

# **How Can We Successfully Measure Well-being Through Measuring Happiness?**

Sam Kieran Jonah Wren-Lewis

Submitted in accordance with the requirements for the degree of  
Doctor of Philosophy

The University of Leeds

Department of Philosophy

January, 2014

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Sam Wren-Lewis to be identified as Author of this work has been asserted by him in accordance with the Copyright, Designs and Patents Act 1988.

© 2014 The University of Leeds and Sam Kieran Jonah Wren-Lewis

## **Acknowledgements**

I would like to thank my family, girlfriend and friends for their continued support throughout my time as a philosopher. In addition, I am grateful to my supervisor, colleagues and other members of my department for all their help over the past few years.

## Abstract

This thesis is an investigation into how successfully we can measure well-being through measuring happiness. We care about our well-being – how well our lives are going for us. And we often treat happiness as a proxy for well-being in our practical lives. It is no surprise, therefore, that there is a burgeoning field of social science that aims to measure well-being through measuring happiness.

There is no consensus in this field, however, over what well-being and happiness are, and how successfully we can measure the former through measuring the latter. It is these issues that I aim to address in this thesis. I will argue in favour of the Indicator View, according to which we should treat happiness as an indicator of *local changes* in well-being. I will further argue that some of the local changes in well-being indicated by happiness constitute an important aspect of our well-being (what I will call our *local well-being*). The upshot is that we can measure an important aspect of well-being – namely *changes in local well-being* – through measuring happiness.

This conclusion is directly relevant to the empirical study of well-being for two reasons. First, I will employ an account of happiness that is similar to one of the main constructs used by social scientists in the measurement of well-being. Second, my conclusion rests on an understanding of well-being that is consistent with all plausible substantive theories of well-being.

## Table of Contents

<b>Acknowledgements</b> .....	<b>i</b>
<b>Abstract</b> .....	<b>ii</b>
<b>Table of Contents</b> .....	<b>iii</b>
<b>List of Figures</b> .....	<b>vii</b>
<b>Introduction</b> .....	<b>1</b>
<b>Chapter 1 Measuring Well-being Through Measuring Happiness</b> .....	<b>4</b>
Abstract .....	4
1.1 Introduction .....	5
1.2 Why measure well-being through measuring happiness? .....	6
1.2.1 Why measure well-being through measuring something else? .....	6
1.2.2 Why measure well-being through measuring happiness?.....	8
1.2.3 Why think that well-being and happiness tend to be strongly correlated? .....	9
1.2.4 Assessing the evidence .....	14
1.3 Well-being: Concept(s).....	15
1.3.1 What is well-being? .....	15
1.3.2 Well-being as a fragmented concept.....	19
1.3.3 Well-being as a distinct concept.....	23
1.4 Happiness: Concept(s).....	27
1.4.1 The fragmentation of happiness.....	28
1.4.2 Summing up where we have got to so far.....	35
1.5 Well-being: Conception(s).....	36
1.5.1 Different conceptions of well-being .....	36
1.5.2 The theory-neutral account of well-being.....	38
1.5.3 The theory-neutral account and different measures of well-being.....	40
1.6 Happiness: Conception(s).....	45
1.6.1 Different conceptions of psychological happiness .....	45
1.6.2 An affective state account of happiness.....	47
Conclusion .....	53

<b>Chapter 2 The Relationship Between Happiness and Well-being .....</b>	<b>55</b>
Abstract .....	55
2.1 Introduction .....	56
2.2 Substantive theories of well-being .....	56
2.2.1 A brief overview of substantive theories of well-being .....	56
2.2.2 Implications of widespread disagreement .....	61
2.3 The theory-neutral account of well-being .....	65
2.3.1 Platitudes about well-being .....	66
2.3.2 Consensus over well-being .....	67
2.3.3 Well-being ingredients.....	70
2.3.4 Summing up and looking forward.....	74
2.4 Prudential Hedonism.....	76
2.4.1 Prudential Hedonism and the Pro-Attitude Requirement ...	77
2.4.2 Prudential Hedonism and the Experience Requirement ....	78
2.4.3 The Problem with Prudential Hedonism.....	82
2.4.4 Another Problem with Prudential Hedonism .....	87
2.4.5 What is Right and Wrong about Hedonism .....	93
2.5 Adaptive Preferences.....	95
2.6 The Theory-Neutral Account Revisited .....	98
2.6.1 The theory-neutral account and the things we care about.....	99
2.6.2 Is the theory-neutral account a substantive theory of well-being?.....	100
Conclusion .....	101
<b>Chapter 3 Happiness as an Indicator of Well-being.....</b>	<b>103</b>
Abstract .....	103
3.1 Introduction .....	104
3.2 Happiness indicates well-being.....	106
3.2.1 The function of affective states .....	106
3.2.2 The historical function of affective states .....	110
3.3 Happiness as a defeasible indicator of well-being .....	113
3.3.1 Non-prudential affective states.....	113
3.3.2 Dysfunctional affective states.....	118
3.3.3 Refining the Indicator View .....	122
3.4 Happiness defeaters .....	124

3.4.1 Undercutting and rebutting defeaters.....	124
3.4.2 Evidence of happiness defeaters.....	127
3.4.3 Potential evidence of happiness defeaters.....	128
Conclusion.....	130
<b>Chapter 4 Happiness as an Indicator of Changes in Well-being.....</b>	<b>132</b>
Abstract.....	132
4.1 Introduction.....	133
4.2 Three views of the function of affective states.....	133
4.2.1 Millgram's View.....	134
4.2.2 Schroeder's View.....	138
4.2.3 Railton's View.....	141
4.3 The Refined Indicator View.....	146
4.3.1 Affective states and (unexpected) changes in well-being	147
4.3.2 Affective states and certain aspects of one's level of well-being.....	150
4.3.3 Putting it all together.....	153
4.4. Using the Indicator View to Interpret Empirical Findings.....	154
4.4.1 The Easterlin Paradox and the Phenomenon of Adaptation.....	154
4.4.2 Happiness promoting strategies.....	158
4.4.3 The key determinants of lasting happiness.....	160
Conclusion.....	165
<b>Chapter 5 Happiness as an Indicator of Local Well-being.....</b>	<b>167</b>
Abstract.....	167
5.1 Happiness as a defeasible indicator of local changes in well- being.....	169
5.1.1 Affective states and local changes in well-being.....	169
5.1.2 Mood, life satisfaction and global changes in well-being .	171
5.2 The instrumental value of the local changes in well-being indicated by happiness.....	175
5.2.1 Aggregating the size of local changes in well-being over time.....	176
5.2.2 Aggregating instances of local changes in well-being over time.....	180
5.3 The intrinsic value of the local changes in well-being indicated by happiness.....	183
5.3.1 Local changes in well-being and intrinsic motivation.....	184

5.3.2 Local changes in well-being and engaging with things of value .....	189
5.4 The global challenge .....	191
5.4.1 Global changes in well-being that are part of a good narrative .....	193
5.4.2 Global changes in well-being that we deeply care about .....	198
5.5 The prudential value of both local and global changes in well-being .....	200
Conclusion .....	204
<b>Conclusion .....</b>	<b>206</b>
<b>Bibliography .....</b>	<b>208</b>
<b>Appendix A: Happiness as an Indicator of Changes in Well-being in the Near-Future .....</b>	<b>227</b>
<b>Appendix B: Happiness as an Indicator of Levels of Well-being .....</b>	<b>232</b>
B.1 The Progress Method .....	232
B.2 The Steady Progress Method .....	233



## List of Figures

<b>Fig. 1.1.....</b>	<b>9</b>
<b>Fig. 1.2.....</b>	<b>23</b>
<b>Fig. 1.3.....</b>	<b>43</b>
<b>Fig. 2.1.....</b>	<b>73</b>
<b>Fig. 4.1.....</b>	<b>161</b>
<b>Fig. 5.1.....</b>	<b>177</b>

## Introduction

Happiness is currently the topic of a wide range of empirical research, and is increasingly becoming the focus of public policy. The interest in happiness largely stems from its connection with well-being. We care about well-being – how well our lives are going *for us*. If we are happy, it seems that, to some extent, we must be doing well. This suggests that we may be able to successfully measure well-being through measuring happiness.

The problem is that both happiness and well-being are elusive and their measurement is far from uncontroversial. What exactly does information about happiness tell us about well-being? Is there more to well-being than happiness? If so, to what extent is happiness connected to well-being? Does the relationship between happiness and well-being hold in certain contexts, such as in the case of deluded, base or immoral forms of happiness? These are controversial questions, but answers to them must be given if we are to make progress in the measurement of well-being.

In this thesis, I will focus on a particular view of happiness that is currently employed by empirical researchers in the measurement of well-being. This is an affective state view of happiness, according to which happiness is a person's (broad, relatively long-term) affective state. In chapter one, I will argue that this conception of happiness adequately accounts for the fact that happiness tends to be used as a proxy for well-being by laypersons (in practical deliberation and evaluation) as well as empirical researchers (in the measurement of well-being). This provides us with good reason to suppose that happiness tends to be strongly correlated with well-being.

How can we further investigate this correlation? In chapter two, I will argue that we can investigate the correlation between happiness and well-being without a substantive theory of well-being. This approach has been relatively unexplored in the philosophical literature, yet has the advantage of avoiding substantive disputes over the nature of well-being that hinder the prospect of measuring well-being through measuring happiness. I will develop a theory-neutral account of well-being, which provides us with an understanding of well-being that is consistent with most substantive theories of well-being. According to the theory-neutral account of well-being, happiness may be correlated with well-being either because it is constitutively or causally related

to well-being. The extent to which happiness is correlated with well-being is determined by its causal connections to “well-being ingredients” (the objects of platitudes about well-being that are causally connected to the objects of other well-being platitudes).

In the remainder of the thesis, I will consider the extent to which happiness is causally connected to well-being ingredients. I will show that, from the function of happiness and the theory-neutral account of well-being, we can infer the extent to which happiness is correlated with well-being.

In chapter three, I will argue that the function of happiness is to inform and guide action, which it does partly by detecting well-being ingredients. It is reasonable to presume that happiness is not generally dysfunctional and thereby that it reliably performs this function. However, in certain contexts, happiness should inform and guide action by detecting non-prudential features of our environment, such as moral or aesthetic features. And, under certain conditions, happiness fails to perform its function. We must be open, therefore, to the empirical possibility that happiness does not generally indicate well-being. Rather, I will argue that we should treat happiness as a *defeasible* indicator of well-being. I will call this the Indicator View.

In chapter four, I will refine this view. I will argue that happiness is a defeasible indicator of *changes* in well-being. This is a largely unappreciated feature of the relationship between happiness and well-being, yet has important implications for the measurement of well-being and happiness. I will additionally argue that, in certain contexts, happiness tends to indicate unexpected changes in well-being or certain aspects of one’s level of well-being. I will show that this refined version of the Indicator View provides us with plausible and informative interpretations of the Easterlin Paradox and the phenomenon of adaptation, as well as other key findings from the empirical study of happiness.

Lastly, in chapter five, I will argue that happiness tends to indicate *local* changes in well-being (i.e. how well we are doing, from moment-to-moment, in our short-term objectives, goals and concerns). These changes will not tend to be indicated by more global measures of well-being, such as measures of life satisfaction. Again, this has important implications of the measurement of well-being. I will argue many of the local changes in well-being indicated by happiness have intrinsic prudential value; I will call these changes in our *local well-being*.

Some theorists have recently argued that our global (in contrast to local) well-being has a distinct kind of value, in virtue of the fact that it is constituted by meaningful life narratives or concerns. I will argue against this view. I will argue that both local and global well-being constitute overall well-being in the same way; neither kind of well-being has a distinct kind of prudential value. Moreover, I will suggest that the relative values of people's local and global well-being vary across their lifespan, personality and cultural environment. The central conclusion of this thesis, therefore, is that we can successfully measure an important aspect of well-being (i.e. changes in local well-being) through measuring happiness.

## Chapter 1

### Measuring Well-being Through Measuring Happiness

#### Abstract

In this chapter, I will argue that we have good reason to suppose that happiness tends to be strongly correlated with well-being. This is shown by the fact that happiness tends to be used as a proxy for well-being: it is used by empirical researchers as a formal proxy in the measurement of well-being; and it is used as an informal proxy by laypersons in practical deliberation and evaluation. This potential correlation provides us with reason to investigate how successfully we can measure well-being through measuring happiness.

In order to investigate the relationship between happiness and well-being, we need a clear understanding of the nature of both well-being and happiness. I will claim that a person's well-being is how well that person's life is going for them, and that this notion is distinct from both a happy life and a good life. I then claim that a person's happiness is a broad, relatively long-term positive psychological condition.

In the remainder of the chapter, I outline a *theory-neutral account* of well-being and an *affective state account* of happiness. A theory-neutral account of well-being provides us with a broadly informative understanding of well-being that is consistent with most substantive theories of well-being. An affective state account of happiness is the view that happiness is constituted by a person's affective state. According to the version of this view that I will defend, a person's level of happiness consists in their balance of positive over negative affective states over a relatively long period of time.

The remainder of the thesis is an investigation into how successfully we can measure well-being (thus understood according to the theory-neutral account of well-being) through measuring happiness (thus understood according to the affective state account of happiness). This thesis can also be viewed as a philosophical assessment of empirical research that aims to measure well-being through measuring happiness in this respect.

## 1.1 Introduction

We care about well-being and its promotion. When we question what makes someone's life go well for them, we are asking questions about their well-being (or their "welfare"<sup>1</sup>). If I am successful in my career, will that make me better off overall? How good is my friend's long-term relationship for them? What can we do to improve the lives of those who are worse off? These kinds of questions are difficult to answer. Nonetheless, we often ask such questions and formulate answers in response: "A life of well-being involves doing something you are passionate about, having caring relationships, appreciating beauty" and so on. The measurement of well-being promises to provide us with answers to questions about well-being that are not readily accessible to us. What things tend to be the key determinants of well-being? Do rich people tend to have higher levels of well-being than poor people? What are the benefits of well-being? We may be able to roughly answer these questions intuitively, from our own experiences, or rationally, from knowledge of human nature, societal conditions, and so on. However, without having a way of measuring people's well-being, we cannot accurately answer such questions. Accurate answers to questions about well-being are the promise of well-being measurement.

The measurement of well-being has perhaps never received so much attention as it does today (Wren-Lewis, forthcoming). Well-being has always been an important part of the rationale for economics, health policy, law, etc. However, more recently, it has become an object of study independently of any single application to which it may be relevant. The study of well-being consists in a broad range of empirical research, from psychology, neuroscience and biology to economics, sociology and political science. These researchers measure well-being in a number of different ways, and have come up with several interesting and important findings over the past few decades.<sup>2</sup>

In addition, the study of well-being has recently become the focus of public policy (Wren-Lewis, 2013). Policy makers around the world have expressed an interest in "National Accounts of Well-being" (Kahneman et al. 2004). For

---

<sup>1</sup> I will treat the terms "well-being" and "welfare" synonymously throughout this thesis.

<sup>2</sup> I will review a number of these findings in chapter four.

instance, in 2010, the British government declared its intention to monitor the nation's progress through measuring the well-being of its citizens. Findings from the UK's National Well-being Programme are used to assess and develop policy (Dolan et al. 2011). Other governments and international organisations are considering similar measures (Helliwell et al. 2012), or have already implemented them (Stiglitz et al. 2009). These institutions claim to measure well-being in an attempt to "measure what matters".

In this thesis I will look at one kind of well-being measure in particular, namely measures of *happiness*. Measuring happiness is an increasingly prominent way in which both social scientists and public policy practitioners aim to measure well-being. This thesis aims to investigate the philosophical foundations of this practice. What are well-being and happiness? What is the relationship between the two notions? To what extent can we measure well-being through measuring happiness? In this chapter, I will argue that we have good reason to suppose that there tends to be a strong correlation between happiness and well-being. This potential correlation provides us with reason to investigate how successfully we can measure well-being through measuring happiness.

## **1.2 Why measure well-being through measuring happiness?**

In this section, I will provide three lines of evidence that suggest that happiness tends to be strongly correlated with well-being.<sup>3</sup>

### **1.2.1 Why measure well-being through measuring something else?**

Before outlining those lines of evidence, however, it is worth asking why we would even think of measuring happiness in the area of well-being measurement. Why not measure well-being directly? If you want to measure how many books you have bought, for instance, you would probably start by counting each of the books on your bookshelf. You would not try to measure how many books you have bought through measuring something else, such as the amount of knowledge you have, or your level of educational attainment. It makes sense to directly measure the amount of books you have bought. So

---

<sup>3</sup> I will discuss what happiness and well-being are later in the chapter. For now, I will assume that each line of evidence concerns similar notions of happiness and well-being.

why not do the same for measuring well-being? Why measure well-being *through* measuring happiness?

There are two answers to these questions. The first reason to think about the indirect measurement of well-being is quite simple: it is not obvious what well-being consists in. We may, for example, think that achievement partly constitutes well-being. But do all kinds of achievement constitute well-being, such as achieving goals that are unskillful or immoral? Does a particular achievement constitute well-being if we are not aware of it, such as the posthumous achievement of one of our goals? And what if we fail to value our own achievements – do they still constitute part of our well-being? These kinds of questions cast doubt on our knowledge of the constituents of well-being. There is no widely accepted theory of well-being amongst well-being theorists. Any current theory of what well-being consists in is controversial (Keller 2009; Heathwood 2010).

If we don't know what well-being consists in, we cannot directly measure well-being. However, even if we did know the constituents of well-being, we may not be able to directly measure those constituents. This is the second reason to think about the indirect measurement of well-being. Suppose, for example, that well-being is largely constituted by a person's valued achievements. In order to measure a person's well-being, therefore, we need to measure their level of valued achievement. How do we measure a person's valued achievements? This may be difficult to do accurately: subjects' actions may not reflect the goals that they most value; subjects may forget about certain kinds of achievements; and so on. It may be easier to measure a person's level of valued achievement *indirectly*, through measuring something else. What we would need to measure is something that is strongly *correlated* with a person's valued achievements. For example, subjects may experience pleasure each time they achieve something they value. If this were the case, we may be able to accurately measure a person's valued achievements through measuring their level of pleasure. No doubt things are likely to be much more complicated than this, but such a relationship would justify the indirect measurement of well-being – in this case, measuring well-being *through* measuring pleasure.

It is for these two reasons that we need to think about the indirect measurement of well-being. Firstly, we may not know what well-being consists in, in which case the direct measurement of well-being is impossible. Secondly, even if we did know the constituents of well-being, those



constituents may be hard to measure. In practice, the constituents of well-being may be most accurately measured through measuring things that are strongly correlated with those constituents.

### 1.2.2 Why measure well-being through measuring happiness?

In the previous section I contrasted two different ways of measuring well-being – directly and indirectly – and considered the virtues of both. In this section, I will show that we may be able to measure well-being directly and indirectly through measuring happiness.

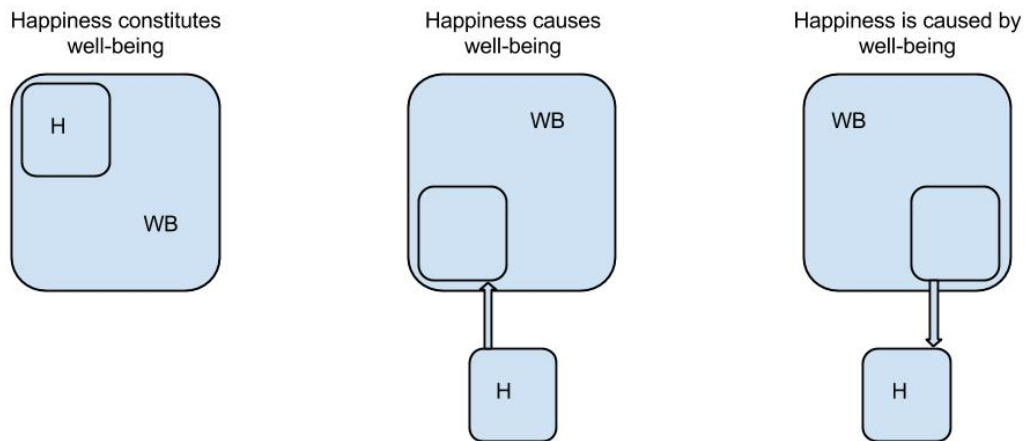
Firstly, we may be able to measure well-being directly through measuring happiness. This is because happiness may *constitute* well-being. Some theorists believe that happiness wholly constitutes well-being; others believe that happiness partly constitutes well-being. I will review the role of happiness in different theories of well-being in chapter two. The point, for now, is that a person's well-being may (at least partly) consist in their level of happiness. If this is the case, we can (at least partly) directly measure well-being through measuring happiness.

