore presently-directed. If a desire is future-directed, then it may be intuitively plausible to suggest that what happens in the future may matter to the welfare effects of the desire's fulfilment or frustration, even if the desire ceases to be experienced before its fulfilment or frustration occurs. We can call those desires that were previously held by a person but that no longer persist within them 'past desires'. To complicate matters, the fulfilment of some future-directed desires seems to matter to us in the present, independently of whether we continue to hold that desire into the future. Dorsey points out that the welfare effects of these types of future-directed desires intuitively seem not to be contingent on the desire's own persistence (Dorsey, 2013: 158). For example, an ardent vegetarian may have a desire in the present that in the future they continue to abide by vegetarian ethical standards, irrespectively of whether they continue to have a desire to abide by those ethical standards in the future. They may also have a meta-desire to continue to have the desire to abide by vegetarian ethical standards, but the actual fulfilment of their desire to abide by vegetarian ethical standards may matter to them in the present independently of whether that meta-desire is fulfilled. In these types of cases, the person experiencing a desire would claim that they presently desire for X to happen in the future, irrespectively of whether they continue to desire X to happen in the future. A prime example of such cases are desires held for events to happen after an individual's death, when the person experiencing the desire knows that they will not be experiencing the continuation of the desire at the time of its fulfilment or frustration. Given that such desires do exist almost universally across persons and given that they can often be felt with intensity, there may be intuitive force behind the claim that the unrestricted non-mental state actualist desire theory – a theory which draws part of its appeal for representing the range of people's desires – ought to allow for fulfilment and frustration of these types of desires to effect welfare. Conversely, the value that we place on other types of desires does appear to be completely contingent upon their own persistence.

---

<sup>26</sup> Indeed, the intuition that we ought to value long-standing desires in a different way to fleeting desires is one of the concerns motivating the adoption of a 'goal theory' (Keller, 2004), or 'global desire theory' (Carson, 2000: 73; Griffin, 1986: 13). These theories only count the fulfilment and frustration of goals or global desires as determining a person's welfare.

For example, when reflecting upon my desire to eat an éclair for breakfast tomorrow, I can observe that it is something that I presently desire to happen in the future only for as long as continue to desire it in the future. With these types of future-directed desires, we can understand of a person that they presently desire for X to happen in the future, only for as long as they continue to desire for X to happen in the future. In these cases, the value that is placed on the desire in the present is tied to the continuation of that desire into the future. Intuitively it seems that without the desire's persistence there seems to be no welfare generated from the fulfilment or frustration of my desire to eat the éclair in the future. The division between which desires are contingent upon their own persistence and which desires seem to involve propositions, at the time of desiring, that fulfilment is important *irrespective* of the desire's persistence, is unclear. Nevertheless, there are examples of desires that fall sharply into either category, as the cases of vegetarianism and the desire for the éclair show.

Consequently, the concurrence view entails that we do not have any reason to fulfil or give motivating weight to past desires that we no longer experience<sup>27</sup>. Yet, if someone holds a desire for the majority of their life, and it is the sort of desire where the value of fulfilment is not intuitively contingent upon the desire's own persistence, but ceases to experience that desire minutes before fulfilment, it is not intuitively obvious that the desire's fulfilment would not benefit their welfare *at all*. To say that past desires may influence someone's welfare is not to say that their fulfilment influences welfare equally to presently existing desires. The claim that the fulfilment and frustration of past desires matters to welfare is compatible with the claim that they matter less to welfare. Whether these desires ought to be treated differently in their welfare determining effects to presently existing desires is an issue beyond the scope of this thesis. Moreover, it is not by any means implausible to claim that the fulfilment and frustration of past desires does not influence welfare non-instrumentally at all. Nevertheless, it should be clear that there is a debate to be had here when it comes to constructing a theory of welfare representative of our intuitions. Adopting the concurrence view in order to answer the problem of Dead Sea apples closes down that debate and should therefore be avoided if possible at this stage of our investigation.

An additional theoretical commitment of the concurrence view is that they entail that present desires we have for past events also cannot affect welfare (Lukas, 2010: 18). Consequently, if I strongly

---

<sup>27</sup> Mark Vorobej defends the thesis that we do have reason to give weight to past desires that we no longer hold when considering the prudential course of action (Vorobej, 1998: 306). His position is that, 'relevant desires held during some period need not be satisfied *during that very period* in order for their satisfaction to be capable of positively affecting someone's overall welfare' (Vorobej, 1998: 310). Therefore, the fulfilment of past desires that someone no longer holds at the time of fulfilment can affect how well someone's life has gone for them overall on his view. However, Vorobej does not think that the fulfilment of a past desire at a later date makes a person's life go better at any one particular moment. If I fulfil a desire I no longer have, my welfare is not improved at the moment of fulfilment, nor during the period in which I held the desire, but my life has gone better overall according to Vorobej. Vorobej also uses the same argument to claim that posthumous desire fulfilment and frustration can affect an individual's overall welfare (Vorobej, 1998: 313). However, Dale Dorsey considers this to be a 'disastrous result' as it entails that we have cases of welfare generation without temporal location, making it impossible to measure welfare at different points of a person's life (Dorsey, 2013: 154). Instead, Dorsey makes an argument for giving weight to the welfare effects of past desires but departs from Vorobej by claiming that our welfare is improved at a specific moment. The moment that he claims welfare is improved at is during the time in which we experience the past desire, and not when the desire that is no longer held is fulfilled (Dorsey, 2013: 152). Therefore, according to this view, actions that we undertake in the future can affect how well our life was going for us in the past, as well as affecting how well our life goes for us overall. Dorsey holds this view partly because he thinks that we ought to give an analogous treatment to temporally distant desires as we do to spatially distant desires (Dorsey, 2013: 158). I mention these positions simply to highlight that there is a vibrant and ongoing debate about past desire fulfilment and posthumous benefits and harms.

desire for a temporally distant ancestor of mine to have had a virtuous or high welfare life, then the objective fact about whether they did lead such a life cannot affect my welfare on the concurrence view. Finally, adopting the concurrence view rules out the possibility of claiming that posthumous events can impact a person's welfare. Consequently, Dorsey points out that the unrestricted non-mental state actualist desire theory would be committed to an interesting asymmetry in its treatment of *spatial* distance in comparison to its treatment of *temporal* distance (Dorsey, 2013: 158). Recall that unrestricted non-mental state actualist desire theories allow for desire fulfilments that do not affect our mental states either directly or indirectly to nevertheless effect our welfare. Such instances Dorsey terms examples of spatial distance in desire fulfilments. Whereas, Griffin describes the fulfilment of spatially distant desires as being those that fail to meet 'The Experience Requirement' that we examined in Chapter One (Griffin, 1986: 13). Conversely, temporal distance is present when a desire is fulfilled or frustrated at any point after the person desiring ceases to hold the desire. Dorsey raises the interesting question of why a theory would be justified in treating these spatial distances in some desire fulfilments as nevertheless not disqualifying them from affecting welfare, while at the same time rejecting the idea that temporal distance in desire fulfilments can ever affect welfare (Dorsey, 2013: 158). If I hold a strong present desire for life to exist on Mars today, and unbeknownst to me that desire turns out to be fulfilled immediately, then the unrestricted non-mental state actualist desire theory claims that my welfare is enhanced. This is despite the fate that this desire fulfilment does not affect my mental states either directly or indirectly. But if I hold a strong present desire for life to exist on Mars *at some point*, and unbeknownst to me that desire turns out to be fulfilled in three hundred years' time, then the concurrence view combined with the unrestricted non-mental state actualist desire theory claims that my welfare is not enhanced. Intuitively it seems wrong that the unrestricted non-mental state actualist desire theory gives a different answer in these two cases, with only the first being granted the status of enhancing welfare. After all, both cases do not involve a person's mental states being affected either directly or indirectly, and both cases describe very similar types of desire being fulfilled. Therefore, it may seem intuitively strange that the temporal location of the desire's fulfilment can be used to disqualify one from affecting welfare but not the other. Dorsey claims that to treat these two different types of distance asymmetrically is inconsistent and bizarre, and therefore we ought to abandon either the concurrence view or the non-mental statism that allows for spatial distance. While people's intuitions diverge sharply on this issue, the theoretical commitment that temporal distance prohibits a desire from affecting welfare but spatial distance does not may be a counterintuitive implication of the concurrence view to some (Sumner, 1996: 127). Consequently, we can see that adopting the concurrence view necessarily entails the commitment to a whole range of conclusions that some advocates of the unrestricted non-mental state actualist desire theorist may want to avoid. Commitment to such a view *as a solution* to the problem of Dead Sea apples should therefore be avoided if possible to allow for a consideration of the merits of these conclusions on an individual basis.

I reject the concurrence view primarily because I find that it commits me to other conclusions that I find to be extensionally and descriptively inadequate and therefore want to avoid. Particularly, I find convincing Dorsey's observation that an asymmetrical treatment of spatial and temporal distance seems unwarranted. There has not been enough space in this discussion to comprehensively make the case that the fulfilment and frustration of past desires can in some cases affect welfare, or to explain in which cases they do affect welfare, or exactly *how and to what degree* they affect welfare. However, I hope to have shown the reader that we have good reasons to at least be hesitant in adopting the concurrence view as a solution to the problem of Dead Sea apples. At the very least, its adoption is by no means as obvious a solution to the problem of Dead Sea apples as Heathwood's brief discussion would have us believe. This is especially so if the problem of Dead Sea apples can be

solved within an unrestricted non-mental state actualist desire theory's framework without appeal to the concurrence view, as I believe that it can be. An alternative solution to the problem of Dead Sea apples for unrestricted non-mental state actualist desire theories can be found by drawing upon the same basic conclusion that concurrence views give of examples of Dead Sea apples, which is to claim that they simply do not exist. Indeed, there is something plausible about the concurrence view's conclusion that examples of Dead Sea apples fail to describe instances of actual desire fulfilments at all and therefore do not undermine unrestricted non-mental state actualist desire theories. What is undesirable about the concurrence view is the reasoning that it gives for claiming that Dead Sea apples do not exist, as this reasoning entails the adoption of other theoretical commitments that we may wish to avoid. There is an alternative position that unrestricted non-mental state actualist desire theories can take instead that involves different reasoning and therefore a different set of accompanying theoretical commitments. I find these additional theoretical commitments to be far more intuitively plausible in maintaining the extensional adequacy of unrestricted non-mental state actualist desire theories than those implied by the concurrence view.

### **The Hubin-inspired desire theory of welfare**

An alternative approach to solving the problem of Dead Sea apples for non-mental state actualist desire theories can be constructed from observations found in Donald Hubin's paper, 'Hypothetical Motivation' (1996). In this paper, Hubin develops and defends an account of reasons which claims that reasons are generated by our actual motivations<sup>28</sup>. Accounts of reasons are concerned with understanding what conditions give rise to reasons for actions. Hubin's account of reasons is developed as an alternative to the adoption of an idealisation account of reasons, which he finds to be inadequate and unnecessary (Hubin, 1996: 40). The adoption of an idealisation account over an actual motivation account of reasons is often driven by problems of 'defective motivations', a problem which mirrors much of our own discussion of defective desires. Idealisation accounts of reasons state that we have reason to do whatever we would be motivated to do in a set of idealised circumstances. The claim that reasons are generated by our actual motivations faces a whole range of counterexamples of seemingly defective motivations that would generate extensionally inadequate reasons for action. Consequently, we may be inclined to adopt an idealisation approach to understanding reasons for action because actual motivation theories seem committed to the claim that reasons are generated by defective motivations. Hubin formulates the problem of defective motivations in this way, 'The view that one has reason to do whatever one is moved to do is too crude. Sometimes our motivation would not hold up to a moment's reflection: appreciation of the cause of our motivation, its nature, or the real effects of acting in accordance with it may annihilate the motivation' (Hubin, 1996: 31). However, in response to the problem of defective motivations, Hubin defends the position that we can account for reasons by claiming that they arise from our own actual *intrinsic* motivations, rather than our actual motivations in general. This position he labels an 'actual intrinsic motivation' account of reasons (Hubin, 1996: 33). His actual intrinsic motivation account of reasons is a rival theory to idealisation accounts of reasons and involves a separate solution of the problem of defective motivations. This solution is based on the claim that actual intrinsic motivations

---

<sup>28</sup> I am using the term motivation instead of desire here because Hubin's account references motivations rather than desires. That said, I take the two to describe similar, if not identical, subjective states of having a favourable attitude towards a proposition, which in turn can lead us to act to bring about the proposition's occurrence. If this analogous treatment between motivation and desire is not what Hubin has in mind, then not too much hangs on it. I examine Hubin's account of reasons and motivation only to show how observations from it can be relevant to our consideration of desires and welfare.

can never be defective, and that when we experience a defective motivation, it is in fact an instrumental motivation. Since instrumental motivations do not generate reasons for action, their defectiveness does not lead the theory to make extensionally inadequate claims about reasons for action. Therefore, Hubin believes that he has solved the problem of defective motivations for actual intrinsic motivational accounts of reasons.

Hubin notes that his position involves a theoretical commitment to certain views about what constitutes an intrinsic motivation, 'This view depends on conceiving of motivation as having a roughly hierarchical structure: some motivation is dependent, in familiar ways, on other motivation together with beliefs about the world' (Hubin, 1996: 44). Nevertheless, the observation that some motivations are purely instrumental to the fulfilment of intrinsic motivations seems to be intuitively plausible. An example of a purely instrumental motivation can be constructed as follows: I am motivated to add the sugar to my coffee, only insofar as I believe that doing so will make the fulfilment of my intrinsic motivation to drink sweetened coffee more achievable. Hubin claims that it is only the fulfilment of these intrinsic motivations, in this case my motivation to drink the sweetened coffee, that generates reasons for action. Moreover, he argues that the instrumental motivations that we have are susceptible to being based on false beliefs in a way that intrinsic motivations are not. According to Hubin, it is this susceptibility to being based on false beliefs that leads to the existence of defective motivations for action. For example, to return to the salted coffee case, my motivation to add the white granular substance before me to my coffee was based on my false belief that this substance was sugar and not salt. Consequently, the instrumental motivation to add the white granular substance to the coffee was defective, but, on Hubin's account, did not generate a reason for action and therefore does not pose a problem to the actual intrinsic motivation account of reasons. Indeed, this example seems to fit Hubin's claim that only instrumental motivations can be defective. In Mendola's words, 'Donald Hubin holds that many intuitively incorrect desires are not incorrect *intrinsic* desires, but rather incorrect *instrumental desires*, rooted in false beliefs about what would satisfy intrinsic desires' (Mendola, 2009: 149). Hubin goes further than this characterisation of his work by Mendola to claim that:

'I believe that the same sort of response is appropriate for all of the "defects" that are supposed to be corrected by the idealizations demanded by hypothetical motivation theories. These can be understood as defects insofar as they are likely to interfere with the promotion of that which the agent is intrinsically motivated to bring about. But, if this is the only sense in which they are defects, the intrinsic motivation theory is not embarrassed by these defects, for it recommends actions in virtue of their connection with those states which we are *intrinsically* motivated to bring about, not in virtue of the particular "derived" motivation we may have as a result of some defect' (Hubin, 1996: 45).