Secondly, we may be able to measure well-being indirectly through measuring happiness. This is because happiness may be correlated with the constituents of well-being in two important ways. First, happiness may *cause* the attainment of certain well-being constituents. For example, happy people may be more productive or sociable, which in turn may tend to lead to worthwhile achievements or relationships respectively. These things may (at least partly) determine one's level of well-being. Second, happiness may *be caused by* a person's well-being. For example, attaining certain things (such as a good job, a healthy lifestyle, etc.) may cause a person to be happy. If happiness were correlated with the constituents of well-being in either of these ways then we could indirectly measure well-being through measuring happiness.

The claim that happiness may be *correlated* with well-being can be justified either because happiness is constitutively or causally related to well-being, or both. That is, happiness may be correlated with well-being either because it constitutes well-being or because it causes or is caused by well-being (or both). This is illustrated by figure 1.1. The claim that happiness may be correlated with well-being is ambivalent over whether this is the result of a constitutive or causal relationship between the two things.

Fig 1.1.

The potential ways in which happiness may be strongly correlated with well-being (where the arrows represent the causal direction between happiness and well-being).



Do we have any reasons to think that happiness tends to be strongly correlated with well-being in any of these ways? In the next section, I will discuss three lines of evidence that suggest that happiness tends to be strongly correlated with well-being.

### 1.2.3 Why think that well-being and happiness tend to be strongly correlated?

So far, I have suggested that we may be able to successfully measure well-being through measuring happiness. Happiness and well-being may tend to be strongly correlated. This potential correlation can be explained by the fact that happiness may constitute, cause or be caused by (at least part of) well-being. Each of these relationships could result in happiness being strongly correlated with well-being. I will investigate each relationship in more detail in chapter two. For now, I will provide an initial case for thinking that happiness and well-being tend to be strongly correlated.

This case rests on three lines of evidence regarding the ways in which we tend to treat happiness as being strongly correlated with well-being. The primary line of evidence, for the purposes of this thesis, is that happiness tends to be used as a proxy for well-being in the formal measurement of well-being. This is often referred to as the “study of happiness” or the “study of subjective well-being”. A host of empirical researchers, from psychologists

and neuroscientists to economists and sociologists, aim to measure people's well-being through measuring their happiness. These researchers tend to treat happiness as a proxy for well-being. The fact that some people are happier than others is often taken to mean the former individuals are better off than the latter individuals (Graham 2010; Layard 2005). Similarly, the fact that various life circumstances, such as income, relationships, health, etc., are correlated with happiness is often taken to mean that those factors are good for us (Diener & Biswas-Diener 2008; Frey 2010). This is reflected in the fact that several prominent researchers have pushed for "National Accounts of Well-being" (Kahneman et al. 2004; Dolan et al. 2011; Diener et al. 2009; Bok 2010; Helliwell et al. 2012). The proposed national accounts consist in measures of happiness, in addition to standard measures of well-being, such as measures of Gross Domestic Product (GDP).

I believe that the study of happiness provides us with strong evidence to think that happiness may tend to be strongly correlated with well-being. Of course, it may *turn out* that such research is based on mistaken beliefs about the relationship between happiness and well-being. However, we do not have reason to presume this is the case from the outset. A large body of empirical research that tends to treat happiness as a formal proxy for well-being may be sufficient evidence of a potential strong correlation between happiness and well-being. This evidence alone may provide us with reason to investigate how successfully we can measure well-being through measuring happiness.

One could object, at this point, that such research may rest on the assumption that well-being is wholly constituted by happiness. For example, when researchers conclude that people with good relationships are happy, perhaps they are assuming that such people are thereby well off. Although this assumption is certainly made by many happiness researchers (Angner 2010), I think it would be a mistake to think that the entire study of happiness is based on this assumption. Such assumptions go against the primary methodology of empirical psychology. According to the psychometric approach, measures are justified on the basis of their correlations with other related measures, and not on the basis of theoretical assumptions (Alexandrova 2012b; Alexandrova 2012a). For example, measures of happiness are often justified on the basis of being strongly correlated with other measures of well-being, such as certain resources and capabilities associated with well-being (e.g. income, health, education, etc.) (Oswald & Wu 2010). Of course, this is not to say that measures of happiness are not influenced by theoretical assumptions. The assumption that happiness

constitutes well-being may have motivated the development of happiness studies. But such assumptions do not, ultimately, justify the practice. It is best to think of empirical research as being effectively theory-neutral when it comes to measuring well-being.

Nonetheless, even if one does doubt that the study of happiness provides us with sufficient evidence that happiness and well-being may tend to be strongly correlated, there are additional lines of evidence that suggest happiness and well-being are related in this way. I will consider two further lines of evidence in the remainder of this section.

The first line of evidence is that happiness tends to be used as an informal proxy for well-being by laypersons in practical deliberation and evaluation. In our practical evaluation, we often take someone's level of happiness as a rough indicator of their level of well-being: knowing that someone is happy makes us think that they are doing well; knowing that someone is unhappy makes us think the opposite. When we want to find out how well someone is doing in an important respect, we often appeal to his or her happiness. Thus, concerned parents inquire as to whether their child is happy or unhappy at school. We ask friends whether or not they are happy in their work or in their relationships. In general, if asked how we are doing, we frequently reply by noting how happy or unhappy we are.

In our practical deliberation, we often appeal to considerations of happiness with regards to important life decisions. Someone trying to decide what kind of career to have, for instance, will often ask which option would be best with regards to their happiness: Will I be happy as a teacher or a lawyer? Similar questions arise about other important choices – whether to get married or to have children, whether to live in the city or the countryside, etc. Policy-makers likewise consider the impact of proposed policies on the happiness of their constituencies.<sup>4</sup>

Note that we often use happiness as a proxy for well-being even when we do not explicitly refer to it as such (Haybron 2003). We may often use the term “happiness” when considering a person's *general* well-being. In contrast, we

---

<sup>4</sup> It is important to note that, in such contexts, people are not merely referring to well-being as “happiness.” If laypersons merely treated happiness as synonymous with well-being, the relationship between happiness and well-being would be an uninteresting one – the terms “happiness” and “well-being” would simply be two ways of referring to the same thing.

may sometimes refer to more specific forms of happiness when considering *particular aspects* of a person's well-being. For example, to say that someone is depressed may be a way of saying that they are unhappy in a particular sense. Likewise, we often claim that we have been "stressed out" over the past week, "feeling anxious" recently, or had a consistent "sense of fulfillment." Indeed, it may well be that *most* talk about happiness does not use the more general terms of "happy" or "unhappy" (Haybron, 2003). This makes sense: it is, after all, more informative to talk of these more specific psychological states than a blank assertion of unhappiness. But the broader category of interest here is nonetheless that of happiness; as Dan Haybron notes: "the report of being depressed, for example, could just as well have been an answer to an explicit query about how unhappy one is." (Haybron, 2003: 315)

The fact that we use happiness as a proxy for well-being in our practical lives is telling. It shows that one of the main ways in which we informally measure well-being is through keeping track of happiness. In order to know how well someone is doing, we ask whether they are happy. Formally measuring well-being through measuring happiness is intuitively plausible for this reason – we frequently (albeit informally) measure well-being through measuring happiness in our practical lives.

The third line of evidence for a strong correlation between happiness and well-being is the philosophy of well-being. On most theories of well-being, happiness is viewed as an important constituent of well-being or something that is strongly caused by well-being. I will consider theories of well-being in more detail in chapter two. For now, it is worth mentioning how happiness features in the three main kinds of theories of well-being influentially outlined by Derek Parfit (1984). Parfit's taxonomy consists in hedonist, desire-satisfaction and objective list theories of well-being, which I will very briefly consider in turn.<sup>5</sup>

---

<sup>5</sup> For each theory, there are many complications, which I will not consider now, but leave until chapter two. For instance, there are many forms of each theory, as well as theories that are not included in Parfit's list (Woodard 2013). For example, *veridical* hedonism is the view that only true pleasures constitute well-being; *idealized* desire-satisfactionism is the view that only the satisfaction of informed and rational desires constitute well-being. Other views, such as Sumner's Authentic Happiness Theory or Haybron's Self-Fulfillment Theory, do not fit neatly into Parfit's taxonomy. In light of such omissions, I intend the above as merely a sketch of how happiness relates to well-being on most philosophical theories of well-being.

According to prudential hedonism, happiness entirely constitutes well-being. Other goods, such as achievement, knowledge, virtue, etc., have merely instrumental value in their ability to bring about happiness. Thus, there is a perfect correlation between happiness and well-being, according to hedonistic theories of well-being.

In contrast to hedonism, desire-satisfactionism is the view that the satisfaction of desire constitutes well-being. Happiness and well-being can come apart, according to desire-satisfactionism, in that we can be unaware of the satisfaction or frustration of our desires. However, it is reasonable to assume that, in general, the satisfaction of our desires often makes us happy. Happiness and desire-satisfaction often coincide. In addition, many people desire happiness. Thus, there is likely to be a strong correlation between happiness and well-being, according to desire-satisfactionism (Angner 2012).

Lastly, according to objective list theories, certain objective goods constitute well-being, such as pleasure, knowledge, achievement, virtue, etc. Happiness and well-being can come apart, according to objective list theories, in that certain goods constitute our well-being regardless of whether or not they make us happy. However, much as with desire-satisfactionism, it is reasonable to assume that the attainment of such goods often makes us happy. Happiness and the attainment of objective goods often coincide. In addition, happiness is often included as one of the objective goods on the list. Thus, happiness will tend to be strongly correlated with well-being, according to objective list theories.

It is perhaps not surprising that philosophical theories of well-being often entail that happiness is strongly correlated with well-being. We have already seen that happiness tends to be used as a proxy for well-being. Insofar as philosophical theories of well-being aim to develop a descriptively adequate account of well-being, therefore, we will tend to see theories that entail that happiness is strongly correlated with well-being.<sup>6</sup> A theory of well-being that

---

<sup>6</sup> On this, Sumner writes: “[T]he best theory about the nature of [well-being] is the one which is most faithful to our ordinary concept and our ordinary experience. That experience is given by what we think or feel or know about well-being, both our own and that of others. The data which a candidate theory must fit, therefore, consist of the prodigious variety of our preanalytic convictions.” (Sumner 1999: p.10-11; see also Tiberius 2004: p.299). In short, a theory that tells us about the nature of well-being must be “faithful” to our common sense judgments about well-being.

results in well-being being too divorced from happiness will fail to account for the ways in which we tend to treat happiness as a proxy for well-being.

In sum, I have presented three lines of evidence that suggest there is a strong correlation between well-being and happiness. We tend to use happiness as a proxy for well-being, on both a formal and informal basis. Formally, empirical researchers use happiness as a proxy for well-being in the measurement of well-being. Informally, laypersons use happiness as proxy for well-being in their practical deliberation and evaluation. This potential correlation provides us with reason to investigate how successfully we can measure well-being through measuring happiness.

#### **1.2.4 Assessing the evidence**

One might think that the method of analysis used in the previous section is problematic. I outlined ways in which empirical researchers, laypersons and well-being theorists tend to treat happiness as a proxy of well-being. Yet, we cannot be sure that, when social scientists, the folk and philosophers use the terms “happiness” and “well-being”, they are all referring to the same things. We need to be clear on what well-being and happiness are in order to further assess the relationship between them.<sup>7</sup>

In the remainder of this chapter, I will consider the nature of well-being and happiness on two different levels: the *concepts* of well-being and happiness, and the *conceptions* of well-being and happiness.

The *concepts* of well-being and happiness concern the broad phenomena in question. What is well-being, for instance? We may disagree over what well-being consists in, yet agree over what well-being is, on a broader level. For example, we may agree that well-being concerns a life that is good for someone. However, you might think that such a life consists in a favourable balance of pleasure over pain. In contrast, I might think a life that is good for someone consists in the achievement of certain goods, such as virtue, knowledge, intimate relationships, etc. We are not disagreeing over the

---

<sup>7</sup> One might worry that, if the folk, social scientists and philosophers all hold different notions of happiness and well-being, the three lines of evidence outlined above are invalid. This may turn out to be the case, but I do not think it makes sense to assume that it is so. In the analysis of the concepts and conceptions of well-being and happiness below, I will argue that there are concepts and conceptions of both things that can account for the evidence outlined above.

concept of well-being; we agree that well-being concerns a life that is good for someone. What we are disagreeing over is what well-being consists in.

The *conceptions* of well-being and happiness concern what these things consist in. Different conceptions of well-being, for instance, concern different views of the constituents of well-being – what well-being consists in. The three different kinds of philosophical theories of well-being outlined above are all different kinds of conceptions of well-being.

In the next two sections of this chapter, I will consider the concepts of well-being and happiness respectively. I will argue the following: Well-being is a life that is good *for* someone; such a life is conceptually distinct from both a happy life and a good life. Happiness is a broad, relatively long-term psychological condition. I will show that these concepts adequately account for the fact that we tend to use happiness as a proxy for well-being.

In the final two sections of this chapter, I will consider different conceptions of happiness and well-being respectively. I will not argue in favour of a particular substantive theory/conception of well-being. Instead, I will outline a theory-neutral account of well-being. This account rests on the causal connections between various “well-being ingredients”. With regards to happiness, I will argue in favour of an affective state account. According to this view, happiness is a person’s affective state. I will show that the affective state account adequately accounts for the fact that we tend to use happiness as a proxy for well-being.

### **1.3 Well-being: Concept(s)**

In this section, I will consider the concept of well-being. What is well-being? Why do we care about it? I will show that a person’s well-being is how well that person’s life is going for them, and that this notion is distinct from both a happy life and a good life. I will suggest that there are a number of reasons why we might care about well-being, and that these reasons justify the role that well-being plays in our practical lives.

#### **1.3.1 What is well-being?**

Most people consider their own and other’s well-being to be valuable. For some, well-being is the only thing of final value. We care about well-being and its promotion. The promotion of well-being is one of the primary goals of individuals, caregivers, developmental charities and organisations, and public



policy. It is for this reason that the “science of well-being” has recently become an area of study in its own right.

In what way do people care about well-being and its promotion? On a personal level, our own well-being is what we aim to achieve when we are being “self-interested,” “looking out for ourselves,” “wanting to get something from a situation,” and concerned to know “what’s in it for us” (Campbell forthcoming). The notion of well-being refers to how good a life is (or how well a life is going) *for the subject whose life it is*. When we aim to benefit our lives in some way, we are aiming to promote our well-being. As Tim Scanlon states, “well-being serves as an important basis for the decisions of a single rational individual, at least for those decisions in which he or she alone is concerned.” (Scanlon 1998, p.108)<sup>8</sup> When we think about whether it would be better for us to have a different career, get married, or make some other major life change, we consider our own well-being. Well-being is one of the main goals of individuals who care about how well their lives are going *for themselves*.

These self-interested motivations can be contrasted with beneficent behaviour, which aims at achieving what is good *for someone else*. When we care about someone, we want that person’s life to go well for him or her – we care about his or her well-being. As Scanlon states, “well-being is what a concerned benefactor, such as a friend or parent, has reason to promote.” (Scanlon, 1998: p.108) When we consider how to respond to the needs or wants of another person (a dependent, spouse, or friend) for his or her sake, we are considering the well-being of others.<sup>9</sup> Well-being is one of the main goals of caregivers and of social policy for governments and developmental organisations.

---

<sup>8</sup> This is not to say that we necessarily care about our well-being, or that the promotion of our well-being provides us with certain reasons to act. For example, Scanlon explains that, “If you ask me why I listen to music, I may reply that I do so because I enjoy it. If you asked me why that is a reason, the reply “A life that includes enjoyment is a better life” would not be *false*, but it would be rather strange.” (Scanlon, 1998: p.126) The things we care about (such as enjoyments, success in one’s main aims, and substantive goods such as friendship) are not desirable because they promote our well-being. They may end up promoting our well-being, but we do not tend to justify their desirability on this basis. Thus, Scanlon labels well-being an “inclusive good” – one that is made up of other things that are good in their own right, not made good by their contributions to it (Scanlon, 1998: p.127).

<sup>9</sup> Indeed, Darwall (2004) argues that a person’s well-being consists in what a benefactor would rationally want for them.

It is worth noting that, in our self- and other-interested behaviour, we often care about promoting different aspects of well-being. We can distinguish between different aspects of well-being on the basis of (a) its temporal dimension, and (b) its contextual dimension. Consider, first, the different aspects of well-being based on the period of time in question. We can distinguish between *synchronic well-being* (how well one's life is going at a particular time) and *diachronic well-being* (how well one's life is going over a certain period of time). Some well-being theorists further distinguish between these two aspects of well-being and *lifetime well-being* (how well one's life is going as a whole) (Velleman 1991; Raz 2004). There are also different aspects of well-being based on the context in question. We can distinguish between different aspects of well-being considered in different practical contexts, such as *mental well-being*, *physical well-being*, *childhood well-being*, *material well-being*, *emotional well-being*, *spiritual well-being*, and so on (Alexandrova 2013). These distinctions are important insofar as our self-interested and other-interested decisions tend to be confined to certain periods of time and contexts. We often think about what is good for us, or someone else, in a specific practical context or period of time. For example, within a medical context, it may make more sense to consider a patient's physical well-being, rather than their overall well-being. As another example, when you ask your friend, who you see roughly once a week, how they are doing, you may only be asking about their well-being over the past week or so. In general, the aspect of well-being that we care about is determined by the relevant period of time or context.

A life of well-being (i.e., a life that is good for the subject whose life it is) can be distinguished from a range of other notions of a good life. A life can go well from an impersonal perspective, such that its existence contributes to the value *simpliciter* of the world. A life can go well from a moral, aesthetic, perfectionist or religious perspective, such that its existence promotes or exemplifies moral, aesthetic, perfectionist or religious values respectively. For instance, a life can go well from a perfectionist perspective, such that its existence exemplifies the excellences characteristic of one's nature. As Wayne Sumner notes, "Lives...are complex things whose value can be assessed along a number of different dimensions or from a number of different standpoints. Welfare represents only one of these dimensions." (Sumner, 1996: p.20)

These other notions of a good life are conceptually distinct from a life that is good *for* someone or something. The life of Mother Teresa, for example, was

undoubtedly a morally good life, but not necessarily a life that was good *for her*. Her life was full of depression and loneliness. This may have enhanced the moral worth of her life, but casts doubt on whether it was good for her, which shows that the two notions are distinct. Of course, it may be that, according to the correct theory of well-being, a morally good life will coincide with (or even *ipso facto* be) a life that is good for the person whose life it is. However, even if a life that is good for someone and a life that is morally good are bound to coincide *substantively*, this does not mean that the two concepts are not distinct. This is shown by the fact that suffering from depression and loneliness casts doubt on whether Mother Teresa had a life that was good for her, but not on whether her life was a morally good one. This would not be possible if the two notions were not conceptually distinct.

What makes a life go well for the subject whose life it is? Well-being theorists refer to the kind of things that contribute towards our well-being as having *prudential value*. The extent to which our life is good for us, therefore, is determined by the prudential value of the things that constitute our life. Prudential value can be distinguished from other kinds of value (such as value *simpliciter*, moral, aesthetic, religious or perfectionist value) in virtue of its *subject-relativity*, or *value for* a certain subject (Griffin 1989; Sumner 1996; Kraut 2009). A certain state of affairs could have prudential value for me while having no such value for you (Taylor 2012). This may not be the case for other kinds of value. For example, a particular action may be morally bad in general, not morally bad for me but morally good for you.

The subject-relativity of prudential value does not entail a subjective account of prudential value, whereby what is good for me is at least partly determined by my attitudes. For example, it may be that friendship is good for someone regardless of his or her attitude towards friendship. Yet, although friendship is good for that person, it may not be good for someone else.<sup>10</sup> Features of a subject ultimately determine whether something is good for the subject, but these features need not be the subject's attitudes.<sup>11</sup>

---

<sup>10</sup> Perhaps because it does not suit someone else's individual *nature*, which may be partly determined by features of the subject other than their attitudes (Haybron 2008a).

<sup>11</sup> Indeed, the relevant features may not even be particular features of the individual; rather, they may be features that the subject shares with other members of their species (Hursthouse 1999; Foot 2001).