Consequently, if Hubin is correct, then *all* defective motivations are explained as simply being defective instrumental motivations, which are motivations that do not by themselves generate reasons for action anyway. We can dismiss these defective instrumental motivations as simply being misplaced instrumental motivations that fail to track effective routes to the fulfilment of our intrinsic motivations. Therefore, the problem of defective motivations for the actual motivation account of reasons is ameliorated in a way that avoids appeal to idealisation. This is done by claiming that there simply are no defective intrinsic motivations.

Whether Hubin's actual intrinsic motivation argument works as an alternative to idealisation accounts of reasons is an issue beyond the scope of this paper. What is important to this paper is whether we can extrapolate observations from his discussion of reasons to our consideration of the problem of Dead Sea apples for unrestricted non-mental state actualist desire theories. Indeed, his

work defending an actual intrinsic motivation account of reasons from the problem of defective motivations can serve as the basis for an alternative response to the problem of Dead Sea apples for unrestricted non-mental state actualist desire theories. If we apply Hubin's observations directly to the issue of welfare, then we end up with a restricted non-mental state actualist desire theory that claims that Dead Sea apples are instances of instrumental desire fulfilments, and that instrumental desire fulfilments do not by themselves non-instrumentally generate welfare. Call this the 'Hubin-inspired desire theory of welfare'. What exactly constitutes a non-instrumental desire is unclear. However, according to the Hubin-inspired desire theory of welfare, the fulfilment of a non-instrumental desire can never be a Dead Sea apple. Consequently, our encounters with Dead Sea apples will give us some indication of the sorts of desires that are instrumental and those desires that are non-instrumental. For example, to return to the salted coffee case, in this scenario the Hubin-inspired desire theory of welfare claims that my desire to consume the substance in the mug was only an instrumental desire to achieve some further end. What exactly the non-instrumental desire motivating that instrumental desire was is unclear. For example, it could be the desire for pleasure in general or the specific desire for sweetened coffee that constitutes the non-instrumental desire motivating my consumption of substance in the mug in this case. Nevertheless, despite this ambiguity, there is something to be said for the idea that my desire to consume the contents of the mug before me was not a truly non-instrumental desire. If we accept this claim, then this argument entails that Dead Sea apples do not exist as a problem for non-mental state actualist desire theories.

Taking this position would be to commit to a restricted non-mental state actualist desire theory because it entails the modest restriction that only those desires that we hold non-instrumentally contribute to welfare when fulfilled or frustrated. This restriction is one that many writers take to be uncontroversial and necessary to preserve a non-mental state desire theory at all, whether idealised or actual. Indeed, many desire theories take it as a premise without a robust consideration of its viability (Carson, 2000: 72; Heathwood, 2005: 489; Hubin, 2003: 318). The reasons motivating the adoption of the restriction that only the fulfilment and frustration of non-instrumentally held desires determines a person's welfare tend not to be based on the claim that these desires are robust to the problems of defective desires or Dead Sea apples. Instead, this particular restriction is seen as necessary for other reasons. Mendola points out that the restriction that only the fulfilment and frustration of non-instrumentally held desires determine welfare seems necessary because, 'otherwise there is the possibility of unintuitive double-counting of satisfactions towards well-being' (Mendola, 2009: 149). This problem emerges because in the pursuit of non-instrumental desire X, a person may form a series of instrumental desires, call them A, B and C, in order to achieve the fulfilment of X. It would seem strange that a proliferation of instrumental desires in this case would contribute to a person's overall welfare, given that these desires are purely held on the basis of a belief that their fulfilment will help facilitate the achievement of the non-instrumentally desired outcome, X. Consequently, most non-mental state actualist desire theories tend to claim that only non-instrumental desires contribute to welfare when fulfilled or frustrated. While I am sympathetic to this restriction, I think that Hubin's position can be convincingly modified so that we can remain agnostic on the claim that only non-instrumentally held desire fulfilments and frustrations determine welfare. Therefore, we can preserve an unrestricted non-mental state actualist desire theory in response to the problem of Dead Sea apples. Moreover, the division between instrumental and non-instrumental desires is more challenging to make systematically than restricted desire accounts often assume. Consequently, avoiding appeal to this division may lead to a more conceptually frugal way to preserve the non-mental state actualist desire account in the face of the problem of Dead Sea apples.

Before we consider how we ought to modify the Hubin-inspired desire theory of welfare, we should first consider whether it really does adequately ameliorate the problem of Dead Sea apples for non-mental state actualist desire theories. I will not consider whether this account can answer all cases of defective desire fulfilments for non-mental state actualist theories of welfare, as we have already seen that the all-things-considered bad for welfare account can answer many examples of ill-informed and irrational desire fulfilments. All that needs to be considered is whether the Hubin-inspired theory of welfare can explain cases of Dead Sea apples through appeal to the claims that they are only instrumental desire fulfilments, and that instrumental desire fulfilments do not enhance welfare in and of themselves. How we assess the viability of these claims will be completely dependent on the criteria we outline for what constitutes a non-instrumental desire. While many non-mental state desire theories rely upon the distinction between instrumental and non-instrumental desires, it is hard to pin down exactly how we categorise actual desires as belonging to either type. Indeed, it seems possible to attribute an instrumental value to the fulfilment of almost any desire: I desire the coffee because it brings me energy to start my day, I desire energy at the start of my day to reply to my emails efficiently, I desire to reply to my emails efficiently to shine at my job, I desire to shine at my job in order to raise my prospects of promotion, I desire a promotion to raise my salary, etc. . Where exactly instrumental desires end, and non-instrumental desires begin on this list is unclear, and without the development of a theory of categorisation there will be reasonable intuitional disagreement about which desires ought to count as non-instrumental and which as instrumental. Indeed, this teleological analysis of reasons for action seems to have motivated Aristotle to claim that all reasons for action point toward the single non-instrumentally valuable end of achieving *eudaimonia* (Barnes & Kenny, 2014: 1094a20-25). Conversely, the teleological analysis of what ends our desires point towards led to John Stuart Mill claiming that all desires are ultimately instrumental desires in the pursuit of the single non-instrumental desire for pleasure, whether that be our own pleasure or other people's (Mill, 1993: 39). Both Aristotle and Mill use these similar arguments, based on reflections on the instrumental functions of common reasons and desires, to identify welfare with *eudaimonia* and hedonism respectively.

These approaches take us far away from the non-mental state actualist desire theory of welfare and into the territory of other types of welfare theories. It seems then that if we consider instrumental desire fulfilments and frustrations to not effect welfare, then we may end up committing to an alternative welfare theory to non-mental state actualist desire theories. One could respond at this point by claiming that just because almost all desires have instrumental value, that does not mean that they are not held for non-instrumental reasons as well. Indeed, the observation that something can be both instrumentally and non-instrumentally valuable for a person goes back to at least Aristotle (Barnes & Kenny, 2014: 1097a30-35). Using this observation, we can claim that it is enough for someone to desire something for partly non-instrumental reasons in order for the fulfilment or frustration of that desire to count towards determining their welfare, regardless of whether that desire is also held for instrumental reasons as well. Nevertheless, it is unclear how we arbitrate between those desires that are held purely instrumentally and those that contain an element of non-instrumental value. Consequently, we ought to be cautious when committing to the claim that only non-instrumental desires matter to determining welfare. In Mendola's words, 'It might be claimed that it is internal to the content of each desire itself whether something is wanted as a means or as an end, or both. But at least in the vast preponderance of cases this is not true in any way that I can introspect' (Mendola, 2009: 149).