Well-being and prudential value are rarely distinguished in the philosophical literature. Indeed, they are often regarded as synonyms (Crisp 2008). Yet, there is a clear distinction between the two notions that can be made, which I will simply stipulate as follows: the notion of well-being describes *how well a life is going* for the subject whose life it is; the things that *contribute towards a life going well* for the subject whose life it is have prudential value. Something is good for you (i.e., has prudential value) if and only if it contributes towards your level of well-being, and something is bad for you (i.e., has prudential disvalue) if and only if it detracts from your level of well-being.

We now have a grasp of what well-being is and the way in which we care about its promotion. However, the notion of well-being has recently been criticized. Some theorists argue that the notion of well-being is a fragmented one – that there is no unified concept of well-being. I will consider these arguments in the next section, and outline three different concepts of well-being (what I will refer to as the Wide, Standard and Narrow Concepts of well-being). In the following section, I will argue that one of these concepts (the Standard Concept) refers to a valuable kind of life that is distinct from both a good life and a happy life. This is the kind of well-being that I will consider in this thesis. Thus, even if well-being is a fragmented concept, the notion of well-being that I will consider in this thesis refers to something that we should care about, and thereby plays the prominent role in our practical lives outlined in this section.

### **1.3.2 Well-being as a fragmented concept**

Recently, there have been several arguments that there is no unified concept of well-being (Griffin 2007; 2000; Raz 2004; 1988; Scanlon 1998). For instance, after listing the three important roles that well-being plays in our practical lives, Scanlon claims that, “it is a mistake to think that there is a single notion of well-being that plays all the roles I have mentioned and that we need a theory of well-being to clarify this concept.” (Scanlon 1998: p.108) In a similar manner, James Griffin gestures that there may be “several different fundamental notions” of well-being (Griffin 2000: p.285).

Of course, it may be that the arguments in favour of the fragmentation of well-being turn out to carry little weight. Michael Bishop, for instance, suggests that the case for fragmentation may simply rest on the fact that people use the expression of “well-being” in diverse ways (Bishop MS: p.14). Although interesting, this linguistic thesis does not necessarily tell us anything about the nature of well-being. If Bishop is correct, I will assume that there is no

problem with the analysis of well-being presented above: a person's well-being is a life that is good for that person; it is because we care about leading lives that are good for us that well-being plays a prominent role in our practical deliberation and evaluation.

However, if Bishop is wrong, and well-being does turn out to be a fragmented category, we need to distinguish between different concepts of well-being and consider which concept(s) we should care about.

We can gain an understanding of potentially different concepts of well-being by considering a thought-experiment called the "Crib Test" (Feldman 2004; Braddock 2010; Kim MS). The Crib Test asks you to imagine you are a loving parent of a newborn baby. You want your baby be well off, and thereby ask yourself the following kind of question: "Would your desire for your baby to be well off be satisfied if your baby grew up to live the life of [someone who lacks a particular good, such as virtue, achievement, knowledge, etc.]" The Crib Test is considered by some theorists to be a reliable method of forming correct intuitions about well-being because parents love their children in the well-being-promoting sense (Darwall 2004). Parents tend to unconditionally care for the existence of their children, and for what is good for their children.

However, not all theorists agree that the Crib Test captures our intuitions concerning well-being. Feldman argues that the Crib Test may end up reflecting our judgements about what is good (in some sense), and not what is good *for someone* (Feldman 2004; 2010). He presents the following example to illustrate this point:

"Suppose a religious fanatic looks into his child's crib. Suppose he wants the child to have a wonderful life. Suppose he thinks that the best imaginable life for the child is one in which the child becomes a martyr for God. This religious fanatic might be filled with love, and he might be thinking about the Good Life for his child. But it is not clear that he is expressing a hope about what we would normally think of as the child's welfare. Perhaps he is thinking about what he takes to be a moral or religious virtue. Perhaps he is thinking about the most beneficial life the child could live. So the mere fact that he is a parent filled with love, and is looking into his child's crib, and is saying something about 'the Good Life', does not absolutely guarantee that he is thinking about well-being." (Feldman 2004: p.10)

Feldman's argument rests on the fact that well-being concerns a value (i.e. prudential value) that is conceptually distinct from other kinds of value (such

as aesthetic value, moral value, religious value, etc.). As explained in the previous section, well-being concerns what is good *for* someone, not what is good in some other kind of respect. One way of putting this point is that the term “well-being” seems to refer to a narrow, non-moral concept. We may want more for our children than for them to have a life of well-being. We may want them to have lives full of meaning and worth. This would be a good life, no doubt, but not necessarily a life that is good *for them* (Smuts MS).

In contrast to this position, Aristotelians tend to understand “well-being” as referring to a concept closer to *eudaimonia*, a concept that embraces a wider range of values than the narrower concept, and allows for the conceptual possibility of moral value to be one of its constitutive elements (Haybron, 2008: chapter two). According to this wider concept of well-being, what you want for your children may be a number of things, such as lives of happiness, meaning and worth. These things all contribute towards your children’s well-being, in that they are in your child’s wider interests. Call this the Wide Concept of well-being (WC). WC refers to the realisation and integration of all of the positive values that can enrich a person’s life. Haybron refers to this concept as “A Good Life”: “a life that is good in *all* respects in which a life can be considered good, including morally – a life that is desirable without qualification, both enviable *and* admirable.” (Haybron 2000: p.211) Under a wider understanding of well-being, “a good life” may simply be referred to as “a life of well-being”.

The Crib test may elicit judgments about the Wide Concept of well-being or about narrower concepts of well-being. We can further distinguish between narrower concepts of well-being. Kagan (1992) argues that the concept of well-being is much narrower in scope that people tend to think. He provides the following argument for thinking that the concept of well-being refers merely to our experiential lives: Something contributes towards a person’s well-being if it is good *for that person*. Persons are nothing other than a body and mind. Thus, according to Kagan, something can only constitute a person’s well-being if it makes a difference to their body or mind. Being genuinely successful, for instance, does not benefit *the person* except insofar as it directly impacts them i.e. through their *experiences* of being successful. Let us call this the Narrow Concept of well-being (NC).

NC concerns only a person’s body and mind, which, according to Kagan, are the intrinsic properties of a person. It does not concern so-called ‘extrinsic’ properties of a person, such as various relational properties (e.g. being in a

certain environment, having certain possessions, attributes, friendships, etc.). Kagan argues that relational properties may make a person's *life* go well, but do not necessarily benefit *the person*. To benefit a person, according to Kagan, it must make a difference in the person i.e. it must make a difference to the person's body or mind. Being genuinely successful may constitute a person's life going well, but does not benefit the person except insofar as it directly impacts them.

If Kagan is right, there may be two different concepts of well-being: one that refers to how well a *person* is doing (i.e. the Narrow Concept) and another that refers to how well a person's *life* is going. How well a person's life is going may include relational properties, such as whether the person is genuinely successful or not. Let us call this concept the Standard Concept of well-being (SC). In contrast to NC, it may include a person's relational properties (e.g. being in a certain environment, having certain possessions, attributes, friendships, etc.). In contrast to WC, it does not necessarily include other kinds of positive values that can enrich a person's life, such as moral, aesthetic, perfectionist or religious values. Both NC and SC are narrower concepts of well-being than WC, but, unlike NC, the Standard Concept of well-being does not reduce what is good for someone to their body and mind.

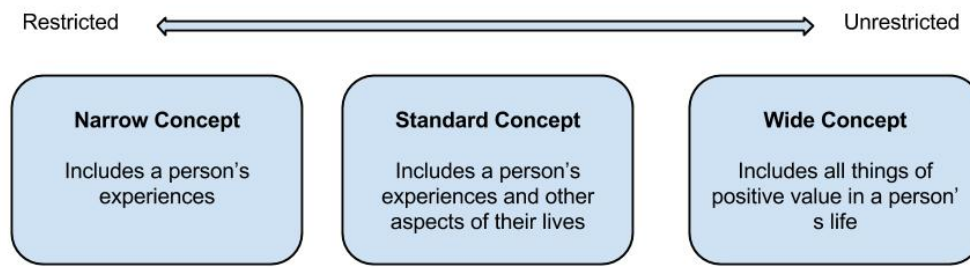
Some well-being theorists have argued that NC and SC essentially refer to a person's well-being over different time-frames. David Velleman (1991), for instance, argues that there are two different kinds of well-being: a person's momentary well-being and a person's lifetime well-being. According to Velleman, a person's lifetime well-being is not simply the lifetime sum of their momentary well-being. Velleman argues that a person's lifetime well-being at least partly consists in the "narrative shape", or "meaningfulness", of a person's life. A person whose life is characterised by a plurality of mutually supporting goals and chapters seems to have a higher level of lifetime well-being than a person whose life is characterised by radically disconnected chapters despite consisting in the same momentary levels of well-being (Kauppinen 2012). It is possible that NC may refer to a person's momentary well-being, whereas SC may refer to a person's lifetime well-being.<sup>12</sup>

These three different concepts of well-being are illustrated in fig 1.2, ranging from highly restricted to unrestricted concepts of well-being:

---

<sup>12</sup> I will reconsider the distinction between a person's momentary well-being and their lifetime well-being in chapter five.

Fig 1.2



In sum, there may at least three different concepts of well-being, which can be categorised in virtue of how restricted or unrestricted they take well-being to be. The most restricted concept (NC) relates to the notion of a happy life – a life of enjoyment, engagement, satisfaction, fulfilment, contentment, etc. In contrast, the least restricted concept (WC) relates to the notion of a good life – a life of moral virtue, meaningful relationships and activities, aesthetic beauty, spiritual depth, and so on.

### 1.3.3 Well-being as a distinct concept

In response to the analysis of the previous section, one could argue that there is no distinct notion of well-being that refers to something we should care about. That is, one could be a *nihilist* about well-being. Nihilists could argue that well-being is a confused concept, which is caught between the concept of a good life and a happy life. The notion of well-being may be a hybrid, “an attempt to find a concept which is half one and half the other” (Raz, 2004: p.270). According to such thinking, we should perhaps discard the notion of well-being, talk simply about the good life and the happy life, and do well not to muddy the waters in-between.

In this section, I will argue that such nihilism would be going too far. It assumes that the notion of well-being does not do any additional theoretical work, beyond taking onboard some parts of the concepts of a good life and a happy life. This is not the case.

The distinguishing feature of well-being and prudential value is that they are both *subject-relative*. There are a number of reasons why we might care about a life that is good in a subject-relative sense. I will briefly note four ways in which a life of well-being may matter in a way that is distinct from both a good life and a happy life.



Firstly, something may be good *for a subject* (i.e. good in a subject-relative sense) because it is related in some way to the *value of the subject*. For instance, Connie Rosati (2006) claims that, because a person is valuable, their well-being consists in those things that *sustain their own value*. This clearly includes meeting one's basic needs for survival (e.g. water, food, shelter, etc.), keeping in good condition and exercising one's valuable capacities and abilities (e.g. being able to form attachments with others, make meaningful contributions to the world, etc.). Similarly, Joseph Raz (2004) claims that, as autonomous agents, a life that is good for us is one in which we appreciate and respond appropriately to things of value. According to Raz, we can appreciate things of value either by engaging with them (e.g. enjoying a good painting, spending time with a friend, etc.) or respecting them (e.g. restoring the good painting, protecting the friend from harm, etc.). It is through appreciating and responding appropriately to things of value, Raz claims, that we promote our own value (as autonomous agents).

Secondly, something may be good *for a subject* because it is related in some way to the *success of the subject*. This is similar to the previous reason, but does not require a view about what makes a subject valuable. Rather it simply views a subject as having certain goals with regards to which it can be more or less successful. For instance, Dan Haybron (2008) notes that organisms can be viewed as having goals of some kind; these goals may take the form of psychological attitudes, such as aims or desires, or may be constituted by tendencies for growth and development. According to Haybron, we promote people's well-being by "helping", "aiding" or "assisting" their success at achieving certain kinds of goals. In contrast, the good life may not be entirely concerned with the goals of the subject. As Haybron claims, one may partly achieve a good life "merely by fulfilling a capacity, even if one hasn't the slightest desire for it, could not be brought to desire it, is in no other way orientated to seek it, and even if one responds with nothing but pain and revulsion towards it." (Haybron, 2008: 169) We may not be able to say the same about achieving a life of well-being.<sup>13</sup>

---

<sup>13</sup> This may relate to Stephen Darwall's "rational care theory" of the concept of welfare (Darwall, 2004). According to Darwall, the concept of welfare is the concept of what it is rational to want for someone insofar as one cares for her. He argues that well-being ultimately concerns the appropriate objects of *sympathetic concern*. It seems reasonable to presume that we should have sympathetic concern towards others with regards to how well they are doing at achieving their goals.

Thirdly, something may be good *for a subject* because it is related in some way to the temporal and physical *limits of the subject*. Again, this is related to the previous reason, but consists in a more nuanced view of how success relates towards well-being. Cheshire Calhoun (MS) has outlined such a view with regards to a meaningful life – a concept that is intertwined with the notion of well-being (Raz, 2004). Calhoun argues that one can achieve a good life through succeeding in one or more of one’s important goals (such as Mother Teresa achieving a good life through succeeding in helping the poor). In contrast, one may only achieve a life of well-being through succeeding in *all* of one’s important goals relative to one’s available time and resources. Another way of putting this point is that well-being may concern how a subject distributes their attention and resources among certain goals – a life of well-being may be a successful life within the bounds of our practical lives.<sup>14 15</sup>

Lastly, something may be good *for a subject* because it can be *justified to the subject* as being good. For instance, Valerie Tiberius (2007) claims that we care about well-being because we care about a life that can be justified as a good life to the person whose life it is. As Tiberius puts it: “To act for your sake, as opposed to acting for morality’s sake or the sake of another person, is to act in a way that I can justify to you in some sense” (Tiberius 2007: p.375) Your life may be justifiably good insofar as it has intrinsic value (e.g. it may have aesthetic value) but this does not count as justifying it as good to a person whom we are trying to persuade of its value. Tiberius argues that,

---

<sup>14</sup> Haybron comes to a similar conclusion in remarking that well-being, “apparently relates to the universal problem of deciding how to distribute resources and attention among those we care about including ourselves. When does a given individual require assistance or special care? Who needs it most (and least)? Who has more than they need? Who has given up the least, or most?” (Haybron, 2008b: 170)

<sup>15</sup> This also supports two further intuitions about well-being. Firstly, it seems that a life of well-being is, to a certain extent, a balanced life. A balanced life seems to require some meaningful relationships such as friendships, some measure of success or achievement in worthwhile projects, and some amount of happiness and (reflective) satisfaction (Tiberius 2005a). A life that does not have some of these things seems to fail as succeeding in certain important goals.

Secondly, it seems that the prudential perspective is the most comprehensive perspective from which to judge a life. Consider, for example, Raz’s (2004: 294) example of an ocean-going sailor who judges himself only by his record-breaking attempts. The sailor has friends and family, and other interests, and they too contribute towards his well-being (even if less than his sailing). Sailing matters to him most. But well-being is the most comprehensive perspective on his life, taking account of everything of importance (sailing, friends and family, and other interests).

with regards to well-being, the person who must be persuaded of the value of a life is the person whose life it is. They must be able to accept or endorse their life if they were to follow a certain procedure (that does not itself consist in particular prudential values).

I am unable to consider these four theories in more detail here. I will assume that at least one of them (or a similar kind of theory) is correct, and therefore that a life of well-being matters in a way that is distinct from both a good life and a happy life.

The Standard Concept of well-being (SC) outlined in the previous section clearly best captures the distinct nature of well-being. In contrast to the Wide Concept of well-being (WC) it does not equate a life of well-being with a good life. And in contrast to the Narrow Concept of well-being (NC) it does not equate a life of well-being with a happy life.

In common with WC, SC can account for the important role that well-being plays in our practical lives. According to SC, our well-being may be constituted by the realisation and integration of many of the positive values that can enrich our life. This is why many of our major life decisions concern our well-being.

In addition to WC, SC can account for situations in which we sacrifice our well-being for other goods that enrich our life. For example, it seems possible (if not probable) for a single mother of unexpected triplets to sacrifice her well-being for more important kinds of positive value such as the moral value of caring for her children. These kinds of situations seem to involve trade-offs between our well-being and other kinds of positive value that enrich our life. The single mother sacrifices her well-being for virtue.<sup>16</sup>

In general, SC can account for the fact that well-being is worth caring about, but that it may not be the only thing worth caring about. As Haybron and Tiberius put it: "it is perfectly ordinary for individuals to care about things they see as having little or no positive bearing on their well-being. Artists, social

---

<sup>16</sup> Indeed, cases of self-sacrifice show two important things about the role well-being plays in our practical lives. Firstly, the fact that a person can choose to sacrifice their well-being for other kinds of positive value that enrich their life shows that these are two different things. Secondly, as Scanlon notes, the fact that it makes sense to make the opposite choice – namely, to not sacrifice one's well-being in favour of other kinds of goods – shows that well-being is an important factor in our major life decisions (Scanlon, 1998: p.131). How well our lives go *for us* may not be all there is to how well our lives go, but it is often the leading factor in our practical deliberation and evaluation.

workers, and dissidents, for instance, sometimes choose paths in life that will leave them, in their eyes, worse off than other options before them. Similarly, people care about things having no relation at all to their own lives, much less well-being: the future welfare of a stranger one briefly met, or the state of the world's ecosystems a thousand years hence. And some individuals may be depressed, detest themselves, or hold religious doctrines on which they actually value their own ill-being." (Haybron & Tiberius MS: p.12)

Lastly, in addition to NC, SC can account for the things we care about beyond our experiential lives. In our practical deliberation and evaluation, we care about the things that increase or decrease our well-being, such as success in our personal objectives, goals and projects. Again, we care about whether we are genuinely successful in these endeavours, not just whether we have certain experiences of success. Moreover, this does not seem to be a mistake. Thought-experiments such as Nozick's experience machine appear to support the idea that actual states of affairs may constitute our well-being beyond our experiences of them (Nozick 1974). I will discuss this issue in detail in chapter two.

In sum, the Standard Concept of well-being refers to a valuable kind of life that is distinct from both a happy life and a good life. In the remainder of this thesis, I will simply refer to it as "the concept of well-being".

## **1.4 Happiness: Concept(s)**

In the previous section on the concept(s) of well-being, I argued that there is a concept of well-being that refers to how well a person's life is going for them, and that this notion is distinct from both a happy life and a good life.

In this section, I will consider the concept(s) of happiness. For the purpose of this thesis, the following analysis of happiness will be much more straightforward than the previous analysis of well-being. This is because I am concerned with the measurement of well-being, and measuring well-being through measuring happiness in particular. In the first main section of this chapter, I showed that that happiness tends to be used as a formal proxy for well-being in the measurement of well-being. I claimed that this is our primary line of evidence in favour of a potential correlation between happiness and well-being. Now, the concept of happiness used by empirical researchers is quite clear. In short, researchers refer to happiness as a broad, relatively

long-term positive psychological condition. This is the concept of happiness that I will be concerned with in this thesis.

However, as mentioned in the first main section of this chapter, one may doubt that the study of happiness provides us with sufficient evidence to suppose that happiness and well-being tend to be strongly correlated. In response, I outlined two additional lines of evidence that suggest happiness and well-being are related in this way. In particular, I showed that happiness tends to be used as an informal proxy for well-being by laypersons in practical deliberation and evaluation.

It is not obvious that the folk use the same concept of happiness as the one used by empirical researchers. In this section, therefore, I will briefly consider a number of different concepts of happiness that may be used by laypersons in their practical lives. I will suggest that social scientists and the folk do tend to use the same concept of happiness. This concept of happiness can adequately account for the fact that happiness tends to be used as both a formal and informal proxy for well-being.

#### **1.4.1 The fragmentation of happiness**

The term “happiness” is ambiguous.<sup>17</sup> It is used to refer to a number of different things (Haybron, 2003). In this section, I will briefly outline eight potential referents of the term.<sup>18</sup> I will suggest that we only tend to use one of them as a proxy for well-being in our practical lives.

The first concept of happiness that I will consider is what Haybron refers to as “Prudential Happiness” (Haybron, 2003). Prudential Happiness is synonymous with well-being – it refers to a life that is good for the person whose life it is.<sup>19</sup> When we ascribe happiness to someone, in this sense, we are making an *evaluative* judgment about how well his or her life is going, namely whether it is going well for him or her. Thus, Haybron notes that people may disagree over whether someone is happy or not if those people have different prudential values: “I might think that Genghis Khan has a happy

---

<sup>17</sup> Though, see Feldman (2010), for an argument against the ambiguity of the term.