While the problem of the proliferation of instrumental desires seeming to improve welfare on the unrestricted non-mental state actualist desire theory ought to concern us, the claim that only the fulfilment and frustration of non-instrumentally held desires determine a person's welfare is one that

also comes with its own set of problems. Moreover, unless we can clarify the division between instrumental and non-instrumentally held desires clearly, then the Hubin-inspired desire theory of welfare risks appearing to be an *ad hoc* solution to the problem of Dead Sea apples. Without giving a full account of what makes a desire instrumentally or non-instrumentally held, it seems unwarranted to dismiss cases of Dead Sea apples as simply belonging to the ill-defined category of instrumental desires and therefore not a problem for the extensional adequacy of our non-mental state actualist theory. This seems especially *ad hoc* if we take a liberal view of which desires count as non-instrumental, as the more desires that we acknowledge count towards determining welfare, the more likely we will find difficult counterexamples in the form of Dead Sea apples emerging in the form of non-instrumentally held desire fulfilments. To claim that these desire fulfilments are instrumental *because* they are Dead Sea apples seems unconvincing. Far better to show that Dead Sea apples all happen to be instrumental desire fulfilments, and that we have other reasons aside from solving the problem of Dead Sea apples to adopt the claim that instrumentally held desires do not determine welfare. Conversely, if we take a more conservative view of which desires are held non-instrumentally and thus count towards determining welfare when fulfilled or frustrated, and claim that most desires are instrumentally held, then we risk constructing a theory of welfare that seems very alienated from the vast majority of desires that motivate people's actions. Indeed, taking this position may lead us closer to an Aristotelean or hedonistic account of welfare than a non-mental state actualist desire theory. Therefore, I think it best to remain agnostic on whether we ought to adopt the restriction account that only non-instrumentally held desires determine welfare when fulfilled or frustrated.

### **The composite desire theory**

Fortunately, it is possible to salvage observations from the Hubin-inspired desire theory of welfare without committing to a restricted desire theory. This can be done by claiming that our actual desires, whether instrumental or non-instrumental, are more *fine-grained* than is often assumed to be the case. If we think of desire fulfilment as being analogous to contract fulfilment, in the way that Griffin does, then we can draw further inspiration from this comparison about what conditions must be present to give rise to desire fulfilment (Griffin, 1986: 14). On a standard explanation of desire fulfilment, desires tend to be conceived as very basic propositions that contain little content and few clauses required for their fulfilment. For example, in the salted coffee case, my action of consuming the contents of the mug that contains the salted coffee is simply conceived of as a desire to consume the contents of the mug (full stop.). Yet it is possible, and indeed reasonable, to claim that desires often track multiple objective conditions at once when constructing the conditions required for their fulfilment. For example, it seems possible to claim that when someone desires X, they desire X only insofar as features A, B, and C are constituent parts of it. The desire itself is therefore a proposition with multiple constituent parts; it is in this sense *fine-grained*. If features A, B and C fail to be constituent parts of the fulfilment of X, then the fulfilment of X by itself does not non-instrumentally enhance welfare on our theory because the desire in its totality has not been fulfilled. In this way, several clauses may go together to make a proposition that is dependent on the fulfilment of all its constituent clauses in order to count as a desire fulfilment. Call this view the 'composite desire theory'. The composite desire theory does not claim that all desires are fine-grained, but it does claim that desires can be fine-grained in this manner. Consequently, we can apply the composite desire theory to the problem of Dead Sea apples by claiming that Dead Sea apples are either completely or largely explained through appeal to the claim that they do not describe examples of actual desire fulfilments, but instead describe only fulfilments of constituent clauses of a wider proposition. By only fulfilling one of the clauses of a composite desire, Dead Sea apples generate the illusion of desire fulfilment to

our pre-theoretical intuitions, but upon reflection we can observe that these experiences do not describe actual desire fulfilments.

In order to illustrate how the composite desire theory could work in practice, we can apply its findings to an example of a Dead Sea apple, such as the case of the salted coffee. In this case, it seems reasonable to claim that my desire for the coffee can be constructed as a composite desire for a cup of coffee. For illustrative purposes, we can claim that this composite desire includes the intuitively plausible clauses stipulating that the coffee contains sugar and is devoid of substances that bring me immediate pain. We remain agnostic on the claim about whether this is a non-instrumental or instrumental desire, and instead claim that it is a composite desire. Consequently, when I drink the coffee with salt, we can explain the Dead Sea apple as arising from the fact that I have simply not got what I desired, as I have not experienced the fulfilment of my desire for a coffee with sugar and devoid of substances that bring me immediate pain. Perhaps some of the clauses of the composite desire have been fulfilled, such as drinking the coffee itself. The fulfilment of this clause generates the illusion of a desire fulfilment, but the desire fulfilment has not occurred because the fulfilment of a composite desire requires that all its component parts be fulfilled. Therefore, based on the observation that desires contain multiple clauses, we can claim that the salted coffee case does not describe an instance of actual desire fulfilment at all because the necessary accompanying features stipulated in the desire's formulation are not present. Therefore, in this paradigmatic Dead Sea apple case, we can claim that I have not actually experienced any desire fulfilment at all, only a defective simulacrum of my desire's fulfilment. In cases such as these, we are of course motivated to action by a desire. But to say that one consumes the salted contents of the mug because they are motivated to do so by a desire, is not the same as saying that desire has necessarily been fulfilled when the contents of the mug is consumed and turns out not to contain what was expected. According to this position, many Dead Sea apples can be explained as not being cases desire fulfilments without the need to appeal to the instrumental/non-instrumental division of desires and the restriction that only the fulfilment and frustration of non-instrumental desires count towards determining welfare.

There are a number of objections that can be raised to this account. I will consider three and sketch brief responses to them. Firstly, an objector can claim that the salted coffee case does seem to involve a desire to bring the mug to my lips, and that clearly that desire has been fulfilled. This objector can argue that while I also had a desire for a cup of sweetened coffee that was frustrated, I had other *concurrent* desires that were fulfilled through the action of consuming the contents of the mug. On this view, desires are not composite entities made up of multiple clauses, but rather concurrent single clauses, each of which is susceptible to individual fulfilment or frustration. Call this position the 'multiple concurrent desires' view<sup>29</sup>. This view claims that we have multiple desires running concurrently about the same object or event, each of which involves simple propositions, some of which are fulfilled and others of which are frustrated when we encounter cases of Dead Sea apples. Bernard Williams highlights this seemingly common phenomenon of multiple concurrent desires when he writes that often we have a number desires about the same object or state of affairs. The example he gives of this phenomenon is that often, the desires 'My wanting that P, my wanting to have some beliefs on the subject of whether P, and my wanting my beliefs to be true' accompany each other (Williams, 1973: 262). His example highlights that we often have multiple desires about the same state of affairs and that, in his view, these desires ought to be treated as *separate* instead of components of a single integrated desire. If we apply the idea of concurrent desires to the salted coffee case, then we can claim that I have a series of desires in play when I drink its terrible saline contents; some of

---

<sup>29</sup> This is not to be confused with the concurrence view discussed earlier, which states that a desire must be fulfilled at the same time as the desire experienced in order for welfare to be generated.

which are frustrated and others of which are fulfilled by that action. Therefore, if we accept the multiple concurrent desires view, then the problem of Dead Sea apples remains for unrestricted non-mental state actualist desire theories<sup>30</sup>.

The response to this objection is simply to double down on the composite desire theory and claim that, while I did have a desire to bring my mug to my lips, the desire was not to bring *that* mug and its saline contents to my lips. Instead, I held a composite desire to bring a mug that contained coffee with sugar and devoid of substances that bring me immediate pain to my lips. While my action of picking up the salted coffee and placing it to my lips was motivated by a desire, on the composite desire theory, that desire was frustrated and not fulfilled by the sip of salted coffee. My proposal is that these features ought to be considered more fine-grained components of a single composite desire, the fulfilment of which depends upon the fulfilment of *all* its constituent parts. The disagreement between the multiple concurrent desires account and the composite desire account seems to be reduced to a different set of intuitions. Nevertheless, the composite desire theory seems in a better place to maintain the normative adequacy of our unrestricted non-mental state actualist desire theory due to its ability to ameliorate the problem of Dead Sea apples. This in turn may give us a reason to prefer it to the multiple concurrent desires view. Indeed, this seems to my mind to be the only viable alternative to a restricted desire theory or the concurrence view in response to the problem of Dead Sea apples for non-mental state actualist desire theories.

The second objection that is worth considering to the composite desire theory is raised by Sumner and is based on some observations about the nature of our desires. Sumner argues that cases of Dead Sea apples are an intractable problem for unrestricted non-mental state actualist desire theories. For Sumner, this problem emerges because the very structure of desire itself involves a knowledge gap on the part of the person desiring between their prospective desires and their actual experiential reaction to attaining subjective desire satisfaction. Sumner points out that we cannot truly know how we will react to a fulfilled desire until the desire itself has been fulfilled and we have perceived its fulfilment. Therefore, because desires are always prospective, they are bound to be approximations of the things that we believe that we want but which may *ex ante* turn out to be not reflective of what we want. For Sumner, the prospectivity of desire shows that desires themselves are not the locus of welfare generation but rather they attempt to track another quality that determines welfare:

‘That mistake resulted from the gap between my *ex ante* expectation and my *ex post* experience. That gap exists by virtue of the prospectivity of desire: my preferences about the future always represent my view *now* of how things will go *then*. Because the gap exists from the very nature of desire, it cannot be closed merely by requiring that desires be rational or considered or informed’ (Sumner, 1996: 131-132)

Sumner makes the intuitively plausible claim that because desires are prospective, full information about how we will feel once we experience subjective desire satisfaction is necessarily unavailable to us. Therefore, according to Sumner, we should expect Dead Sea apples to arise occasionally. Call this epistemic gap between *ex ante* expectation and *ex post* experience the ‘problem of prospectivity’ for unrestricted non-mental state actualist desire theories. The problem of prospectivity would seem to undermine the composite desire theory’s solution to Dead Sea apples because it suggests that even

---

<sup>30</sup> Indeed, not only does it remain but it seems magnified. If desires tend to be single clauses rather than composite entities, then inevitably we will encounter more Dead Sea apples, as some of the clauses when fulfilled seem completely unrelated to welfare. It is by clustering individual clauses together under the rubric of composite desires that we minimise or remove the possibility of this happening.

composite desires are necessarily prospective and therefore approximations about what we believe may enhance our welfare. These approximations may turn out to be Dead Sea apples upon fulfilment.

Sumner raises an important problem for non-mental state actualist desire theories. However, the problem of prospectivity need not undermine the viability of the composite desire theory's response to the problem of Dead Sea apples. This is for two reasons. Firstly, an unrestricted non-mental state actualist desire theory does not need to claim that desire fulfilment will necessarily always produce all the benefits anticipated by the person desiring. It is perfectly consistent with the unrestricted non-mental state actualist desire theory to claim that often we have false beliefs about the intensity of our desires and that therefore outcomes do not produce the anticipated amount of welfare generation. The theory is able to claim that often we get it wrong in the *degree* of welfare generated by a composite desire's fulfilment. Indeed, all that an unrestricted non-mental state theory needs to claim in order to navigate the problem of prospectivity is that *some* welfare is always generated by composite desire fulfilments. It seems intuitively plausible to suggest that this minimal requirement of some welfare being generated is always present in cases of genuine composite desire fulfilments. Therefore, to make the composite desire theory robust to the problem of prospectivity, we will have to claim that composite desires always track objective conditions that if brought about will lead to an amount, however small, of positive welfare generation occurring. In this way, the problem of prospectivity is reduced to one that effects only the degree to which a desire fulfilment enhances welfare, and not the premise that desire fulfilments always do have some positive non-instrumental effect on welfare, even if those positive effects are outweighed by other negatives effects. The second reason why the problem of prospectivity is not completely devastating to the unrestricted non-mental state actualist desire theory is found in the all-things-considered bad for welfare explanation of defective desires. By applying this argument, we can claim that often the reason why our subjective experience of desire fulfilment does not match up to our anticipation of a desire's fulfilment is to do with the other desire's that are frustrated as a consequence of our initial desire's fulfilment. Therefore, the problem of prospectivity need not undermine the composite desire theory's approach to explaining Dead Sea apples, providing we are willing to make these claims.

The third problem worth considering for the composite desire theory is raised by Lauinger, who claims that it is simply intuitively implausible to explain all Dead Sea apples through appeal to the composite desire theory. In Lauinger's words, 'It goes too far to say that people *never* desire to have certain jobs (full stop), that people *never* desire to be in certain relationships (full stop), etc.' (Lauinger, 2011: 328). The issue he is raising here is that sometimes our desires are not full of elaborate clauses and are instead simply desires for states of affairs without a series of caveats specifying exactly how those states of affairs ought to look. The point he is raising here is somewhat diminished by his examples, as these examples all track quite major desire fulfilments such as taking a new job or entering into a new relationship. Consequently, these examples seem to not need the argumentative resources provided by the composite desire theory, as they can instead be explained by appeal to the all-things-considered bad for welfare theory. When we experience major desire fulfilments, we will inevitably find that they lead to a whole host of other desire fulfilments and frustrations. We can appeal to the aggregate desire frustration to explain cases where taking that job or entering the new relationship seems to be a Dead Sea apple. The all-things-considered bad for welfare argument in this case claims simply that, while your desire was fulfilled, its fulfilment was vastly outweighed to the whole host of desire frustrations that were invited into your life by its fulfilment. Therefore, if someone does have the desire to take that job or enter that relationship (full stop), then we do not need a composite desire theory to explain why it was bad for their welfare. The account becomes more necessary in small scale-cases of desire fulfilments when there are no further desire fulfilments or frustrations present to explain why the desire's fulfilment was all-things-considered bad for welfare.

In the more small-scale cases, such as drinking a cup of salted coffee, the claim that the desire is composite and therefore not truly fulfilled seems much more plausible. This is because the desire does not require a great number of clauses linked to the desire's fulfilment in order for us to rule out cases of drinking salted coffee as not describing actual desire fulfilment. Lauinger could claim at this point that it is possible to simply desire to drink the substance in the mug (full stop) without consideration for its contents and to experience a Dead Sea apple as a result. The counterclaim from the unrestricted non-mental state actualist desire theorist at this point is to say, in the event that someone really held that desire (full stop), then its fulfilment would not have been completely devoid of welfare. We explain the complete absence of welfare by claiming that you never got what was desired, only a simulacrum of that desire. This explanation seems, to my mind at least, intuitively plausible.

Nevertheless, I do think that Lauinger's concern that sometimes we do just desire things (full stop) serves as a cautionary warning that we should not allow our account of composite desires to become too broad in its ability to stipulate large amounts of clauses in order for a desire to be fulfilled. Desires cannot contain too many clauses necessary for fulfilment, or else we would never achieve the fulfilment of any desires. This seems intuitively wrong. Moreover, sometimes I may not get exactly what I desire but I get something relevantly similar to it. In these cases, it seems that the approximation to my desired state of affairs is close enough for me to claim that my desire has been fulfilled despite it not matching exactly my *ex ante* expectation of what fulfilment would look like. For example, those who order shopping online will be familiar with situations where replacement shopping items arrive in place of specifically ordered items when they are out of stock. Often these replacement items serve to fulfil our desires just as well as the originally selected item. If we allow our account of composite desires to specify too many clauses necessary for the fulfilment of composite desires, then we will struggle to explain cases like this, where a desire is fulfilled but not in the way that we necessarily expected it to be. One way of explaining this situation on the composite desire theory is to claim that the initial desire for item X was frustrated, but that on appearance of item A, we immediately experienced a desire for item A that was fulfilled instantly upon its arrival. This seems profoundly counterintuitive to explain cases where items X and A are relevantly similar. Perhaps a better way of explaining these cases is to claim that my initial desire for item X was *reformulated* on appearance of item A rather than frustrated. This may be a more convincing approach if we allow such reformulations to reshape the clauses necessary for a composite desire's fulfilment. If we consider desire fulfilments as analogous to contract fulfilments, then we can consider cases like these as renegotiations of the contract's terms in response to new phenomena. Nevertheless, we have good reasons to be sceptical of overly specific desires, and any composite desire theory will need to limit the scope of just how specific the conditions for a desire's fulfilment can be.