<sup>18</sup> It is worth noting that each notion of happiness has a converse notion of unhappiness. I will focus solely on notions of happiness here, and leave the reader to infer the notions of unhappiness.

<sup>19</sup> Recent notable works by theorists that use the term “happiness” in its prudential sense include: (Almeder 2000; Annas 1995; McMahan 2006; Noddings 2004; White 2006).

life, because I think what matters for well-being is getting what you want; while you may deny this because you think a life of evildoing, however “successful,” is sad and impoverished.” (Haybron 2011a)<sup>20</sup>

It seems clear that we do not tend to treat happiness synonymously with well-being in our practical lives; we tend to treat them as separate things. It is for this reason that we are able to use happiness as a proxy for well-being. We can see this in two ways. First, happiness is often used as a defeasible proxy for well-being. We say things such as, “Even though I was happy, I was not doing well”, “Knowing the truth is good for me, even if it doesn’t make me happy”, and so on. Second, we say things such as “I wish my child would be happy and healthy”, though we may not say things such as “I wish my child to have well-being and be healthy”. Being healthy may not be valuable beyond the well-being that it provides. However, being healthy may be valuable beyond the happiness that it provides. This shows that the two things – happiness and well-being – tend to be distinct.<sup>21</sup>

In contrast to Prudential Happiness, the term “happiness” is often used to refer to an entirely descriptive psychological state. I will briefly consider three concepts of happiness of this kind. The first descriptive concept of happiness refers to a particular episodic emotion. Happiness, in this sense, is a particular positive emotion, similar to other emotions such as joy and excitement. The emotion of happiness can be viewed as one of the garden-variety emotions, such as emotions of anger, disgust and fear (Ekman & Davidson 1994). Such emotional episodes are typically fleeting, and

---

<sup>20</sup> In the discussion of the concept of well-being above, we saw that there may be at least three different concepts of well-being: narrow, wide and standard senses of well-being. It follows, therefore, that there may be three different concepts of Prudential Happiness. Indeed, Haybron notes that happiness is sometimes used to refer to a life that is most choiceworthy – a life that is good overall, or all things considered (Haybron, 2003). More succinctly, we can refer to this kind of life as “the good life”. Haybron refers to this concept of happiness as “Perfectionist Happiness.” But, it is worth noting that, according to the wide sense of well-being discussed above, Perfectionist Happiness is synonymous with well-being. We might think that Genghis Kahn was well-off according to the standard sense of well-being, for example, but that he was not *truly* (or *really* or *genuinely*) well-off, according to the wide sense of well-being. Similarly, we might think that he had a happy life, but not a *truly* happy life.

<sup>21</sup> Indeed, the fact that the term “happiness” is sometimes used synonymously with the term “well-being” is further evidence that happiness (in its non-prudential sense) may be strongly correlated with well-being. That is, the two notions may be so closely bound up with each other that they are sometimes treated as the same thing.

correspond to a favourable thought, action, event or situation. For example, one might say that, “eating the ice-cream made me happy.”

Of course, some emotional episodes of happiness (and, indeed, emotional episodes in general) can last for much longer time periods – perhaps days, or weeks. For example, one might say that they have been feeling happy since their partner agreed to marry them last week. Nonetheless, emotional episodes are generally short-lived, often occurring in response to a particular event or situation that requires us to act in a certain way (Griffiths 1997). I will refer to this concept of happiness as Happiness<sup>Episodic Emotion</sup>.

The second descriptive concept of happiness that I will consider can be viewed as an extension of the previous concept. In contrast to viewing happiness as a particular positively valenced episodic emotion, we sometimes view happiness as all (or almost all) kinds of positively valenced emotions. Thus, experiences of happiness include experiences of joy, contentment, satisfaction, fulfilment, tranquillity, peace of mind, stimulation, excitement, attachment, love, admiration, interest, and so on.

All these positively valenced emotions share two features in common. First, they involve representing things as being good in some respect. For example, an episode of satisfaction involves representing the achievement of a goal, which is also represented as being good. Similarly, an episode of attachment involves representing a connection with another, which is also represented as being good. The second feature shared by most positively valenced emotions is that, as well as representing things as being good in some respect, they also tend to *feel* good. We often want to experience emotions of joy, contentment, satisfaction, etc., because such experiences are pleasurable. In sum, the term “happiness” sometimes refers to all positively valenced emotional states: episodic emotions that (a) represent things as being good in some way, and (b) tend to feel good. I will refer to this concept of happiness as Happiness<sup>Positive Emotion</sup>.

The third descriptive concept of happiness can again be viewed as a further extension of the previous concept. This sense of happiness concerns a person’s broader psychological state constituted by their balance of pleasurable states over displeasurable states. Being happy, in this respect, consists in experiencing a greater amount of positive emotion (or pleasure) than negative emotion (or displeasure). One may be happy momentarily, or happy over a longer period of time. If a person has spent more time in a pleasurable state than in a displeasurable state over a certain period of time

then we can say that they were happy over that period of time. I will refer to this concept of happiness as Happiness<sup>Positive State</sup>.<sup>22</sup>

I take it that this concept is the closest concept so far to the concept that refers to something we tend to use as a proxy for well-being in our practical lives. When we think about how well we have been doing, we may tend to think about whether we have, in general, been feeling good or bad. Knowing that a person predominantly feels good may give us a fairly good indication that their life is going well. However, the problem with this concept is that it may refer to what is sometimes called “smiley-face happiness” – a psychological state that may lack both the *breadth* and *depth* of something that tends to be used as a proxy for well-being. In contrast to Happiness<sup>Positive State</sup>, we tend to think of happiness as a broad, relatively long-term, positive psychological condition. If a depressed individual happens to experience a day of predominantly pleasurable states, this does not tend to indicate that they are doing well. The point is that happiness seems to have “deep, far-reaching, and typically lasting consequences for a person’s state of mind and behaviour.” (Haybron, 2008: p.69)

Concepts of happiness that are partly descriptive and partly evaluative may be able to better account for both the breadth and depth of happiness. I will briefly consider three concepts of happiness that fit this description.

Firstly, the term “happiness” is sometimes used to refer to a person’s preferred mental state.<sup>23</sup> This may simply be a person’s balance of pleasurable states over displeasurable states. Or it may be a different kind of mental state, such as an overwhelming state of contentment, tranquillity or

---

<sup>22</sup> Note that Happiness<sup>Positive State</sup> is a descriptive concept of happiness. Some theorists (namely quantitative prudential hedonists) hold that a person’s well-being is constituted by their overall balance of pleasurable states over displeasurable states (Feldman, 2004; 2010; Crisp, 2006). If this were the case, Happiness<sup>Positive State</sup> would be substantively identical to well-being. However, it is conceptually possible that Happiness<sup>Positive State</sup> is not substantively identical to well-being. Thus, Happiness<sup>Positive State</sup> remains a descriptive concept. Our values make no difference to happiness viewed in this sense. Happiness is constituted by a person’s balance of pleasurable states over displeasurable states, according to Happiness<sup>Positive State</sup>, no matter what we think constitutes well-being.

<sup>23</sup> The notion is similar to the idea of preference hedonism put forward by Parfit (1984). According to preference hedonism, a person’s well-being is constituted by their desired state of consciousness. As Scanlon puts it: “the experience of living a life is made better by the presence of those mental states, whatever they may be, which the person living the life wants to have” (Scanlon, 1993: p.186). In contrast to preference hedonism, the concept of happiness in question may have little to do with well-being. In that, well-being may have little to do with a person’s desired state of consciousness.



peace of mind. People may tend to desire states of consciousness that cohere with lives that are good for them. For example, Buddhists tend to emphasise the importance of contentment and peace of mind over satisfaction and pleasure because they believe that the former mental states tend to be beneficial, whereas the latter mental states tend to be harmful (Flanagan 2011). But, this is not a necessary feature of the concept in question; this sense of happiness merely refers to the psychological condition that people desire to have. I will refer to this concept of happiness as Happiness<sup>Desired Positive State</sup>.

The second descriptive/evaluative concept of happiness is similar to Happiness<sup>Desired Positive State</sup>, but subtly different. According to this concept, happiness is the psychological condition that a person *believes* accompanies his or her well-being. Alternatively put, if well-being is characterised as how well one's life goes, then happiness is characterised as how well one's life goes *from the inside* (Campbell MS). Unlike Happiness<sup>Desired Positive State</sup>, happiness in this sense is not necessarily desirable. It is the states of the world that one's psychological state *points towards* that are valuable. This concept of happiness, for instance, refers to the psychological state of feeling that one's self and one's activities, projects, relationships, etc., are worthwhile. One need not, in addition, feel that one's happiness is worthwhile. The value of happiness, in this sense, is derived from its contents.<sup>24</sup> It is a psychological condition that consists in the (perhaps true) belief that one's life, activities, etc., are worthwhile (Raz, 2004). I will refer to this concept of happiness as Happiness<sup>Worthwhile Life</sup>.

Lastly, the term "happiness" is sometimes used to refer to whatever psychological condition (at least partly) constitutes well-being. This is not necessarily the psychological condition that people either desire to have or the condition that accompanies a worthwhile life. For example, a person's well-being may be entirely constituted by feelings of life satisfaction caused by virtuous activity. In this case, the psychological condition that constitutes a person's well-being is being satisfied with their life as a result of being virtuous. Yet, it is possible that people desire pleasurable mental states beyond such feelings of life satisfaction. Similarly, it is possible that people

---

<sup>24</sup> This kind of happiness may be constituted merely by a person's beliefs about the value of her self and her life. Alternatively, it may be constituted by a person's *appropriate* beliefs, namely true beliefs about the value of her self and her life. This is a debate to be had between different conceptions of this kind of happiness.

feel their life is going well when they witness the achievement their life-long projects. Thus, Happiness<sup>Desired Positive State</sup> and Happiness<sup>Worthwhile Life</sup> can, on a conceptual level, come apart from the psychological condition that (at least partly) constitutes well-being. I will refer to this concept of happiness as Happiness<sup>Prudential Experience</sup>.

All three of these partly descriptive and partly evaluative concepts of happiness may refer to psychological conditions that tend to be used as proxies for well-being by laypersons in their practical lives. However, I doubt that we primarily use these kinds of happiness as proxies for well-being. This seems doubtful for two reasons. First of all, due to their partly evaluative nature, each concept permits to a wide range of variation. For instance, with regards to Happiness<sup>Desired Positive State</sup>, people from more “collectivist” cultures tend to value mental states that occur through the realisation of positive social relationships (e.g. compassion), whereas people from more “individualist” cultures tend to value mental states that occur through self-achievement (e.g. pride). With regards to Happiness<sup>Prudential Experience</sup>, people’s happiness will be determined by one’s view of what is good for them (which, as I will mention in the next section, is far from uncontroversial). It does not seem that the kind of happiness that we tend to use as a proxy for well-being permits of this level of variation. Paradigm cases of happiness are psychological states such as being in high spirits, elated, carefree, contented, at ease, self-assured, a sense of fulfilment, and so on. In contrast, paradigm cases of unhappiness are states such as being depressed, melancholy, anxious, stressed out, lonely, empty, etc. (Haybron, 2008) It seems that these states constitute happiness/unhappiness regardless of whether people desire them or whether they contribute towards well-being.<sup>25</sup>

---

<sup>25</sup> Against this view is a series of studies by Sven Nyholm investigating the role that moral judgements play in folk’s attributions of happiness (Nyholm 2007; Phillips et al. forthcoming). His studies suggest that the folk concept of happiness involves some normative or evaluative element such that, of the things that determine whether we take somebody to be happy, one is whether we think that person is living a good life or not. This suggests, according to the folk concept of happiness, happiness is not entirely descriptive.

On closer inspection, I think that Nyholm’s results may turn out to be entirely compatible with the notion of happiness employed by empirical researchers. It may be that people tend to incorporate normative judgements into judgments about happiness because being a good person typically results in being happy. That is, various components of a good life (such as being virtuous, having contact with reality, etc.) may be part of a *prototypical* happy life, whereby happiness is an entirely descriptive psychological state.

The second reason why we do not tend to use the above three kinds of happiness as proxies for well-being is that they are inefficient. What we want from a proxy for well-being is something that is more epistemically accessible than well-being itself – otherwise we would simply refer to our well-being in our practical deliberation and evaluation. Partly descriptive and partly evaluative concepts of happiness, such as Happiness<sup>Worthwhile Life</sup>, require us to know a considerable amount about how well our lives are going for us in order to know whether or not we are happy. This is inefficient. An entirely descriptive notion of happiness is a much more efficient proxy for well-being. A proxy of this kind is more likely to get things wrong – such as in the case of the deliriously happy mental patient – but such costs will tend to be outweighed by the fact that it is epistemically accessible. Assessments of happiness, according to a purely descriptive notion of happiness, are easier to come by: in order to get an idea of how well we are doing, we merely need to evaluate our psychological states.

A purely descriptive notion of happiness is used by empirical researchers in the measurement of well-being. Although social scientists do not consistently refer to one psychological condition as “happiness” (more on this below), they nonetheless study psychological conditions that are entirely descriptive in nature. They also study psychological conditions that have more breadth and depth than states such as Happiness<sup>Positive State</sup>. Haybron refers to such conditions as “Psychological Happiness” (Haybron, 2003). Psychological Happiness is a broad, relatively long-term positive psychological condition. I suggest that, in-line with empirical researchers, we primarily use this kind of happiness as a proxy for well-being in our practical lives. This is the kind of psychological condition that we refer to with phrases such as, “I just want my children to be happy and healthy”, “I’ve not been happy since I lost my job”, “Did you have a happy childhood?”, and so on.

The aim of this overview of different concepts of happiness has been to support the claim that social scientists and the folk tend to use the same kind of happiness as a proxy for well-being. The term “happiness” may refer to a number of different things, but it does not seem that any of these things are well suited for us to primarily use as an informal proxy for well-being in our practical lives. In contrast, the concept of happiness used by empirical researchers (Psychological Happiness) does seem to be well suited in this respect. It seems reasonable to suppose, therefore, that the same kind of happiness is used both by empirical researchers and laypersons as a proxy for well-being.

Of course, it may turn out that this analysis is incorrect. Happiness is an elusive concept and may refer to a number of things that we use in inconsistent ways. If so, I am happy to drop the term “happiness” and simply talk about the kind of broad, relatively long-term psychological conditions that are studied by empirical researchers. Not much rests on the term. The phenomena in question is the psychological condition that we tend to use as a proxy for well-being. If this condition does not turn out to be worthy of the term “happiness” then this thesis can be viewed as investigation into how successfully we can measure well-being through measuring such a psychological condition.

In sum, I have suggested that happiness – viewed as a broad, relatively long-term positive psychological condition – can adequately account for the fact that happiness tends to be used as both a formal and informal proxy for well-being. In the final main section of this chapter I will consider what this psychological condition consists in. I will argue in favour of an “affective state account” of happiness, whereby happiness consists in a person’s affective state.

#### **1.4.2 Summing up where we have got to so far**

We now have a clear idea of what well-being and happiness are. I have argued in favour of the Standard Concept of well-being and the concept of Psychological Happiness as the primary concepts of happiness and well-being. Well-being concerns a life that is good for someone, and its value is distinct from both a good life and a happy life. Happiness is a broad, relatively long-term psychological condition that both empirical researchers and laypersons use as a proxy for well-being.

Moving forward, in the next two main sections, I will consider different *conceptions* of both well-being and happiness respectively. I will not argue in favour of a particular substantive theory/conception of well-being. Instead, I will outline a *theory-neutral account* of well-being. With regards to happiness, I will argue in favour of an *affective state account*. According to this theory, happiness is constituted by a person’s affective state; this adequately accounts for the fact that we tend to use happiness as a proxy for well-being.

## 1.5 Well-being: Conception(s)

In this section, I will consider different conceptions of well-being. Crucially, I will not argue in favour of one particular conception of well-being. What constitutes well-being is a controversial issue – there are no widely accepted substantive theories of well-being (Keller, 2010; Heathwood, 2011). Instead of arguing in favour a particular substantive theory, I will outline a *theory-neutral account* of well-being – a set of core features of well-being that are consistent with most substantive theories of well-being.

I will outline the theory-neutral account of well-being as follows. Firstly, I will briefly show that the three main kinds of theories of well-being are controversial. Secondly, I will describe the theory-neutral account. I will consider both of these issues at greater length in chapter two. In the remainder of this section, I will illustrate the theory-neutral account with regards to the three main kinds of measures of well-being.

### 1.5.1 Different conceptions of well-being

In chapter two, I will provide a comprehensive overview of substantive theories of well-being in order to illustrate that there are currently no widely-*acceptable* (in contrast to widely-*accepted*) theories of well-being. In this section, I will simply show that the three main kinds of substantive theories of well-being are controversial – none of the theories are widely accepted. I will assume that this lack of consensus amongst philosophers of well-being generalises to all current substantive theories of well-being.<sup>26</sup>

Recall from the first main section of this chapter, that the three main kinds of substantive theories of well-being (based on Parfit's influential taxonomy) are mental state, desire-satisfaction and objective list theories. Let us consider mental state theories, first. Mental state theories begin with the intuitive claim that for something to be good *for me* it must make some kind of impact on me. In order to make this impact, I must notice it, or be aware of it – it must impact my experiences. The problem with such a view, however, is that it is neutral over *how* my experiences are generated. I may have genuinely meaningful relationships, and feel good about that, or I may merely think that I have meaningful relationships, despite the fact that the people I care about talk badly about me behind my back. We intuitively think that the latter situation is

---

<sup>26</sup> For overviews, see Keller (2010) and Heathwood (2011)

worse than the former, even if I never find out about the things people say about me when I'm not around. Mental state views, however, do not distinguish between the two scenarios insofar as I have the same experiences in both.

The second main kind of well-being theory – desire-satisfaction theories – make up for the counterintuitive nature of mental states theories by claiming that it is the *actual* satisfaction of our desires that constitute well-being whether we are aware of our desires being satisfied or not. However, such a view faces the following problem: we can desire things that do not seem to be good for us. We may have malicious, stupid or self-destructive desires (e.g. desires to torture innocent others, count blades of grass, eat gravel, etc.) the satisfaction of which does not seem to be of any benefit. Even the satisfaction of our idealised desires (i.e. desires formed under full information and rational capacities) may be not be good *for us*, such as the desire for a distant stranger to be healthy or for human life to flourish over the next few millennia. In short, the problem with desire-satisfaction theories is that they do not guarantee that the objects of our desires are those things that seem to be good for us.

Objective list theories (the final main theory of well-being) solve this problem by simply listing the kinds of goods that do seem to be good for us. These include goods such as love, virtue, knowledge, health, achievement, enjoyment, and so on. According to such theories, the attainment of these goods constitutes our well-being. Yet, despite the fact that these kinds of goods tend to be good for us, we can still imagine certain individuals who do not benefit from such goods. For this reason, objective list theories tend to consist in only a few, very broad objective goods. The problem with this strategy, however, is that we can imagine certain individuals who benefit from certain goods that are not on the list. An objective list theory must find an intuitive balance between these two strategies – something that is yet to have been achieved.

Of course, much more can be said about the merits of each of these three kinds of theories, and there are additional theories of well-being that do not fit into this traditional tripartite classification. The aim of this brief overview was merely to show that none of the current main theories of well-being are widely-accepted. I will consider in further detail in chapter two, whether any current theory of well-being is widely-acceptable. For now, it is enough to show that there is widespread disagreement amongst philosophers of well-

being – this level of disagreement is nicely illustrated by Simon Keller as follows:

“Intuitions about welfare are firm, but neither consistent nor widely held. If your theory of welfare implies that writing a great novel or having a successful marriage makes you better off, just in itself, then you can expect someone to complain that the theory is obviously ridiculous. (How could it make you better off regardless of whether you value or want it?) If your theory does not imply that writing a great novel or having a successful marriage in itself advances your welfare, then you can expect to hear that that is obviously counterintuitive. (Is it not obvious that a life with such things goes better than one without them?) Some think it obviously possible to live a good life on the experience machine – you just need to want to be on it and know that you are on it; and some think it obvious that no life on the experience machine is worth living. It is depressing.” (Keller, 2010: 663)

This level of disagreement motivates the need for an understanding of well-being that is consistent with most substantive theories of well-being. In the next section, I will briefly outline the theory-neutral account, which I believe can provide us with such an understanding.

### **1.5.2 The theory-neutral account of well-being**

I will discuss the theory-neutral account at length in chapter two. For now, it will help to briefly outline how it is that the account can provide us with a broadly informative understanding of well-being that is consistent with most substantive theories of well-being.

To begin with, we can note that we already know the kinds of things that are good for people, even if we don't know whether or not these things constitute people's well-being. As Dan Hausman writes: “We need to know something about what is good for people...but it does not follow that we need a philosophical *theory* of well-being for this purpose...Platitudes concerning what makes people better or worse off like the claims that enjoyment contributes towards well-being and illness diminishes it depend on no philosophical theory that specifies what things are intrinsically good for people and why.” (Hausman 2012: p.7)

I believe that Hausman is correct in noting that platitudes about well-being are broadly informative. The understanding of well-being that platitudes provide us with, however, quickly run into trouble when such platitudes come into conflict with each other. Suppose, for example, that we all agree that material

wealth and enjoyment tend to be good for us. Things start to get complicated when we have to evaluate the well-being of two people who are equal in all respects, except one has a greater level of material wealth and the other has a greater level of enjoyment. Who is better off? Without a substantive theory of well-being to fall back on, relying on platitudes about well-being does not enable us to answer such questions.

Fortunately, I think there is the potential for a more sophisticated understanding of well-being, which remains consistent with most substantive theories. This understanding remains neutral over whether the object of any particular platitude about well-being is constitutively or causally related to well-being. To continue the example, material wealth may either constitute well-being or cause well-being; enjoyment may either be caused by well-being or constitute well-being. For the objects of both well-being platitudes, it seems as if we can say that they tend to correlate with well-being, even if we cannot be sure whether they are constitutively or causally related to well-being.

We can say this because the objects of different platitudes about well-being tend to be *causally connected* with each other. This is the key point of the theory-neutral account. It is because of these causal connections that we can remain neutral over whether the object of a particular well-being platitude is either constitutively or causally related to well-being. I will call the objects of platitudes about well-being that are causally connected with each other in this way “*well-being ingredients*.” Well-being ingredients are causally connected with each other in such a way that we don’t need to know whether any particular well-being ingredient constitutes well-being or not. Returning to the example of material wealth and enjoyment, even if material wealth does not constitute well-being, it is likely to be causally related to something that does. The same goes for enjoyment. As Michael Bishop points out, we do not need to know which well-being ingredients constitute well-being because, “the world has already joined them together with causal bonds.” (Bishop, MS: p.1). Thus, we can say that both well-being ingredients will tend to be correlated with well-being, while remaining neutral over whether either constitutes well-being.

I will discuss the theory-neutral account at length in chapter two. For now, in the next section, I will illustrate how we might apply the theory-neutral account to the three main kinds of measures of well-being.



### 1.5.3 The theory-neutral account and different measures of well-being

How do researchers currently measure well-being? In the main, measures of well-being can be divided into three categories: measures of (a) preference-satisfaction, (b) capabilities and functionings, and (c) subjective well-being (Angner, 2008). I will very briefly review each kind of measure in turn.

Let us consider measures of preference-satisfaction, first. These measures aim to measure how well an individual is doing through measuring the satisfaction of his or her preferences. Researchers typically discover people's preferences either by asking them or observing their behaviour. *Stated-preference* measures typically consist in asking subjects how much they are willing to pay for a particular good or service. If a subject is willing to pay more for one good than another good, it is assumed that she prefers the former good to the latter one. If the subject attains the preferred good, it is assumed that her preference is satisfied. In contrast to stated-preference measures, *revealed-preference* measures typically consist in observing subjects' consumer behaviour. If a subject chooses to pay a certain amount of money for a particular good or service, it is assumed that she prefers that good to other goods that cost the same or less amount of money. It is also assumed that consuming the preferred good satisfies the subject's preference. In general, a subject's level of wealth and access to preferable non-market goods is taken to be a rough proxy for her level of preference-satisfaction.

Now let us consider measures of capabilities and functionings. The measures aim to measure how well an individual is doing through measuring certain things that they do or be (i.e., her functionings) and their abilities to do and be certain things (i.e., her capabilities). Examples of functionings include being secure, nourished, socially connected, having a job, raising a family, having self-respect, being happy, and so on. Capabilities consist in people's abilities to achieve these things. Researchers typically measure a similar list of central capabilities, such as the ability to live a life of normal length, have good health, have bodily integrity, have attachments to things and people outside ourselves, form a conception of the good, and participate effectively in political choices (Nussbaum 2001). These capabilities are either measured subjectively or objectively. Subjective measures typically consist in asking people about their level of each capability on a set-point scale. In contrast, objective measures tend to measure the material conditions related to each

capability, such as measuring a person's ability to live a life of normal length through measuring their life expectancy.

Lastly, let us consider measures of subjective well-being. These measures aim to measure how well an individual is doing through measuring how well they are doing from their own point of view. This typically includes measures of a subject's (affective and cognitive) evaluations of their own well-being.<sup>27</sup> Researchers measure a person's subjective well-being in three kinds of ways. Measures of *affect balance* aim to measure a subject's overall balance of positive over negative affective experiences (Kahneman & Krueger 2006; Fredrickson & Losada 2005). Measures of *life* and *domain satisfaction* aim to measure a subject's satisfaction with their overall life and particular life domains respectively (Diener et al. 1985; Diener et al. 1999). Measures of *well-functioning* (often referred to as measures of *psychological* or *eudaimonic well-being*) aim to measure a subject's attitudes towards particularly important aspects of her functioning, such as her sense of growth, purpose, self-acceptance, mastery, etc. (Keyes et al. 2002; Ryan & Deci 2001). Typically, these different kinds of attitudes are measured using self-reported surveys, whereby subjects are asked to report their level of positive and negative affect, life or domain satisfaction or specific aspects of well-functioning.

We now have an idea of the three main kinds of measures of well-being. I think it is clear that each kind of measure nicely maps onto the three main substantive theories of well-being reviewed above. Measures of preference-satisfaction relate to desire-satisfaction theories of well-being.<sup>28</sup> Measures of capabilities and functionings relate to objective-list theories of well-being.<sup>29</sup>

---

<sup>27</sup> In addition, researchers may aim to measure certain mental states that can be viewed as non-evaluative, such as chronic pleasures and pains. Such mental states may not consist in their subject having some kind of evaluative attitude towards their own well-being. Nonetheless, I think we can make the general claim that measures of subjective well-being aim to measure a subject's (affective and cognitive) evaluations of their own well-being.

<sup>28</sup> Whereas desires can be aimed at a single good or state of affairs, preferences are rankings over alternative bundles of goods or states of affairs. Thus, satisfying preferences often entails the satisfaction of particular desires.

<sup>29</sup> For instance, Martha Nussbaum's list of central capabilities is based on a normative conception of human nature and the capacities that are objectively good for people (Nussbaum 1987). Indeed, measures of capabilities and functionings partly developed out of critiques of preference-satisfaction measures. Proponents of the capability approach note that people can often desire things that are not in their interests (Sen 1999; Nussbaum 2001). This mirrors the debate between desire-satisfaction and objective-list theories of well-being briefly discussed above.

And measures of subjective well-being relate to mental state theories of well-being.<sup>30</sup> Due to the controversial nature of each of the three main substantive theories of well-being, one might conclude that there will tend to be widespread disagreement over the justification of each of the main kinds of measures of well-being.

This need not be the case, however. After all, empirical researchers continue to measure well-being and discover interesting findings without having resolved the debate over the nature of well-being. The theory-neutral account can explain why this practice is justified.

On the basis of the three main kinds of measures of well-being, we can develop a version of the theory-neutral account that consists in three main well-being ingredients, namely our:

- actual prudential achievements
- opportunities for prudential achievements
- awareness of prudential achievements

Each of the main kinds of measures of well-being corresponds to one or more of these three well-being ingredients. Measures of subjective well-being are measures of our awareness of certain prudential achievements (i.e. the objects of our subjective well-being). We may be happy with our relationship, for example, or satisfied with our career success. Measures of preference-satisfaction are measures of our certain prudential achievements (i.e. our preferred prudential achievements). When we act to promote our well-being we will generally aim to satisfy our preferences. Similar to such measures, measures of functionings are measures of certain prudential achievements; though these need not be our preferred prudential achievements. Lastly, measures of capabilities are measures of certain opportunities for prudential achievements. We may have the capacities and opportunities to maintain our relationships, achieve success in our career, satisfy our preferences, be well-nourished, have self-respect, and so on.

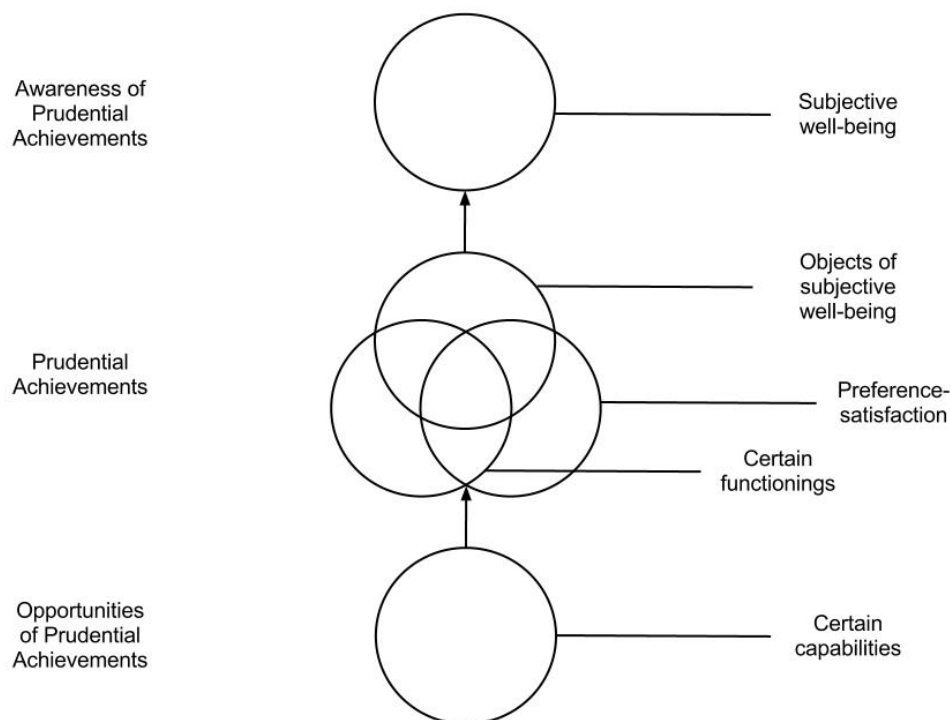
---

<sup>30</sup> Indeed, subjective well-being researchers often talk about how they are aiming to “directly” measure well-being, in contrast to supposedly less direct measures, such as measures of preferences satisfaction or capabilities and functionings (Angner 2010). This kind of claim seems to presuppose that well-being is constituted by subjective well-being, which is a particular version of a mental state theory. In other instances, subjective well-being researchers explicitly assume a mental state theory of well-being: As Ed Diener and Eunkook Suh write: “Subjective well-being research [...] is concerned with individuals’ subjective experiences of their lives. The underlying assumption is that well-being can be defined by people’s conscious experiences – in terms of hedonic feelings or cognitive satisfactions” (Diener & Suh 1997: p.191).

These three main well-being ingredients are causally connected to each other in such a way that we can remain neutral over which well-being ingredient(s) in fact constitute well-being. We are unlikely to have one of these well-being ingredients without the others. Opportunities for prudential achievements, awareness of prudential achievements and actual prudential achievements are causally connected in such a way that, whichever one of these ingredients constitutes well-being, the other ingredients are either caused by or cause well-being.

This is perhaps best seen by the fact that each well-being ingredient plays an important causal part in our prudential decision-making process. First, we need opportunities to attain certain prudential achievements. Second, with the help of such opportunities, we may in fact have certain prudential achievements. Third, upon attaining certain prudential achievements, we may become aware of such achievements. These causal connections are illustrated as follows:

Fig. 1.3



Note that it is only in exceptional cases that these three main well-being ingredients do not tend to be causally connected in this way (e.g. in cases of blissful ignorance or consistent good luck). This is why the three main kinds of measures of well-being will tend to be justified – in general, the things

measured by each kind of measure will tend to be causally connected to the things measured by the other kinds of measures.

Indeed, different measures of well-being may be more or less relevant depending on the practical context in which well-being is measured. In certain contexts the causal connections between ingredients may break down or at least weaken. For example, within the context of childcare, a child's *opportunities* for prudential achievements may be more relevant than their awareness of prudential achievements. A child's opportunities for prudential achievements will have a much greater causal influence on their future well-being. Conversely, within the context of elderly care, an elderly person's *awareness of* prudential achievements may be more important than their opportunities for prudential achievements. An elderly person's opportunities for prudential achievements may have a much weaker causal impact on their future well-being. I believe that variation in the causal connections between well-being ingredients accounts for one of the most notable features of the measurement of well-being, namely that there is a considerable amount of variation between measures used in different contexts (Alexandrova 2013). I think it is because the causal connections between well-being ingredients are weaker or stronger in certain contexts that empirical researchers tend to use different measures of well-being (e.g. measures of child well-being, subjective well-being, physical well-being, etc.) depending on the context in question.

In sum, I have outlined a theory-neutral account of well-being that is consistent with most substantive theories of well-being and different kinds of measures of well-being. The theory-neutral account does not consist in a view of the constituents of well-being – such a view would be overly controversial. Rather, it consists in a number of well-being ingredients – the objects of platitudes about well-being that tend to be causally connected with the objects of other well-being platitudes. According to the theory-neutral account, each well-being ingredient tends to be either constitutively or causally related to well-being. This provides us with a theory-neutral understanding of the things that tend to be correlated with well-being.<sup>31</sup>

---

<sup>31</sup> It is worth noting that the theory-neutral account is not entirely agnostic over substantive theories of well-being. The account may be inconsistent with some substantive theories, such as prudential hedonism. This, however, is a big issue, and one that I shall leave to chapter two.

## **1.6 Happiness: Conception(s)**

In the final main section of this chapter, I will look at different conceptions of (psychological) happiness. Again, for the purpose of this thesis, the following analysis of happiness will be much more straightforward than the previous analysis of well-being. This is because I am concerned with the measurement of well-being. In the first main section of this chapter, I showed that happiness tends to be used as a formal proxy for well-being in the measurement of well-being, and claimed that this is our primary line of evidence that suggests happiness tends to be strongly correlated with well-being. Although empirical researchers refer to a number of different constructs as “happiness”, each of the conceptions of happiness used by such researchers is quite clear. One of the broad, relatively long-term psychological conditions that social scientists refer to as “happiness” is a person’s affective state. This is the conception of happiness that I will be concerned with in this thesis.

As with the analysis of the concept(s) of happiness, one may doubt that this conception of happiness used by empirical researchers matches up with the conception of happiness used by laypersons in their practical lives. This is important insofar as one doubts that the study of happiness provides us with sufficient evidence that happiness and well-being may tend to be strongly correlated. In this section, therefore, I will briefly consider a number of different conceptions of happiness that may be used by laypersons in their practical lives. I will suggest that social scientists and the folk do tend to use the same conception of happiness – the conception that refers to a person’s affective state. This conception of happiness can adequately account for the fact that happiness tends to be used as both a formal and informal proxy for well-being.

### **1.6.1 Different conceptions of psychological happiness**

In the philosophical literature, there are four main accounts of psychological happiness, namely (a) hedonism, (b) life satisfaction theories, (c) affective state views, and (d) hybrid theories (Brülde 2007). I will briefly outline each of these candidates in turn before arguing in favour of a particular version of the affective state theory.

Hedonism is the view that happiness consists in one’s balance of pleasure over displeasure (Feldman, 2010), or enjoyment over suffering (Crisp 2006a). Theories of hedonism differ in numerous ways. For instance, theories differ

over their view on the nature of pleasure. Hedonism Internalism consists in the view that pleasure is a distinct feeling or sensation (Crisp, 2006). In contrast, Hedonism Externalism consists in the view that pleasure is a distinct pro-attitude towards one's experience (Feldman, 2006). Hedonistic theories also differ over how one's balance of pleasure over displeasure is calculated. Following Bentham, some hedonists aggregate pleasures in accordance with the duration and intensity of each pleasure (Feldman, 2010). However, following Mill, other hedonists aggregate pleasures in accordance with additional qualities, such as the complexity and depth, as well as the duration and intensity, of each pleasure (Crisp, 2006). In general, hedonism is largely discredited as a theory of well-being but continues to be widely accepted as a theory of happiness (Feldman 2010; Morris 2011).

Life satisfaction theories consist in the view that happiness is constituted by one's judgements of overall life satisfaction. Theories of life satisfaction differ over the particular kind of judgment involved. For instance, some theorists hold that judgments of life satisfaction concern one's whole life, including one's past, present and (likely) future (Tatarkiewicz 1976; Telfer 1980). Other theorists hold that judgments of life satisfaction concern only some important aspect of one's life (Sumner 1996; Tiberius 2008). In general, however, judgments of life satisfaction involve a global attitude towards one's life. Such attitudes consist in the assessment of how one is doing with regards to the things that one cares about.

Affective state views emphasise the importance of one's emotional condition in contrast to merely having a favourable balance of pleasant over unpleasant experiences (Haybron 2001; Sizer 2010). Such views differ in respect to the kinds of emotional states that constitute happiness, often with an emphasis put on moods, or mood-constituting affective states (Haybron 2005; Sizer 2010). According to such accounts, a person's level of happiness is determined by the extent to which their emotional state, or mood state, is a positive one. Some accounts further emphasise the importance of one's disposition, or relatively long-term propensity, to have an overall positive emotional state (Haybron, 2005). The general idea is that happiness concerns the way in which one emotionally experiences the world. This includes how one typically feels, but also includes how one tends to think and act in accordance with one's perceptions and judgments of how well one's life is going.

Affective state views attempt to accommodate the main features of hedonism and life satisfaction theories. Like hedonism, affective state views emphasise one's affective experiences. Like life satisfaction theories, affective state views emphasise the relationship between psychological happiness and the things that we care about. In contrast to this approach, the remaining theory of happiness that I will consider – the hybrid view – explicitly combines both hedonism and life satisfaction theories (Haybron 2011a).

Hybrid views consist in the view that happiness is constituted by both affect balance and life satisfaction. Such accounts may also include feelings of well-functioning, e.g., a sense of relatedness, purpose, growth, mastery, and so on. According to hybrid theories, a person's level of happiness is determined by the combined weighted contribution of their affect balance, life satisfaction and well-functioning respectively. These three constructs are often combined to form the construct of subjective well-being outlined above.

### **1.6.2 An affective state account of happiness**

In this section I will defend an affective state view of happiness. The affective state account is simply the view that our happiness is constituted by our affective state. Individual affective states are mental states that (a) are representational, (b) represent things as being good/bad, and (c) include non-conscious states. I will consider each of these features in turn, before considering the plausibility of a particular version of this view.

#### *What are affective states?*

Firstly, affective states are representational mental states. That is, our affective states represent certain kinds of state of affairs. For example, when we fear a dog we represent it as being dangerous. Almost all theories of affect hold that affective states represent certain things in this way. Indeed, one of the major criticisms against “feeling theories” of affect – which maintain that affective states are simply constituted by certain feelings – is that such views cannot accommodate the fact that affective states represent certain kinds of states of affairs (Prinz 2004a; Whiting 2011; Kriegel 2011). Now, feeling theories have well-known responses to this criticism, but almost all of these responses consist in explaining how feelings can also be representational mental states. For instance, Prinz (2004a) argues that the bodily responses that constitute emotional feelings can represent “organism-environment relations”. Alternatively, Kriegel (2011) argues that emotion feelings involve a kind of phenomenal intentionality; they are “feelings-



towards” an object (Goldie 2002).<sup>32</sup> Both responses argue in favour of the representational nature of affect.