More work is required to understand exactly issues around how the components of any one desire are tied together in cases of composite desires. Nevertheless, the outline of this argument does illustrate how the unrestricted non-mental state actualist desire theory can claim that some instances of Dead Sea apples fail to describe desire fulfilments at all because the outcome does not fulfil crucial clauses that were central to the actual desire held. Therefore, we have a robust alternative to the concurrence theory in order to explain many Dead Sea apples on the unrestricted non-mental state actualist desire theory. This alternative does not require us to make theoretical commitments to the effects of past desire or posthumous desire fulfilments and frustrations on welfare. Neither does it require us to make theoretical commitments to whether only the fulfilment of non-instrumental desires enhance welfare, and indeed whether the instrumental / non-instrumental categorisation can be made in a satisfactory way. Instead it involves a theoretical commitment to some claims about the structure of our actual typical desires, which are then argued as justified through reference to the composite desire theory.

**Final remarks on Dead Sea apples**

I believe that the composite desire theory can be used to explain most Dead Sea apple examples. Others I believe can be accounted for through the all-things-considered bad for welfare explanation. Nevertheless, while I think there are good reasons for adopting these approaches to explain many cases of Dead Sea apples, I do not believe that it is necessary or convincing to evoke this response to explain every example given of Dead Sea apples. Indeed, to do so would undermine some of the plasticity of the unrestricted non-mental state actualist desire theory, which enables us to value things as welfare affecting outside of our own subjective desire satisfaction. Conversely, the problem of Dead Sea apples rests solely on reference to the lack of a feeling of subjective desire satisfaction when we perceive the fulfilment of one of our actual desires. Yet, if occurrences outside of our mental states can nevertheless affect how well our life has gone, then it seems fair to suggest that some Dead Sea apples may also positively affect our welfare while failing to arouse the feeling of subjective desire satisfaction within us. There may be cases of Dead sea apples, where despite our immediate discontinuation of the desire at the moment that we subjectively perceive the desire's fulfilment, there are reasons to think that the desire fulfilment may have provided some non-instrumental benefits to our welfare. Therefore, the third way that the non-mental state actualist desire theory can account for cases of Dead Sea apples is to claim that there is, even in cases like these, a degree of welfare is generated from fulfilment thus making them all-things-considered bad for welfare and unproblematic to the unrestricted non-mental state actualist desire theory. This welfare generation need not be tied to subjective mental states. Indeed, one of the virtues of unrestricted non-mental state actualist desire theories is that they can value things that do not meet the Experience Requirement. While it may be initially imperceptible to our pre-theoretic intuitions, I think that there is a good reason to believe that such cases do involve an element of non-instrumental welfare enhancement. Therefore, my approach involves adopting a mixture of the composite desire theory argument against Dead Sea apples, the all-things-considered bad for welfare account of defective desires, and a degree of 'biting the bullet' to explain those examples of Dead Sea apples that survive these two other arguments. While either the composite desire theory approach or the 'biting the bullet' position alone will suffice to explain these types desires, I believe that there are good reasons for us to distinguish between different cases where we may want to apply each explanation separately. This preserves the plasticity of the unrestricted non-mental state actualist desire theory in allowing it to value things outside of subjective desire satisfaction as non-instrumentally welfare enhancing, while also allowing us to dismiss other cases, such as the salted coffee Dead Sea apple, as not being cases of desire fulfilment at all.

## Chapter 4: The problem of intrinsically quirky desires

### Introduction

This chapter seeks to build upon the work of the previous two chapters by briefly surveying the problems presented by intrinsically ‘quirky desires’ for unrestricted non-mental state actualist desire theories. Quirky desires are desires where it is claimed that fulfilment does not non-instrumentally enhance a person’s welfare. We can contrast quirky desires with defective desires, which claim that the fulfilment of these desires actively detracts from a person’s welfare. Defective desires are best understood as a subset of quirky desires, and in practice there is a lot of overlap between the two. The case of Dead Sea apples for example can be either defective *and* quirky if negative welfare is generated by their fulfilment or alternatively they can be simply quirky if no welfare is non-instrumentally generated by their fulfilment. Quirky desires are another type of desire said to undermine the extensional adequacy of unrestricted non-mental state actualist desire theories of welfare. This is because these theories claim that desire fulfilment ought to always non-instrumentally enhance welfare. The desires that I am considering in this chapter are labelled ‘intrinsically’ quirky because they are all desires that writers have found fault with due to some inherent feature they are said to possess or that is inherent to their acquirement. I have labelled base desires, poorly cultivated desires, pointless desires and artificially aroused desires as all being examples of intrinsically quirky desires in this discussion (Heathwood, 2005: 487). However, my labelling of them as intrinsically quirky rather than defective is contestable, for an objector can claim that the fulfilment of some, or indeed all, of these desires not only fails to non-instrumentally enhance welfare but rather actively detracts from welfare instead. In that case, these desires would become defective rather than simply quirky. Nevertheless, there is an additional benefit to grouping all these desires together under the banner of intrinsically quirky for the sake of this discussion, which is that they can all be answered with very similar arguments.

The reason given for intrinsically quirky desires failing to non-instrumentally generate welfare upon fulfilment is not because they necessarily lead to a lack of subjective desire satisfaction or a sense of regret in the agent, but rather because there seems to be something *intrinsically* wrong about claiming that these types of desire fulfilments ought to contribute non-instrumentally to welfare, or at least contribute non-instrumentally to welfare to the same extent as regular desires of the same intensity. Moreover, this intrinsic wrongness seems to be present *irrespective* of whether their fulfilment gives rise to subjective desire satisfaction or fulfils long-standing and subjectively important desires to the person experiencing them. Consequently, it makes sense to treat these cases together because they all merit a similar response. These issues are all given a relatively fast treatment, for I think that adequate responses have been elaborated within the pre-existing literature on these topics. Overall this chapter is much briefer than the preceding chapters because it draws heavily on well-trodden responses to solve these problems for unrestricted non-mental state actualist desire theories. Nevertheless, a consideration of these problems is necessary if we are to successfully defend the claim that unrestricted non-mental state actualist desire theories can maintain extensional adequacy. In this chapter I consider poorly cultivated desires, pointless desires, base desires and artificially aroused desires together and claim that their existence does not undermine the unrestricted non-mental state actualist desire theory. I conclude that we ought to essentially ‘bite the bullet’ in claiming that all these types of desires do non-instrumentally enhance welfare when fulfilled and outline a set of reasons to ameliorate intuitions that run counter to that conclusion.

### Intrinsically Quirky Desires

Let us begin with a consideration of pointless desires. Pointless desires are those desires which seem very strange and without purpose to outside observers (Heathwood, 2005: 488). A classic example of a pointless desire is provided by Rawls, who outlines the case of a person who dedicates their life to counting blades of grass<sup>31</sup>. Carson outlines Rawls' case in this way, 'Rawls asks us to consider the case of a person who possesses the capacity to do important work in applied math, but instead prefers to devote himself to counting blades of grass. Some contend that this person would be better off (or have a better life) if he became a mathematician, even though he prefers to count blades of grass' (Carson, 2000: 80). The question then arises as to whether we ought to claim that, other things being equal, the fulfilment of pointless desires non-instrumentally generates as much welfare as the fulfilment of more relatable desires would do. Most writers in the literature on unrestricted non-mental state actualist desire theories simply answer this question in the affirmative (Keller, 2004: 31; Heathwood, 2005: 500; Mendola, 2009: 162). Indeed, it is a necessary commitment of unrestricted non-mental state actualist desire theories that they affirm the autonomy of the individual in placing value on whatever they desire irrespective of any objective standards. Nevertheless, writers in this tradition have given additional reasons to ameliorate the intuitive concern that it is extensionally inadequate to claim that the fulfilment of pointless desires non-instrumentally enhances welfare to the same degree as non-pointless desire fulfilments.