The representational nature of affective states entails that such states can be evaluated with regard to their epistemic correctness. When I feel fear, an aspect of my environment strikes me as threatening, and either that portion of the world really is threatening, in which case my fear is epistemically adequate, or it only seems threatening to me; in the latter case, my fear contains a misconception of reality, a cognitive error. Note that this does not entail that affective states are nothing but cognitive states i.e., beliefs or judgements. So-called cognitivist views of affect may not be able to account for the hedonic character of affective states, or the fact that affective states often persist even though the person knows better (e.g. my fear can persist even after I have recognized that the object of my fear is not dangerous). I will not debate the merits of cognitivist and non-cognitivist theories of affect here. The point is that all plausible views of affect maintain that affective states are representational states, which can either fail or succeed at representing certain kinds of states of affairs. I will assume, for now, that affective states generally succeed in accurately representing their objects, though I will discuss this issue in detail in chapter three.

What do affective states represent? Affective states have two representational components; affective states represent (a) certain states of affairs, and (b) those states of affairs as being (*pro tanto*) good or bad. Thus, affective states are evaluative states. An episode of satisfaction, for example, involves representing the achievement of a goal, and that the achievement is good. Similarly, an episode of attachment involves representing a connection with another, and that the connection is good. In contrast, an episode of fear involves representing danger, and that such danger is bad. Likewise, an episode of sadness involves representing an irreparable loss, and that such loss is bad. The kinds of states of affairs represented as being good or bad by our affective states vary considerably in virtue of their breadth. As an example of a very broad affective state, we may feel a general anxiety towards our lives, representing our lives as generally being under threat. In contrast, we may feel an itch towards a specific part of our body, representing a particular disruption of the skin. Both states of affairs are represented as

---

<sup>32</sup> See also Doring (2007), on *affective perceptions*, Roberts (2003) on *concern-based construals* and Helm (2009) on *felt evaluations*.

being bad, yet the represented states of affairs vary considerably with regards to their level of generality.

Interestingly, these two representational components of affective states can sometimes come apart (Shriver 2014). For instance, people given morphine will often report that they still feel pain but that it no longer feels bad. Similarly, lesions to particular brain regions have been shown to diminish the unpleasantness of felt pain – again, patients say that they still feel pain but that it “is no longer bothersome.” A particular instance of this is the condition known as pain asymbolia, where people report feeling pain but show no aversive reaction to painful stimuli. One way of interpreting such findings is that people’s affective states still consist in a representation of a certain state of affairs (e.g. a particular bodily dysfunction) but no longer represent that state of affairs as being bad (e.g. as something to avoid or prevent).

The final feature of affective states worth mentioning is that affective states can unconsciously represent things as being good or bad. We can view our affective states as a constant monitor of how well we are doing with regards to our objectives, goals and projects (Railton 2008). It is possible that a large amount of this monitoring takes place on a non-conscious level (Winkielman & Berridge 2004). This is not to say that affective states are not often conscious experiences. It also does not rule out the possibility that all affective states are accessible to consciousness. The point is merely that affective states are wide-ranging phenomena, which constantly represent the evaluative status of perceived events, situations, actions, thoughts, etc. It is reasonable to presume that a large amount of this activity takes place beyond a person’s level of conscious awareness (Chen & Bargh 1999).

#### *Towards an affective state account of happiness*

Now that we have an understanding of the nature of affect, we are in a position to see what the affective state account of happiness consists in. Firstly, according to the version of the account that I will consider in this thesis, happiness is a *broad* psychological condition. That is, it consists in a wide range of affective states, including emotions, moods, pains and pleasures. At any given point in time, a person’s affective state will consist in a range of positive and negative affect states; some of these may be consciously felt affective experiences, others may be unconscious. For example, at a certain point in time, we may be feeling angry about a particular insult, but also satisfied with our response to the insult. At the same time, we may be excited about our evening plans while feeling a bit anxious about

running late. Underlying these conscious or sub-feelings may be states of fatigue or irritability, and perhaps further anxiety caused by the combination of low energy, the insult and running late. Together, these affective states (and perhaps several more) make up our affective state during this point in time.

The second feature of this version of the affective state view is that happiness is a *relatively long-term* psychological condition. As mentioned above, if a depressed person has a day in which they predominantly experience positive affective states, this does not mean they are happy. Happiness is constituted by our affective state over a relatively long period of time. I suggest that the extent to which happiness is a long-term condition is determined by how efficient it is to treat happiness as such when using it as a proxy for well-being. In the short-term, our affective state is not an efficient proxy for how well we are doing, as the example of the momentarily cheery depressed person shows. However, in the long-term, our affective state may also not be an efficient proxy for well-being. In order to evaluate how well we are doing, we do not want to have to consider how we've been feeling for the past few years (though, in certain contexts, this may make sense, e.g. how well we are doing since we moved house a year ago, or changed job, had children, and so on). In general, it may be efficient to view happiness as our affective state over a relatively long (but not too long) period of time.

Putting these two features of the affective state view together, a person's level of happiness consists in their affective state over a relatively long period of time. If our positive affective states outweigh our negative affective states (over that period of time) then we have a positive affective state. Conversely, if our negative affective states outweigh our positive affective states then we have a negative affective state. A person's level of happiness is determined by the extent to which their affective state is either positive or negative.<sup>33</sup>

---

<sup>33</sup> This affective state view may resemble hedonism more than it resembles other affective state accounts of happiness, such as Haybron's emotional state account (Haybron, 2008). Haybron distinguishes between central and peripheral affective states, arguing that only the former constitutes happiness. In addition, he argues that a person's relatively long-term propensity to experience central affective states may constitute happiness. In response, to the distinction between central and peripheral affective states, I believe that Morris (2011) is correct in claiming that this distinction breaks down upon further investigation. I also side with Feldman's (2010) critique of a person's affective dispositions partly constituting happiness. The result is that I believe the best formulation of Haybron's emotional state account is effectively the affective state view outlined in this section.

It is worth noting how two different kinds of affective states – emotions and moods – fit into this picture. I will consider the difference between emotions and moods in more detail in chapter five. For now, I will simply clarify that both kinds (indeed, all kinds) of affective states constitute happiness, conceived of as a person’s affective state over a relatively long period of time. Yet, we should expect that emotions, in contrast to moods, tend to *dominate* our affective state. This is because emotions tend to be fast-responding, fast-dissipating affective states needed for moment-to-moment fine-grained action guidance (Railton, 2008). In contrast, moods set the tone for our life. As Laura Sizer puts it: “[mood] is more subtle [than emotion], acting as background to our ongoing activities. [It] has more to do with the way one approaches life as a whole than with reactions to particular objects or events.” (Sizer, 2010: 147) The result is that moods respond to general changes in our circumstances, in contrast to emotions, which occur at a frequent rate to meet our current needs of the moment. Moods may influence our general outlook on life, but emotions will primarily constitute our affective state.<sup>34</sup>

#### *Assessing the account*

According to this version of the affective state account, a person’s level of happiness consists in their affective state over a relatively long period of time. This is one of the main constructs of happiness used by empirical researchers in the measurement of well-being. As mentioned above, social scientists aim to measure well-being through measuring happiness, whereby happiness is often (though not always<sup>35</sup>) defined as a subject’s affective state, or “affect balance.” For instance, in a major review of the benefits of happiness, Sonia Lyubomirsky et al., refer to happiness as a person’s balance of positive over negative affective states (Lyubomirsky et al. 2005). In addition, Daniel Kahneman’s construct of “Objective Happiness” consists in a subject’s affective state as measured by either real-time self-reports or physiological and neurological responses (Kahneman 1999; Davidson 2004). The fact that

---

<sup>34</sup> It is also worth noting that moods and emotions will tend to interact in the following two ways: (a) moods dispose us to experience mood-congruent emotions, and (b) repeated or intense emotions can contribute to the onset of a mood (Sizer, 2010; Prinz, 2004).

<sup>35</sup> Empirical researchers sometimes define happiness as a subject’s life satisfaction. At other times, the term “happiness” is used to refer to a subject’s subjective well-being, consisting of their affect balance, life satisfaction, and maybe also their feelings of well-functioning.

happiness tends to be used as a formal proxy for well-being in this way is evidence that happiness may be strongly correlated with well-being.

In addition, I believe that this version of the affective state theory adequately accounts for the way in which laypersons tend to use happiness as an informal proxy for well-being. Firstly of all, the account fits with many of our intuitions over happiness. It adequately accounts for what Haybron (2005) claims are the paradigmatic cases of happiness/unhappiness. On the positive side, Haybron lists feeling elated, carefree, contented, at peace and delighted with one's life. On the negative side, Haybron lists feeling worried, lonely, empty, low, worthless and deeply dissatisfied with life. These are all affective states.

More importantly, however, is that our affective state may tend to be strongly correlated with our well-being in the following two ways. First, as mentioned above, affective states represent aspects of our lives that are good or bad. We tend to feel sad when we are separated from loved ones, satisfied when we achieve a goal, proud when we live up to our values, and so on. Thus, when our lives are going well, we can assume that we will experience a range of positive affective states. Conversely, when our lives are going badly, we are likely to feel consistently bad. In chapters three and four, I will discuss the thesis that affective states tend to indicate well-being. For now, it is enough to note that our affective state may tend to indicate how well our life is going.

Second, affective states have far-reaching consequences for a person's state of mind and behaviour – they have what Haybron calls *causal depth* (Haybron, 2003). Affective states cause us to attend, think and act in certain ways. This is because our affective states motivate us to think and act in response to the good or bad things that they represent. We are motivated to avoid or prevent the things that we represent as being bad, and are motivated to sustain or seek out the things that we represent as being good. These attentive, cognitive and motivational features of affective states can have profound effects. Consider the difference between an overwhelmingly negative affective state and its opposite. As Haybron notes: "Depression strips pleasure from life, diminishing our functioning in myriad respects, and makes our projects, plans and relationships go worse; Severe anxiety similarly undermines our well-being in many ways." (Haybron, forthcoming: p.11) In contrast, an overwhelmingly positive affective state tends to have the opposite effects.

Lastly, the affective state view defended here adequately accounts for the *efficiency* with which we use happy as an informal proxy for well-being in our practical lives (Haybron, 2003). People have limited cognitive resources and information. Yet, they still need to assess various aspects of their lives, such as how well they have been doing recently, the benefits of a particular decision, and so on. Referring to our affective states is an efficient way of determining how well we are doing. Rather than gathering a large amount of data and performing complicated equations to determine how well we are doing, we can simply recall how we have been feeling recently. It is presumably for this reason that subjective well-being researchers aim to measure well-being through measuring people's affective state – because it is something that we tend to do already in our practical lives.

Again, as mentioned in the main section above on the concept(s) of happiness, it may turn out that this account of happiness is incorrect. If so, I am happy to drop the term “happiness” and simply talk about a person's affective state. Not much rests on the term. The phenomenon in question is our affective state, which we tend to use as a proxy for well-being. If our affective state does not turn out to be worthy of the term “happiness” then this thesis can be viewed as investigation into how successfully we can measure well-being through measuring our affective state.

In sum, I have suggested that social scientists and the folk do tend to use the same conception of (psychological) happiness – the conception that refers to a person's affective state. I believe that this conception of happiness adequately accounts for the fact that happiness tends to be used as both a formal and informal proxy for well-being.

## **Conclusion**

In this chapter, I have argued that we have good reason to suppose that there tends to be a strong correlation between happiness and well-being. This potential correlation provides us with reason to investigate how successfully we can measure well-being through measuring happiness.

I have clarified the concepts and conceptions of well-being and happiness that I will be using throughout this thesis. I have argued that a person's well-being is how well that person's life is going for them, and that this notion is distinct from both a happy life and a good life. Rather than arguing in favour of any

particular (controversial) substantive theory of well-being, I outlined a theory-neutral account. This account provides us with a broadly informative understanding of well-being that is consistent with most substantive theories of well-being.

With regards to happiness, I am interested in the broad, relatively long-term psychological condition that empirical researchers often use as a formal proxy for well-being. In particular, I am interested in the psychological condition constituted by a person's affective state. I have argued that this concept and conception of happiness adequately accounts for the way in which laypersons tend to use happiness as an informal proxy for well-being. We have sufficient evidence, therefore, to suppose that happiness (conceived of as one's affective state) tends to be strongly correlated with well-being.

## Chapter 2

### The Relationship Between Happiness and Well-being

#### Abstract

In this chapter, I will consider how we can further investigate the correlation between happiness and well-being. I will argue that we can do so without a substantive theory of well-being. I will develop the theory-neutral account of well-being outlined in chapter one. The theory-neutral account provides us with an understanding of well-being that is consistent with most substantive theories of well-being.

In the first section of the chapter, I will show why we need an understanding of well-being that is consistent with most theories of well-being. The reason is that there are no widely acceptable substantive theories of well-being. I will show that we need a widely acceptable account of well-being in order to investigate the correlation between happiness and well-being.

In the second section of the chapter, I will consider how we can investigate the correlation between happiness and well-being without a substantive theory of well-being. I will consider three different kinds of methods for developing a theory-neutral understanding of well-being. I will argue that only one method results in an understanding of well-being that is both consistent with most substantive theories of well-being and broadly informative. I will use this theory-neutral account in the remainder of the thesis to investigate how successfully we can measure well-being through measuring happiness.

In the third section of the chapter, I will consider a particularly influential substantive theory that is not consistent with the theory-neutral account, namely prudential hedonism. Prudential hedonism maintains that happiness correlates with well-being to a greater extent than the theory-neutral account. I will argue that such a substantive theory is implausible. In the fourth section of the chapter, I will use a similar argument to claim that the opposite kind of substantive theory – one that maintains happiness correlates with well-being to a lesser extent than the theory-neutral account – is also implausible. The upshot is that the theory-neutral account is consistent with all plausible substantive theories of well-being.



## 2.1 Introduction

Perhaps the most straightforward way of investigating how successfully we can measure well-being through measuring happiness would be to do the following: (a) provide an account of what happiness is, (b) provide an account of what well-being is, (c) show how the two things are related, and (d) assess the implications of this relationship for the measurement of happiness. In chapter one, I covered step (a). First of all, I argued that happiness refers to a (broad, relatively long-term) positive psychological condition that empirical researchers and laypersons tend to use as a proxy for well-being. I then argued that this psychological condition could plausibly be viewed as a person's affective state. This thesis, therefore, can be viewed as an investigation into how successfully we can measure well-being through measuring a person's affective state.

Unfortunately, however, covering step (b) is less straightforward. This is because the nature of well-being is controversial. There are no widely accepted substantive theories of well-being in the philosophical literature. Although well-being theorists agree on the things that tend to cause well-being (such as intimate relationships, friendships, professional and academic success, etc.) such theorists disagree over the constituents of well-being.

In this chapter, therefore, I will aim to provide an adequate understanding of well-being while saying as little as possible about what well-being consists in. I will do this by further outlining the theory-neutral account of well-being introduced in chapter one.

## 2.2 Substantive theories of well-being

In this section, I will provide a brief overview of substantive theories of well-being to give an idea of the disagreement that pervades the philosophical literature. I will classify theories of well-being on the basis of two widespread disagreements in particular, namely disagreements over the existence of the experience requirement and the pro-attitude requirement. I will then consider the implications of such widespread disagreement.

### 2.2.1 A brief overview of substantive theories of well-being

There are two different kinds of philosophical theories of well-being, namely *enumerative* and *explanatory* theories of well-being (Crisp 2006b; Woodard 2013; Taylor 2012; Fletcher 2013). Enumerative theories attempt to answer

the following question: what are the constituents of (a person's) well-being? In contrast, explanatory theories attempt to answer the following question: in virtue of what is any given item supposed to be a constituent of (a person's) well-being? For example, an enumerative theory might hold that pleasure is a constituent of well-being; a complementary explanatory theory might hold that pleasure is a constituent of well-being in virtue of its pleasurable nature, or because it satisfies a subject's desires (Woodard, 2013: p.5).

I will not focus on explanatory theories of well-being in this chapter. For the purposes of *measuring* well-being we need to have an idea of the *constituents* of well-being. A complete understanding of well-being would include an explanation of why these constituents are good for us. But such an explanation is not necessary for the measurement of well-being. Instead, I will focus solely on enumerative theories of well-being – what well-being consists in.

As mentioned in chapter one, enumerative theories of well-being are largely divided by two kinds of disagreements, namely disagreements over the *experience requirement* and the *pro-attitude requirement* (Woodard, 2013: p.8-9). I will very briefly consider each requirement.

Consider the experience requirement, first. Those in favour of the experience requirement argue that the only constituent of a subject's well-being are their experiences (Griffin, 1986: 13, 16-19; Scanlon, 1993: 186-187; Sumner, 1996: 127-128). According to such theories, if some fact about your life does not affect your experience, it cannot affect your well-being.

There are intuitively strong arguments both in favour of and against the experience requirement. Kagan (1992) makes the following argument in favour of accepting the requirement. Something contributes towards a person's well-being if it is good for that person. According to Kagan, persons are nothing other than a body and mind. Thus, something can only constitute a person's well-being if it makes a difference to their body or mind. Being genuinely successful, for instance, does not benefit the person except insofar as it directly impacts them i.e. through their experiences of being successful.

In response, theorists who reject the experience requirement argue that we care about whether we are genuinely successful in such endeavours, not just

whether we have certain experiences of success.<sup>36</sup> Moreover, we do not seem to be mistaken in this respect. Thought-experiments such as Nozick's experience machine appear to support the idea that actual states of affairs may constitute our well-being as well as our experiences (Nozick, 1974). We do not seem to think that a life divorced from reality would be a life of well-being.

Now consider the pro-attitude requirement. Those in favour of the pro-attitude requirement argue that the only constituents of a subject's well-being are the satisfactions of some (actual or hypothetical) pro-attitude. According to such theories, if I do not have a certain kind of (actual or hypothetical) pro-attitude towards some fact about my life, it cannot affect my well-being.

As with the experience requirement, there are intuitively strong arguments both in favour of and against the pro-attitude requirement. In favour of accepting the requirement, it does not seem that certain goods (such as contact with reality, health or long-term relationships) constitute a person's well-being when that person is (actually or hypothetically) averse to those goods. The goods in question may be good in some other respect, such as morally or aesthetically good, but it seems objectionably paternalistic to insist that they are also good for someone who lacks (or would lack) certain pro-attitudes towards them. Alternatively put, it does not seem that we can justify the value of such goods to people who do not have certain pro-attitudes towards those goods (Tiberius, 2007).

In response, theorists who reject the pro-attitude requirement argue that certain goods constitute people's well-being regardless of their attitudes towards them. It seems that goods such as achievements, knowledge, virtue, etc., are valued because they are good for people, rather than being good for people because they are valued (Hausman, 2011). Moreover, people's attitudes are systematically prone to error or bias. Desires and values are formed with limited information, overly influenced by the present, emotionally salient stimuli, and so on (Kahneman 2011). Appealing to the attitudes that people would have with full information or full rational capacities seems to be either ad-hoc or incoherent with the intuitions behind endorsing the pro-attitude requirement (Rosati 1995; Hawkins 2010).

---

<sup>36</sup> Of course, the fact that we care about something (e.g. being genuinely successful) does not necessarily mean that it is good for us (Baber 2008). I will discuss this issue in more detail below.

The experience and pro-attitude requirements divide theorists of well-being because of the strong intuitions in favour of both their acceptance and rejection (Keller 2009). It is further worth noting that the acceptance/rejection of one requirement does not necessarily entail anything about the acceptance/rejection of the other. One may accept the experience requirement yet reject the pro-attitude requirement, or vice versa. Or one may either accept both requirements or reject both requirements. Any of these theories of well-being are possible (though, of course, some may be more plausible than others). In the remainder of this section, I will outline the different kinds of theories of well-being possible based on the rejection/acceptance of the experience and pro-attitude requirements (Woodard, 2013: p.10).

Firstly, a theory of well-being may accept both the experience requirement and the pro-attitude requirement. Such theories would hold that a subject's well-being is constituted by experiences of states of affairs that he or she has certain pro-attitudes towards. There are several influential theories of well-being that meet this description. For instance, Heathwood's *Subjective Desire Satisfaction* theory of well-being is the view that well-being is constituted by pleasure, which Heathwood views as the believing that one's present desires are being satisfied (Heathwood 2006). Similarly, Feldman's *Attitudinal Hedonism* theory is the view that well-being is constituted by instances of attitudinal pleasure, which Feldman views as a pro-attitude towards some state of affairs (Feldman 2002). Sumner's *Authentic Happiness* theory also fits into this category (Sumner, 1996). Sumner views well-being as constituted by authentic (i.e. informed and autonomous) attitudes of life satisfaction.<sup>37</sup>

Secondly, a theory of well-being may accept the experience requirement and reject the pro-attitude requirement. Such theories would view a subject's well-being as constituted by certain experiences, but these experience needn't be

---

<sup>37</sup> Of course, this introduces *additional* requirements into a theory of well-being, namely that experiences of states of affairs that one has a pro-attitude towards must also be "authentic." Sumner states that these additional requirements are a positive feature of his account, in that they make up for the problems of mental states theories that reduce well-being entirely to experiences. He claims that, "since [the Authentic Happiness theory] incorporates an information requirement (as part of its conditions of authenticity), it is a state-of-the-world theory. That I experience a state of affairs is necessary in order for it to benefit me, but (since the experience may be illusory or deceptive) it is not sufficient." (Sumner, 1996: p.175)

directed at states of affairs that he or she has certain pro-attitudes towards. Classical hedonism is a case in point. It is the view that well-being is constituted by experiences of pleasure, whether a subject cares about the objects of pleasure or not (Bradley 2009). Hedonistic theories can be further divided by the kinds or varieties of pleasure that they view as constitutive of well-being (Crisp 2006a; Bramble, MS). Another kind of theory that falls into this second category description is Kagan's *Enjoyment of the Good* theory (Kagan 2009). Kagan views well-being as constituted by enjoyable experiences caused by the attainment of certain objective goods, such as achievement, virtue, knowledge, etc.

Thirdly, a theory of well-being may reject the experience requirement and accept the pro-attitude requirement. Such theories would view a subject's well-being as constituted by states of affairs that he or she has certain pro-attitudes towards. These states of affairs need not have an impact on the subject's experiences.<sup>38</sup> Perhaps the most standard version of this kind of account is the Simple Desire-Satisfaction theory of well-being, which is the view that well-being is constituted by the satisfaction of a subject's desires (Murphy 1999; Heathwood 2005). Other theories maintain that different kinds of pro-attitudes (other than desires) matter. Aim-Achievement theories consist in the view that well-being is constituted by the achievements of a subject's goals (Keller 2004; Scanlon 1998; Raz 1988; Portmore 2007). Value-Fulfilment theories consist in the view that well-being is constituted by the realization of a subject's values (Raibley 2010; Tiberius 2013a). Theories in this category also differ with regards to additional requirements that they may accept or reject. The Ideal Desire-Satisfaction theory of well-being is the view that well-being is constituted by the satisfaction of a subject's desires under certain conditions, such as full information and rationality (Griffin, 1986: p.26-34; Brandt 1998; Railton 2003). Lastly, another kind of theory that falls into this third category is the Hybrid Desire-Satisfaction theory of well-being. It is the view that well-being is constituted by the attainment of certain objective goods (achievement, virtue, knowledge, etc.) that a subject desires (Lauinger 2013; Olsaretti 2006).

Fourthly, a theory of well-being may reject both the experience requirement and the pro-attitude requirement. Such theories would view a subject's well-

---

<sup>38</sup> Unless, of course, a subject had a pro-attitude towards having a certain kind of experience, such as desiring a pleasurable life.

being as constituted by certain states of affairs; he or she needn't have certain pro-attitudes towards them, nor need these states of affairs have an impact on the subject's experiences.<sup>39</sup> Objective-List theories (Fletcher 2013; Sarch 2009; Arneson 2009; Brink 1989: p.231-240) consist in the view that well-being is constituted by the attainment of certain goods, such as achievement, knowledge, virtue, quality relationships, health, etc. Similarly, Nature-Fulfilment theories consists in the view that well-being is constituted by the exercise of a person's (cognitive, affective, sensory and social) capacities (Haybron 2008a; Kraut 2009; Nussbaum 2001). Such theories differ in virtue of their focus on either individual or species-specific capacities (Haybron, MS). Lastly, another kind of hybrid theory that falls into this fourth category is the Disjunctive Hybrid theory of well-being (Woodard, 2012: p.11). It is the view that well-being is constituted by either certain experiences *or* the satisfaction of certain pro-attitudes (or something else, such as the attainment of certain goods). Such theories thereby reject both the experience and pro-attitude requirement, yet maintain that experiences and the satisfaction of pro-attitudes are central to well-being (Hawkins, 2010).

### **2.2.2 Implications of widespread disagreement**

In the previous section, I provided a brief overview of the kinds of substantive theories of well-being that are possible, categorized on the basis of the acceptance/rejection of the experience requirement and the pro-attitude requirement. I hope to have shown that there is widespread disagreement within the philosophy of well-being literature. In this section, I will consider the implications of such disagreement.

One might be tempted to conclude, on the basis of widespread disagreement over substantive theories of well-being, that there is something dubious about the *concept* of well-being. Perhaps there is no unified notion of well-being that can account for all of the important roles that well-being plays in our practical lives. I considered this possibility in chapter one, and argued that there is at least one concept of well-being (what I called the Standard Concept) that plays a prominent role in both our self- and other-interested practical deliberation and evaluation. This concept of well-being does not necessarily entail the acceptance/rejection of either the experience requirement or pro-attitude requirement. Thus, narrowing our focus to this

---

<sup>39</sup> Unless, of course, the kinds of states of affairs that matter include either having certain experiences (e.g. pleasure) or satisfying certain pro-attitudes (e.g. value-fulfilment).

concept of well-being does not help us resolve the widespread disagreement discussed above.

Instead, one might think that such widespread disagreement is the result of the philosophical methodology used to generate substantive theories of well-being. For instance, philosophers tend to rely heavily on thought-experiments in order to argue either in favour or against particular theories of well-being. These thought-experiments are often outlandish or unrealistic, requiring us to imagine, for example, individuals who only care about counting blades of grass (Rawls 2005) or who are plugged into a machine that perfectly simulates reality (Nozick, 1974). Our intuitions towards such scenarios may be influenced by a number of biases, thereby rendering them unreliable guides to the nature of well-being (Elster 2011).

I think that this methodological point is reasonable. However, recognizing the unreliability of our intuitions towards outlandish thought-experiments does not straightforwardly resolve the disagreement that pervades the philosophical literature. For, all substantive theories of well-being tend to rely heavily on thought-experiments. A ban on unrealistic thought-experiments may not achieve the effect of rendering certain kinds of theories of well-being implausible, leaving only a few plausible kinds of substantive theories behind. This is not to say that methodological changes in philosophical theorising over well-being could not have this effect. The point is that such changes would still require considerable argument in order to resolve the disagreements over the acceptance/rejection of the experience requirement and the pro-attitude requirement.

The upshot is that there is no easy way in which the widespread disagreement discussed above might be resolved. We should conclude that no current substantive theories of well-being are widely acceptable, at least not without considerable argument. In the remainder of this section, I will show that this matters for investigating how successfully we can measure well-being through measuring happiness.

We need a widely acceptable understanding of well-being in order to further investigate the relationship between happiness and well-being. This is illustrated by the current debate surrounding a set of findings from the study of happiness referred to as the “Easterlin Paradox” (Easterlin 1974; 2001). These findings show that, although significant increases in average income correlate with significant increases in happiness in the short-term, increases in income do not correlate with changes in happiness in the long-term (>10

years).<sup>40</sup> Simply put, the Easterlin Paradox suggests that income makes little difference to lasting happiness.

Now, the validity of the findings that make up the Easterlin Paradox has been challenged. For instance, Inglehart et al. (2008) have shown that significant increases in income make little difference to life satisfaction in the long-run, but do make a difference to lasting happiness. Additionally, Sacks et al. (2010) have shown that relative increases in income correlate with lasting increases in subjective well-being.<sup>41</sup> Of course, these challenges have also been challenged themselves (see Easterlin et al. 2010 for some replies).

However, assuming that the findings that make up the Easterlin Paradox are valid, we need to interpret the prudential relevance of these findings. That is, we need to assess the implications of the Easterlin Paradox for the relationship between income, happiness and well-being. On this issue, there is considerable disagreement. On the one hand, some theorists argue that such findings show income does not correlate with well-being; on the other hand, some theorists argue that such findings show the opposite result: namely, that it is happiness that does not correlate with well-being.

For instance, one of the most prominent “happiness economists”, Andrew Oswald, in his article in the Financial Times entitled, “The Hippies Were Right all Along about Happiness”, ends his discussion of the Easterlin Paradox saying as follows: “Happiness, not economic growth, ought to be the next and more sensible target for the next and more sensible generation.” (Oswald, 2006) These sentiments are echoed throughout the happiness economics literature. For example, economists Richard Layard (2005) and Robert Frank (2001) have both proposed that, on the basis of the Easterlin Paradox, governments should increase or introduce new taxes to discourage people pursuing wealth over goods that do bring lasting happiness, such as relationships, volunteering, leisure, and so on. All of these interpretations presume that happiness strongly correlates with well-being, and that income

---

<sup>40</sup> These seemingly paradoxical results are supported by two different kinds of findings: (a) the correlation between income and happiness *within* countries (whereby increases in average income correlate with significant increases in happiness in the short-term) and (b) the correlation between income and happiness between countries (whereby increases in national income do not correlate with significant increases in happiness in the long-term).

<sup>41</sup> See also Deaton (2010), Hagerty & Veenhoven (2006) and Stevenson & Wolfers (2008) for further critiques of the Easterlin Paradox.



does not (or at least that happiness correlates with well-being to a greater extent than income does).

In response, other theorists have (not surprisingly) hit back at the prudential relevance of happiness (for an overview see Booth et al., 2012). Theorists who view economic growth as an effective way of improving people's welfare, such as through alleviating poverty, tend to therefore argue that income strongly correlates with well-being. Such theorists often say something along the lines of, "there are more important things to life than happiness." If people can be doing badly, such as being in poverty, but remain happy all the same, then it seems that happiness does not correlate with well-being to the extent that happiness economists tend to make out (Sen, 1999; Nussbaum, 2000). This interpretation presumes that income strongly correlates with well-being, and that happiness does not (or at least that income correlates with well-being to a greater extent than happiness does).

These kinds of disagreements are problematic for investigating how successfully we can measure well-being through measuring happiness. In the contexts in which happiness does not correlate with certain goods that may be prudentially relevant, we need to make a judgment: either happiness or the good in question correlates with well-being. For example, who is better off: the unhappy rich person or the happy poor person? The debate over the implications of the Easterlin Paradox shows that such judgements depend on our understanding of well-being. In order to make judgements that are widely acceptable, we need an understanding of well-being that is not based on any particular (controversial) substantive theory of well-being. In short, we need an understanding of well-being that is consistent with most substantive theories of well-being.

This is not to say that there is a problem with attempting to determine the correct substantive theory well-being. The problem is that attempts to do so have not as yet been successful at being widely acceptable. For the purpose of measuring well-being through measuring happiness, we need an understanding of well-being that is widely acceptable in order to make judgements about the contexts in which happiness does and does not correlate with well-being. Thus, for the purpose of this thesis, it will be helpful to start from an understanding of well-being that is consistent with most substantive theories of well-being.

## 2.3 The theory-neutral account of well-being

In this section, I will consider how we can investigate the correlation between happiness and well-being without a substantive theory of well-being. I begin by considering two potential methods. Firstly, we can construct a list of all the goods that have prudential value according to *any* plausible substantive theory of well-being. What we are left with is a number of *platitudes* about what is good for people (e.g. positive experiences, the satisfaction of pro-attitudes, objectively valuable goods, etc). Secondly, we can construct a list of all the goods that have prudential value according to *all* plausible substantive theories of well-being. What we are left with is a *consensus* over what is good for people (e.g. positive experiences of the satisfaction of pro-attitudes towards objectively valuable goods).

I will argue that the list of goods generated by the first method is too *unrestricted* to be consistent with most substantive theories of well-being. For example, if both income and happiness are goods on the list, theorists may reasonably disagree over who is better off: a rich unhappy person or a poor happy person. In contrast, I will argue that the list of goods generated by the second method is too *restricted* to be broadly informative. For example, neither income nor happiness may be goods on this list – neither good may be causally related to the goods that have prudential value according to *all* plausible substantive theories of well-being.

I will then consider a third method that is neither too unrestricted nor too restricted, and is therefore both consistent with most substantive theories of well-being and broadly informative. According to this method, we can construct a list of all the goods that (a) have prudential value according to any plausible substantive theory of well-being *and* (b) tend to be causally connected to other goods that meet condition (a). What we are left with is a number of *platitudes* about what is good for people, the objects of which are *causally connected* to each other (e.g. positive experiences, which tend to be caused by the satisfaction of pro-attitudes, which tend to be caused by objectively valuable goods, etc.). I will refer to these goods as *well-being ingredients*.

This method provides us with a theory-neutral account of well-being. For any particular well-being ingredient, it remains neutral over whether it is either constitutively or causally related to well-being. For example, insofar as positive experiences and the satisfaction of pro-attitudes are causally related, it may either be that (a) positive experiences constitute well-being and the

satisfaction of pro-attitudes are causally related to well-being, or (b) the satisfaction of pro-attitudes constitute well-being and positive experiences are causally related to well-being. Either way, we can say that a particular well-being ingredient *correlates* with well-being to the extent that it is causally related to other well-being ingredients. I will use this theory-neutral account in the remainder of the thesis to investigate the correlation between happiness and well-being.

### 2.3.1 Platitudes about well-being

In this section, I will discuss one way in which we can develop a theory-neutral account of well-being. It involves constructing a list of all the goods that have prudential value according to *any* plausible theory of well-being. For example, mental state theories maintain that certain positive experiences, such as pleasure, have prudential value. Pleasure tends to be caused by certain goods and activities, such as intimate relationships, skill-development, good health, and so on. These things will be on the list. Similarly, objective list theories maintain that certain goods and activities, such as quality friendships, meaningful achievements, knowledge, virtue, etc., have prudential value. These goods tend to cause certain mental states or states of affairs, such as enjoyment, interest, engagement, and so on. These things will also be on the list.

At the end of this process, we will be left with a list of *platitudes* about what is good for people. We know that the objects of these platitudes are good for people without knowing whether or not they constitute well-being. As Daniel Hausman notes:

“We need to know something about what is good for people ... but it does not follow that we need a philosophical *theory* of well-being for this purpose. For example, urban economists can easily see that a new community swimming pool contributes to individual welfare because they know that people enjoy swimming and socialising at swimming pools. This conclusion assumes that enjoyment contributes towards well-being. Similarly, economists can conclude that the prevalence of malaria-carrying mosquitoes diminishes well-being, if they assume that sickness and death diminish well-being. Platitudes concerning what makes people better or worse off like the claims that enjoyment contributes towards well-being and illness diminishes it depend on no philosophical theory that specifies what things are intrinsically good for people and why.” (Hausman, 2011: p.7)

I think Hausman is right to note that platitudes about well-being can provide us with a list of prudential goods that is broadly informative. However, the problem with this method is that the list generated is unlikely to be robust to most substantive theories of well-being. To see why, let us reconsider the relationship between income, happiness and well-being.

I will assume that, according to certain plausible theories of well-being, income helps cause some things that constitute well-being. In addition, I will assume that, according to certain other plausible theories of well-being, happiness is caused by some things that constitute well-being (and perhaps constitutes well-being itself). The facts that income and happiness are good for people are platitudes about well-being – we know that these things are good for us without having to know what constitutes our well-being. However, in the contexts in which income and happiness come apart, these platitudes do not enable us to determine which of the two goods correlates with well-being (or, at least, correlates with well-being to a larger extent). For example, if we were to compare one poor but happy individual with one rich but unhappy individual, we would have no way of knowing which individual is better off, all other things being equal. Simply knowing that income and happiness tend to be good for people does not help us determine which good correlates with well-being to a greater extent.

In the context in which income and happiness come apart, the platitude method is likely to lead to disagreements over how successfully we can measure well-being through measuring either good. Indeed, this is what we saw in the debate over the Easterlin Paradox, discussed briefly in the previous section. On the basis of different substantive theories of well-being, some theorists claim that happiness correlates with well-being to a larger extent than income does, while other theorists claim the opposite. A list of goods derived from platitudes about well-being does not provide us with a way of resolving such dilemmas. In short, in certain contexts, theorists may reasonably disagree over which goods on the list matter for well-being – in such contexts, such a list will not be consistent with most substantive theories of well-being.

### **2.3.2 Consensus over well-being**

In the previous section, I considered a theory-neutral account of well-being that consists in a list of goods that have prudential value according to *any* plausible theory of well-being. In contrast, in this section, I will consider a method that involves constructing a list of goods that have prudential value

according to *all* plausible theory of well-being. For example, mental state and objective list theories may both maintain that pleasure caused by certain goods and activities, such as quality friendships, meaningful achievements, knowledge, virtue, etc., have prudential value. If so, these things may be on the list.

At the end of this process, we will be left with a *consensus* over what is good for people. We need not adjudicate between plausible substantive theories of well-being if there is a broadly informative overlap between those theories. That is, substantive theories of well-being may tend to agree over the kinds of things that have prudential value.

Tim Taylor (forthcoming) explicitly adopts this method in developing a theory-neutral account of well-being for the purposes of measuring well-being. He considers the areas of overlap between four kinds of substantive theories of well-being: (a) mental states theories, (b) desire-satisfaction theories, (c) Aristotelian (or nature-fulfilment) theories, and (d) objective list theories. For instance, he shows that both Aristotelian theories and mental states theories hold that well-being is constituted by pleasure caused by the development and exercise of human capabilities. He further claims that identifying the objects of a person's idealised desires is likely to lead to "a significant area of overlap with Aristotelian theories." (Taylor, forthcoming: p.8) If we consider only the four kinds of substantive theories of well-being that Taylor does, we can develop something like the following consensus of things that have prudential value: a person's (a) positive experiences of (b) actually satisfying their idealized desires towards (c) the development and exercise of human capabilities that (d) are objectively good.

Like Hausman, I think Taylor is correct to emphasise that we can develop an account of well-being without having to rely on any particular substantive theory. Discovering the areas of overlap between theories of well-being is certainly likely to be consistent with most substantive theories. However, the problem with this method is that the list of goods generated is unlikely to be broadly informative. Whereas a list of goods derived from platitudes about well-being is broadly informative but unlikely to be consistent with most substantive theories of well-being, a list of goods derived from consensus over well-being is likely to be the opposite. To illustrate this, let us again consider the relationship between income, happiness and well-being.

Recall that the platitude method included both income and happiness as things on the list, but was unable to distinguish between them when they

came into conflict. In contrast, the consensus method may not include anything as broad as happiness or income on the list. Happiness will only be on the list insofar as it is caused by (b) the actual satisfaction of idealised desires towards (c) the development and exercise of human capabilities that (d) are objectively good. Income will only be on the list insofar as it causes these things. We cannot say whether or not a rich/poor person, or a happy/unhappy person, is doing well unless we know more about how they use their income or how they receive their happiness.

In response, Taylor might reply that we can generally assume that happiness is caused by (b), (c) and (d), or that income causes (a), (b), (c) and (d). That is, Taylor could claim that (a) is likely to overlap with (b), (b) is likely to overlap with (c), and (c) is likely to overlap with (d). Thus, we can generally assume that any good (such as happiness or income) related to either (a), (b), (c) or (d) contributes towards well-being. This amounts to saying that, in practice, there is not much difference between substantive theories of well-being. Indeed, at various points, Taylor states that the prescriptions of different well-being theories only tend to come apart in atypical circumstances, such as in the case of delusional pleasures or “whimsical, idiosyncratic or culpable ones” (Taylor, forthcoming: p.10).

I think there are two problems with this response. First, it may not simply be the case that there is a significant overlap between any two plausible substantive theories of well-being. For instance, in chapter five, I will argue that, in general, happiness does not tend to correlate with how well we are doing in our long-term goals and projects. This is a large area in which certain positive experiences (i.e. positive affective states) do not overlap with the satisfaction of certain idealized desires (i.e. life-long projects). It does not seem unreasonable to suspect that similar-sized differences exist between other plausible substantive theories of well-being.

The second problem with Taylor’s claim that there is likely to be a significant overlap between plausible substantive theories of well-being is as follows. Even if there is a significant overlap between any two plausible theories, the more theories we consider, the less overlap we are likely to find. That is, even if (a) is likely to overlap with (b), (b) is likely to overlap with (c), and (c) is likely to overlap with (d), it may end up *unlikely* that (a) overlaps with (d).<sup>42</sup>

---

<sup>42</sup> For example:  $0.8 \times 0.8 \times 0.8 \times 0.8 = 0.41$ .

The result is that there may not be a significant overlap between *most* substantive theories of well-being.