Mendola likens the fulfilment of pointless desires to many other much more common desires that people hold in order to draw support for the intuition that the degree of pointlessness does not affect the welfare generated by its fulfilment, 'It seems to me anyway that drinking mud or counting grass blades are no more objectively worse for one than some highly revered religious practices' (Mendola, 2009: 162). Heathwood takes a similar approach by claiming that the line between the desires that we consider pointless and those which we do not seems to be unclear, 'We can't say that piano playing, stamp collecting, and rock climbing escape pointlessness because their point is the fun of doing so. Counting the blades of grass and throwing the ball against the wall can have a point in that sense, too' (Heathwood, 2005: 499). Consequently, it appears that generating a criterion for pointlessness would involve tricky theoretical commitments that may lead us to devalue some of the desires that we currently have, and which do not intuitively seem to be pointless. An additional reason why we might intuitively recoil at the idea that the pointlessness of a desire does not non-instrumentally effect the amount of welfare generated by its fulfilment is because we often have instrumental reasons to avoid the fulfilment of such desires (Bruckner, 2016: 2; Heathwood, 2005: 492; Mendola, 2009: 162). This is especially the case in some examples of pointless desires such as eating a bowl of gravel, which inevitably seem to be tied to the frustration of other desires – such as, for example, the desire not to be in pain. In this way, the intuition that the pointlessness of a desire devalues the amount of welfare non-instrumentally generated by its fulfilment may arise from us identifying good instrumental reasons for being averse to pointless desire fulfilments. However, the unrestricted non-mental state actualist desire theorist is committed to the claim that some compensatory welfare is to be found even in such cases of pointless desires. In Keller's words, 'on the Unrestricted View there is, from the point of view of your welfare, something to be said for eating the gravel' (Keller, 2004: 28). Another reason why we may want to revise our intuition that the fulfilment of pointless desires does not non-instrumentally enhance welfare to the same extent as the fulfilment of non-pointless desires is provided by Keller. He claims that, when we consider someone

---

<sup>31</sup> There is not uniformity in the literature on the terminology used for these types of desires. Heathwood labels them pointless desires, Keller calls them 'crazy goals' (Keller, 2004: 30), and Griffin labels them irrational desires (Griffin, 1986: 25). I have opted to use Heathwood's classificatory system.

endeavouring to fulfil a pointless desire, it is surely better for them that they actually fulfil that desire, all things being equal, than that they fail to fulfil it (Keller, 2004: 31). In this way, the good grass counter who endeavours successfully to accurately record grass numbers would have more welfare generated by this activity than the bad grass counter who hopelessly loses track of the total number of grass counted. If we subscribe to this intuition, then Keller thinks that we ought to assent to the idea that pointless desires non-instrumentally enhance welfare to the same extent as non-pointless desires when fulfilled.

A final reason that we may want to revise the intuition that the fulfilment of pointless desires does not non-instrumentally enhance a person's welfare, or does not non-instrumentally enhance a person's welfare to the same extent as the fulfilment of a non-pointless desire of the same intensity does, is found in Heathwood's position that, 'I explain our feeling that the lives described in the arguments are lacking by pointing out that the lives rank low on other scales of evaluation that we care about, scales different from the welfare scale' (Heathwood, 2005: 500). Consequently, Heathwood claims that we intuitively recoil at the idea that pointless desire fulfilments non-instrumentally enhance welfare to the same degree as non-pointless desire fulfilments of the same intensity because our intuitions are mistaking welfare value with other scales of value that are not enhanced by the pointless desire's fulfilment. He points to other writers who have made a similar point that welfare is not the only way of evaluating an individual's life that seems to matter normatively to us (Feldman, 2004: 8; Griffin, 1986: 23; Sumner, 1996: 20). We may intuitively believe that pointless desire fulfilments do not non-instrumentally enhance welfare to the same extent as non-pointless desire fulfilments because they are devoid of, for example, ethical or aesthetic value. Therefore, we have a host of arguments that may go some way to ameliorating the intuition that pointless desire fulfilments do not non-instrumentally enhance welfare to the same extent as non-pointless desire fulfilments of the same intensity.

These considerations can be analogously applied to the case of base desire fulfilments. These are desires which seem inherently immoral to fulfil. For example, the sadistic desire of taking pleasure in the suffering of non-consenting innocent people may be a prime example of a base desire. The immorality of the desire is sometimes said to undermine the amount of welfare non-instrumentally generated by its fulfilment. Heathwood explains the intuition that base desire fulfilments do not non-instrumentally enhance welfare by claiming that our intuitions are tracking ethical value rather than welfare values in these cases, 'The reason, I submit, that we react strongly against a life of perpetual indulgence in bestiality is that we care about more than just welfare. We don't merely want ourselves (or those we love) to be well off; we also want (or want them) to do good things, to be good people, to achieve worthwhile goals' (Heathwood, 2005: 498). This strikes me as a satisfactory argument. It is intuitively plausible that someone's life could be good for them despite their life being a blight upon the lives of others. Indeed, this may be part of the reason why we react so intuitively strongly against unethical actions.

Similar considerations also seem relevant to cases of so-called poorly cultivated desires. These are the desires for states of affairs that it seems we would not have had we taken more time to develop a sophisticated appreciation of other categorically similar states of affairs, which seem to outside observers to be more intrinsically valuable. For example, someone's appreciation of terrible art or music may be said to be poorly cultivated if it stems partly from a lack of exposure to, and time considering, objectively good art and music. Once again, we can apply the argument that our intuitions suggest that these poorly cultivated desire fulfilments do not non-instrumentally enhance welfare to the same extent as well cultivated desire fulfilments because our intuitions are picking up on alternative scales of value, such as aesthetic value. It may be that listening to objectively terrible music

detracts from the total amount of aesthetic value experienced over the course of a lifetime in a way that listening to good music does not do. However, this has no bearing on the welfare value of the life according to the unrestricted non-mental state actualist desire theory. Heathwood adds to this argument the additional intuition that well cultivated desires tend also to have the instrumental benefit of increasing the intensity of our desires, 'Since the desires are going to be satisfied anyway, one gets more out of them, in terms of welfare, the more intense they are. My mediocre desires never light my fire so much, and I'm made worse off for it' (Heathwood, 2005: 495). If Heathwood is right, then it may be that well cultivated desires are in general more intense than poorly cultivated desires and thus generate more welfare when fulfilled. Consequently, we would have instrumental reasons to cultivate our desires adequately.

The final subtype of intrinsically quirky desires that I survey in this chapter are artificially aroused desires. These are the desires that people have which intuitively seem to have their origin in problematic causes. These causes are said to instil desires within persons that they previously had no desire for and in a way deemed intuitively problematic. Heathwood's example of a potential cause of artificially aroused desires is found in cases of advertising (Heathwood, 2005: 493-494). In this case, we have a whole industry designed to stimulate artificially aroused desires within the population. However, Heathwood argues that we ought not to be paternalistic in claiming that some desires have intrinsically problematic causes and therefore ought not to count as non-instrumentally enhancing of welfare. Instead, he appeals to the anti-paternalistic nature of desire theories in allowing individual's subjective desires to determine their own conditions for welfare generation. Moreover, as with cases of pointless desires, it would be theoretically challenging to establish a criterion for differentiating between artificially aroused desires and legitimately aroused desires. Mendola goes so far as to claim that, 'There is no determinate truth about what you would desire if you were suitably insulated from distorting social influences, because it is not objectively fixed in some prior way what is proper education and what is distorting influence' (Mendola, 2009: 161). However, Heathwood does provide an instrumental reason that may lead us to claim artificially aroused desire fulfilments are generally not as beneficial to welfare as legitimately aroused desire fulfilments would be. This is because he thinks they are much more likely to lead to Dead Sea apples when fulfilled (Heathwood, 2009: 493).

I think that arguments from pointless desires, base desires, poorly cultivated desires and artificially aroused desires are much less damaging to the extensional adequacy than cases of Dead Sea apples and defective desires. Therefore, I have not granted as much of a detailed consideration to these problems. What I have attempted to do here is provide a selection of arguments that ameliorate the intuitive force of the claim that these intrinsically quirky desires undermine the extensional adequacy of unrestricted non-mental state actualist desire theories. These arguments essentially amount to 'biting the bullet' in accepting that these types of desires do, all things being equal, non-instrumentally enhance welfare when fulfilled to the same degree as normal desire fulfilments do.

## Conclusion

This paper has considered a set of problems that have been claimed undermine the extensional adequacy of unrestricted non-mental state actualist desire theories. I have divided these problems into problems of defective desires, Dead Sea apples, and intrinsically quirky desires. What I have set out to do is show how unrestricted non-mental state actualist desire theories have the resources to explain these phenomenon without recourse to idealisation or restricted variations of the desire theory. In my first chapter, I examined the features of unrestricted non-mental state actualist desire theories and some of the reasons motivating their adoption. My second chapter then considered the problem of defective desires and showed how the all-things-considered bad for welfare argument could be employed to answer many of them, while also explaining how ethical and other-regarding actions fit into unrestricted non-mental state actualist desire theories. My third chapter examined the problem of Dead Sea apples and argued against Heathwood's position that we ought to adopt the concurrence view to solve it. Instead, I argued that adopting a 'composite desire theory' can ameliorate this problem substantially and that any remaining examples of Dead Sea apples can be explained as involving some non-instrumental welfare enhancement despite not giving rise to subjective desire satisfaction. My final chapter surveyed the problem of intrinsically quirky desires and surveyed the wealth of pre-existing arguments against the claim that they undermine unrestricted non-mental state actualist desire theories. Between them these arguments have provided a set of reasons against thinking that unrestricted non-mental state actualist desire theories are made extensionally inadequate by these sorts of problematic desire fulfilments. However, that is not to say that there are no other problems that may prove fatal to the viability of these types of theory.

## Bibliography:

- Barnes, Jonathan & Anthony Kenny (2014). *Aristotle's Ethics: The Complete Writings*. Oxfordshire: Princeton University Press.
- Bradley, Ben (2007). 'A Paradox for Some Theories of Welfare'. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*. 133(1): 45-53.
- Brandt, Richard B. (1979). *A Theory of the Good and the Right*. Clarendon Press: Oxford.
- Brandt, Richard B. (1955). 'The Definition of an "Ideal Observer" Theory in Ethics'. *Philosophy and Phenomenological Research*. 15(3): 407-413.
- Bruckner, Donald W. (2016). 'Quirky Desires and Well-Being'. *Journal of Ethics and Social Philosophy*. 10(2): 1-34.
- Carson, Thomas L. (2000). *Value and the Good Life*. University of Notre Dame Press: Notre Dame.
- Cowen, Tyler and Derek Parfit (1992). 'Against the Social Discount Rate' in Fishkin and Laslett (eds.). *Justice Between Age Groups and Generations*. Yale University Press: New Haven.
- Dorsey, Dale (2013). 'Desire-satisfaction and Welfare as Temporal'. *Ethical Theory and Moral Practice*. 16(1): 151-171.
- Donner, Wendy (1991). *The Liberal Self: John Stuart Mill's Moral and Political Theory*. Cornell University Press: New York.
- Enoch, David (2005). 'Why Idealize?'. *Ethics*. 115(4): 759-787.
- Fehige, Christoph (1998). 'A Pareto Principle for Possible People' in Fehige and Wessels (eds.). *Preferences*. Walter de Gruyter: New York.
- Fehige, Christoph and Ulla Wessels (1998). *Preferences*. Walter de Gruyter: New York.
- Feldman, Fred (2004). *Pleasure and the Good Life: Concerning the Nature, Varieties, and Plausibility of Hedonism*. Oxford University Press: Oxford.
- Feldman, Fred (2002). 'The Good Life: A Defense of Attitudinal Hedonism'. *Philosophy and Phenomenological Research*. 65(3): 604-628.
- Griffin, James (1986). *Well-Being: Its Meaning, Measurement, and Moral Importance*. Clarendon Press: Oxford.
- Heathwood, Chris (2006). 'Desire Satisfactionism and Hedonism'. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*. 128(3): 539-563.
- Heathwood, Chris (2005). 'The Problem of Defective Desires'. *Australasian Journal of Philosophy*. 83(4): 487-504.
- Hubin, Donald C. (2003). 'Desires, Whims and Values'. *The Journal of Ethics*. 7(3): 315-335.
- Hubin, Donald C. (1996). 'Hypothetical Motivation'. *Noûs*. 30(1): 31-54.
- Hubin, Donald C. (1991). 'Irrational Desires'. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*. 62(1): 23-44.

- Keller, Simon (2004). 'Welfare and the Achievement of Goals'. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*. 121(1): 27-41.
- Lauinger, William (2011). 'Dead Sea Apples and Desire-Fulfillment Welfare Theories'. *Utilitas*. 23(3): 324-343.
- Loeb, Don (1995). 'Full-Information Theories of Individual Good'. *Social Theory and Practice*. 21(1): 1-30.
- Lukas, Mark (2010). 'Desire Satisfactionism and the Problem of Irrelevant Desires'. *Journal of Ethics & Social Philosophy*. 4(2): 1-24.
- Mendola, Joseph (2009). 'Real Desires and Well-Being'. *Philosophical Issues*. 19, Metaethics: 148-165.
- Mill, John Stuart (1993). *Utilitarianism, on Liberty, Considerations on Representative Government*. Geraint, Williams (ed.). Everyman: London & Rutland.
- Nozick, Robert (1974). *Anarchy, State, and Utopia*. Blackwell Publishing: Malden, Oxford & Carlton.
- Parfit, Derek (1984). *Reasons and Persons*. Oxford University Press: New York.
- Portmore, Douglas W. (2007). 'Desire Fulfillment and Posthumous Harm'. *American Philosophical Quarterly*. 44(1): 27-38.
- Rosati, Connie S. (1995). 'Persons, Perspectives, and Full Information Accounts of the Good'. *Ethics*. 105(2): 296-325.
- Rosati, Connie S. (1996). 'Internalism and the Good for a Person'. *Ethics*. 106(2): 297-326.
- Sidgwick, Henry (1981) [1907]. *The Methods of Ethics*. 7<sup>th</sup> ed. Hackett Publishing Company: Indianapolis.
- Sobel, David (1997). 'On the subjectivity of Welfare'. *Ethics*. 107(3): 501-508.
- Sumner, L. W. (1996). *Welfare, Happiness, and Ethics*. Clarendon Press: Oxford.
- Vorobej, Mark (1998). 'Past Desires'. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*. 90(3): 305-318.
- Williams, Bernard (1973). 'Egoism and Altruism' in *Problems of the Self*. Cambridge University Press: Cambridge. 250-267.
- Woodard, Christopher (2013). 'Classifying theories of welfare'. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*. 165(3): 787-803.