In sum, the problem with the consensus platitudes method is that, in generating a list of prudential goods that all theorists can agree on, we end up with a list that is too restricted to be broadly informative.

### **2.3.3 Well-being ingredients**

In the previous two sections, I have discussed two methods for developing a theory-neutral account of well-being. The first method consists in generating a list of goods that have prudential value according to *any* plausible theory of well-being. The problem with this method is that, although the list would be broadly informative, it would not be consistent with most substantive theories of well-being. In contrast, the second method consists in generating a list of goods that have prudential value according to *all* plausible theories of well-being. The problem with this method is that, although the list would be consistent with most substantive theories of well-being, it would not be broadly informative. It seems, then, that we have a dilemma – we either generate a list of prudential goods that is informative but not consistent with substantive theories or vice versa.

In this section, I will argue in favour of a third method, which I believe manages to avoid this dilemma. It involves constructing a list of all the goods that (a) have prudential value according to any plausible theory of well-being *and* (b) tend to be causally connected with other goods that meet condition (a). For example, mental state theories maintain that certain positive experiences, such as pleasure, have prudential value. Objective list theories maintain that certain goods and activities, such as quality friendships, have prudential value. Now, insofar as pleasure and quality friendships are causally connected, both pleasure and quality friendship will be on the list.

At the end of this process, we will be left with a number of *platitudes* about what is good for people, the objects of which are *causally connected* to each other. I will refer to these goods as *well-being ingredients*. Well-being ingredients are likely to be consistent with most substantive theories of well-being. For any well-being ingredient (e.g. pleasure, quality friendships, etc.) we can maintain that it either constitutes well-being or is causally connected to a good that constitutes well-being. For example, mental state theories of well-being may be correct in maintaining that pleasure constitutes well-being, in which case quality friendships correlate with well-being insofar as they either are caused by or cause pleasure. Alternatively, objective list theories of

well-being may be correct in maintaining that quality friendships constitute well-being, in which case pleasure correlates with well-being insofar as it is either causes or is caused by quality friendships. We can make the same claims about any well-being ingredient.

The list of goods generated by this method will be more restricted than the one generated by the platitudes method and less restricted than the one generated by the consensus method. I will briefly compare and contrast these methods in turn.

In common with the platitudes method discussed above, a list of well-being ingredients consists in goods that have prudential value according to *any* plausible theory of well-being i.e. the objects of platitudes about well-being. However, in contrast to the platitude method, the list consists only in the objects of platitudes about well-being that are causally connected to the objects of other well-being platitudes.

To illustrate this difference, consider two substantive theories of well-being: a mental state theory that maintains pleasure constitutes well-being, and an objective list theory that maintains quality friendships, intimate relationships, meaningful achievements, and virtue constitute well-being. According to the platitude method, all of these goods are part of our theory-neutral account of well-being. However, let us further suppose that, although there are strong causal connections between pleasure, quality friendships, intimate relationships and meaningful achievements, there are no causal connections between those goods and virtue. In such a case, virtue would not be a well-being ingredient. Thus, according to the method outlined in this section, we cannot say that an unhappy virtuous person is doing better than a happy non-virtuous person – such a judgement would not be consistent with most theories of well-being. Although the objective list theorists may agree with this judgement, the mental state theorist would not. The only way in which a list of platitudes about well-being can be consistent with most substantive theories of well-being is if it were restricted to a list of well-being platitudes, the objects of which were *causally connected* to each other; that is, if it were restricted to a list of well-being ingredients.

In common with the consensus method discussed in the previous section, a list of well-being ingredients will include the goods that have prudential value according to *all* plausible theories of well-being. However, in contrast to the consensus method, the list will be broader in scope. It will also include the kinds of goods mentioned in the previous two paragraphs: the objects of



platitudes about well-being that are causally connected to the objects of other well-being platitudes.

To illustrate this difference, consider again the relationship between income, happiness and well-being. Recall that, according to the consensus method (outlined by Taylor), happiness will only be on the list insofar as it is caused by (b) the actual satisfaction of idealized desires towards (c) the development and exercise of human capabilities that (d) are objectively good. Income will only be on the list insofar as it causes these things. We cannot say that whether or not a rich/poor person, or a happy/unhappy person, is doing well unless we know more about how they use their income or how they receive their happiness. This is not broadly informative. In contrast, the method outlined in this section will include happiness and income as things on the list insofar as they are causally connected with the objects of other platitudes about well-being, such as pleasure, quality friendships, intimate relationships and meaningful achievements. We can say that a rich person is doing well insofar as income *tends to* be causally connected to the objects of other well-being platitudes. Likewise, we can say that a happy person is doing well insofar as happiness tends to be causally connected to such goods.

The list generated by this method is, therefore, both broadly informative and consistent with most substantive theories of well-being. It provides us with a way of resolving disagreements over how successfully we can measure well-being through measuring any particular prudential good. For example, when income and happiness come apart, we can determine how successfully we can measure well-being through measuring either good in the following way. We can investigate the causal relations between each good and other well-being ingredients. If happiness tends to be caused by (or cause) other well-being ingredients (e.g. quality friendships, intimate relationships and meaningful achievements) whereas income does not, we can conclude that happiness tends to correlate with well-being to a greater extent than income does, and vice versa.

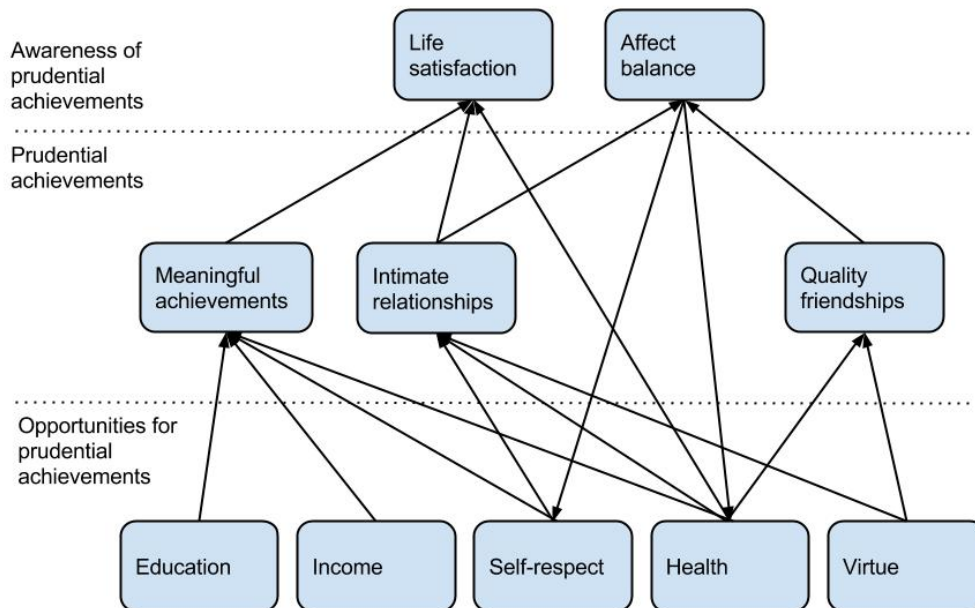
In the remainder of this thesis, I will simply refer to this understanding of well-being as the “theory-neutral account of well-being”. The theory-neutral account provides us with a list of well-being ingredients that tend to be correlated with well-being. It remains neutral over whether each well-being ingredient is either constitutively or causally related to well-being.

It is important to note that the strength of correlations and causal relations are a matter of degree. The theory-neutral account is only able to maintain that a

particular well-being ingredient (e.g. income, happiness, etc.) tends to be correlated with well-being because it tends to be causally connected to other well-being ingredients. It follows that it is only able to maintain that a particular well-being ingredient tends to be correlated with well-being *to a certain extent* because it tends to be causally connected to other well-being ingredients *to that extent*. For example, happiness may be strongly correlated with well-being insofar as it may be causally connected to a number of well-being ingredients, such as a person's quality friendships, intimate relationships, meaningful achievements, and so on. If happiness were not causally connected to a number of well-being ingredients, the claim that happiness strongly correlates with well-being would not be consistent with most substantive theories of well-being.

According to the theory-neutral account, therefore, well-being ingredients correlate with well-being to the extent that they are causally connected with other well-being ingredients. Lastly, it is worth noting that, in certain contexts, the causal connections between different well-being ingredients can break down. For example, happiness may tend to be caused by meaningful achievements, but may not be in contexts with limited information. The theory-neutral account consists merely in *general* causal connections between well-being ingredients (e.g. happiness may generally be caused by meaningful achievements), accepting that, in certain contexts, such connections can break down. Thus, if happiness were causally connected to a number of well-being ingredients, we would only be able to claim that happiness generally strongly correlates with well-being. In the following diagram, arrows represent general causal connections between well-being ingredients (whereby the direction of the arrows represent the direction of causation between different ingredients):

Fig. 2.1



We are now in a position to sum up the method outlined in this section. The theory-neutral account of well-being provides us with a list of well-ingredients – objects of platitudes about well-being that are causally connected to the objects of other well-being platitudes. For any particular well-being ingredient we can remain theory-neutral over whether it is either constitutively or causally related to well-being. For example, insofar as happiness is causally related to certain well-being ingredients, it may either constitute well-being or be caused by well-being. Either way, we can say that happiness correlates with well-being to the extent that it is causally related to other well-being ingredients – we can remain neutral over which of these ingredients in fact constitute well-being.

### 2.3.4 Summing up and looking forward

In the second section of this chapter, I showed that there are no widely acceptable substantive theories of well-being. Thus, for the purpose of investigating how successful we can measure well-being through measuring happiness, we cannot rely on any particular substantive theory of well-being. We need an understanding of well-being that is consistent with most substantive theories of well-being.

In the third section of this chapter, I considered ways in which we can understand well-being without having to rely on a particular substantive

theory. I outlined a theory-neutral account of well-being that is both consistent with most substantive theories of well-being and broadly informative. According to this account, we can understand well-being in virtue of “well-being ingredients” – list of well-being platitudes, the objects of which are *causally connected* to each other. Well-being ingredients correlate with well-being to the extent that they are causally connected to other well-being ingredients.

Despite the fact that the theory-neutral account is consistent with most substantive theories of well-being, some substantive theories are inconsistent with it. The substantive theories of well-being that are inconsistent with the theory-neutral account are simply those that entail certain well-being ingredients are correlated with well-being either to a greater or lesser extent. For example, an objective list theory may consist in the view that virtue constitutes well-being, even if virtue does not tend to be causally connected to other well-being ingredients. This would be inconsistent with the theory-neutral account.

In the remainder of this chapter, I will focus on happiness in particular. I will assume that happiness is a well-being ingredient. Happiness tends to be caused by a number of other well-being ingredients, such as the quality of one’s relationship and health.

According to the theory-neutral account, happiness correlates with well-being to the extent that it tends to be caused by other well-being ingredients. This is inconsistent with substantive theories of well-being that consist in the view that happiness is correlated with well-being either to a greater or lesser extent. For example, prudential hedonism entails that happiness perfectly correlates with well-being. But it is very unlikely to be the case that happiness tends to be causally connected to all well-being ingredients. Thus, prudential hedonism entails that happiness correlates with well-being to a greater extent than is likely according to the theory-neutral account.

For the purposes of this thesis, additional attention is merited by substantive theories of well-being that are inconsistent with the theory-neutral account in this way. Although such substantive theories are not widely acceptable (for the reasons given in the second section of this chapter) if we have some good reasons for believing that such theories may be correct, we also have some good reasons to doubt the theory-neutral account.

In the remainder of this chapter, I will argue that the substantive theories of well-being inconsistent with the theory-neutral account are implausible. I will

consider prudential hedonism in particular, in that it is the most influential substantive theory of well-being inconsistent with the theory-neutral account. Prudential hedonism maintains that happiness correlates with well-being to a greater extent than the theory-neutral account. I will argue that such a theory is implausible because it fails to adequately account for some of the things that we most care about.

I will then use a similar argument to claim that the opposite kind of substantive theory – one that maintains happiness correlates with well-being to a lesser extent than the theory-neutral account – is also implausible.

In contrast, I will argue that the theory-neutral account of well-being is likely to adequately account for things we most care about (e.g. one's level of income and educational attainment, the quality of one's health and relationships, etc.). The things we most care about typically coincide with well-being ingredients. The upshot is that the theory-neutral account is consistent with all plausible substantive theories of well-being.

## **2.4 Prudential Hedonism**

Hedonism about well-being (what I will refer to as “prudential hedonism”) and the theory-neutral account of well-being are inconsistent in the contexts in which happiness does not tend to be causally connected with other well-being ingredients. For example, a severely deluded patient may be deliriously happy despite lacking certain well-being ingredients (such as intimate relationships, health, autonomy, virtue, knowledge, etc.). In such contexts, the hedonist will maintain that the patient's happiness nonetheless (perfectly) correlates with the patient's well-being. The proponent of the theory-neutral account will maintain the opposite. In this section, I will argue that the hedonistic position is implausible.

I will argue against the plausibility of prudential hedonism as follows. First of all, I will consider its merits. I will show that the plausibility of prudential hedonism does not rest on either the acceptance or rejection of the pro-attitude requirement. In addition, I will show that hedonism has a reasonable rationale for accepting the experience requirement. However, I will then argue that, despite these merits, prudential hedonism is implausible because it does not adequately account for some of the things we most care about.

### 2.4.1 Prudential Hedonism and the Pro-Attitude Requirement

Hedonism about well-being is the view that a person's well-being is entirely constituted by their balance of pleasure over pain. According to prudential hedonism, insofar as someone is happy, they are well-off. Hedonism involves the acceptance of the experience requirement, namely that, for something to matter to our well-being, it must impact our experiential lives.

Different versions of hedonism emphasise the different kinds of experiences or mental states that matter to well-being. For 'quantitative hedonists', well-being is constituted by all experiences of pleasure and pain, which are differentiated simply by their intensity and duration. In contrast, for 'qualitative hedonists', pleasures and pains are further differentiated in virtue of their 'quality'. Further requirements may be added in determining the kinds of experiences that constitute well-being. For instance, 'veridical hedonism' is the view that only true or accurate pleasures constitute well-being. For the purposes of this thesis, I will only consider what I will call 'affective state hedonism'. Affective state hedonism is the view that a person's affective experiences constitute their well-being. According to affective state hedonism, researchers who accurately measure a subject's affective state will thereby measure the constituents of well-being (which, from now on, I will simply refer to as "hedonism").

How plausible is hedonism as a theory of well-being? First of all, it consists in the acceptance of the experience requirement, which is controversial. Many theorists have argued that our well-being is constituted by things beyond our experiential lives, thereby rejecting the experience requirement. I will discuss these arguments in detail below.

The second point worth making about the plausibility of hedonism is that, depending on how the theory is formulated, it may consist either in the acceptance or rejection of the pro-attitude requirement (Dorsey 2011b). I will refer to these different forms of hedonism as Attitude-Dependent Hedonism (ADH) and Attitude-Independent Hedonism (AIH) respectively.

According to ADH, a person's experiences of pleasure and pain constitute their well-being because (a) they are caused by the satisfactions of certain pro-attitudes, and (b) it is only the satisfactions of certain pro-attitudes *that impact our experiential lives* that constitute well-being. Thus, ADH restricts the satisfactions of certain pro-attitudes that constitute well-being to those that make a difference to a person's experiences.

According to AIH, a person's experiences of pleasure and pain constitute their well-being because they have intrinsic prudential value, and indeed are the only things that have intrinsic prudential value. All other prudential values, such as knowledge, achievement, virtue, etc., have merely instrumental value, in that they tend to bring about experiences of pleasure (or prevent experiences of pain). Thus, AIH also restricts the things that constitute well-being to a person's experiences.

The fact that hedonism can be formulated in these two different ways is a point in favour of the theory. If the acceptance of the pro-attitude requirement proves to be implausible, this does not necessarily rule out the plausibility of AIH. Conversely, if the rejection of the pro-attitude requirement proves to be implausible, this does not rule out the plausibility of ADH. In short, the plausibility of hedonism does not rest on whether or not constituents of well-being must satisfy a person's certain pro-attitudes.

However, both ADH and AIH rely on similar arguments in favour of restricting the constituents of well-being to a person's experiences. That is, both versions of hedonism accept the experience requirement and provide similar arguments in favour of doing so. In the next section, I will outline what I take to be the most plausible argument put forward by hedonists in favour of accepting the experience requirement.

#### **2.4.2 Prudential Hedonism and the Experience Requirement**

Both ADH and AIH restrict the constituents of well-being to a person's experiences. What arguments can the hedonist provide for doing this? Let us consider ADH, first. According to ADH, for something to matter to a person's well-being, it must (a) impact their experience and (b) be the satisfaction of a certain pro-attitude. There are two ways in which this may be the case. First, it may be that people ultimately only desire one thing: happiness. By "ultimately," I mean that all other desires people have are instrumental – they are desired merely as a means towards attaining happiness. If this were the case, we would desire things such as friendship, health and security solely in order to be happier (through the attainment of such things). This claim constitutes the theory of Psychological Hedonism, which I will not consider here.<sup>43</sup> Psychological hedonism has been widely refuted in both the

---

<sup>43</sup> There are, however, some key points in the theory's favour, which presumably explain its ongoing attraction. In the main, pleasure and pain may be the mechanisms we use to value certain states of affairs. Thus, we may have certain pro-attitudes towards things

philosophical and empirical literature.<sup>44</sup> I will assume that we intrinsically care about things beyond our own experiential lives.

Assuming that we intrinsically care about things beyond our experiences, ADH needs an alternative argument for restricting the satisfactions of pro-attitudes that constitute well-being to those that impact our experiential lives (Heathwood, 2006). One point in favour of ADH is that it seems plausible that the satisfactions of pro-attitudes that constitute well-being should be restricted in some way or another. Proponents of the pro-attitude requirement often refer to this issue as the “problem of remote desires.” (Parfit, 1984; Griffin, 1986) We have desires (or other kinds of pro-attitudes, such as preferences or values) that extend in time and space far beyond our awareness. Examples of remote desires include the desire for posthumous fame, the desire for a stranger to flourish, the desire for some distant future scenario (e.g. the prevention of run-away climate change), or some quirky desire whose satisfaction is epistemically inaccessible (e.g. a prime number of atoms in the universe). The satisfaction of such desires does not seem to constitute well-being.<sup>45</sup>

Well-being theorists have failed to reach any kind of agreement over how the desire-satisfactions that constitute well-being can be restricted in a way that adequately captures our intuitions about remote desires.<sup>46</sup> This leaves the

---

insofar as we take pleasure in them (though see Schroeder 2004 for a critique of this view). Alternatively put, we may only be able to value something insofar as that thing causes pleasure in us. It is perhaps for this reason that people say things such as, “I just want my children to be happy”, or “There is in reality nothing desired except happiness” (Mill 2002: chapter two).

<sup>44</sup> For a philosophical example see Rachels (2010). For an empirical example see Berridge (1996).

<sup>45</sup> Though see Lukas et al. (2010) for an exception. Lukas et al. argue that the satisfaction of every actual desire, including remote desires, promotes well-being, although they seem to readily admit that this requires that one “embrace the absurdity and simply deny the intuition that some desires are irrelevant to well-being” (Lukas et al. 2010: p.21).

<sup>46</sup> For instance, Mark Overvold argues that the only desires whose satisfaction promote a person’s well-being are those whose satisfaction require that person’s existence; remote desires do not, therefore, promote well-being (Overvold 1980). Other philosophers, however, argue that satisfaction or frustration of posthumous desires can affect a person’s well-being (Brandt 1979, Portmore 2007). Griffin distinguishes between informed satisfied desires that can and cannot count towards a person’s well-being as follows: “What counts for me, therefore, is what enters my life with no doing from me, what I bring into my life, and what I do with my life” (Griffin, 1986: p.22). I agree with Bishop (MS: chapter two) in interpreting this to mean that as long as a remote informed desire is properly connected to one’s life plan or narrative, its satisfaction promotes the



kind of restriction offered by the proponent of ADH (namely, to restrict the desire-satisfactions that constitute well-being to those that a person is aware of) as a live option.

The argument in favour of this kind of restriction has been stated clearly by Shelly Kagan (1992). (This is the same argument outlined in chapter one, in favour of a narrow concept of well-being.) Kagan provides the following rationale for the view that desire-satisfactions must impact a person's experiential life in order to constitute his or her well-being. He begins by noting that something contributes towards a person's well-being if it is good *for that person*. Kagan then claims that persons are nothing other than a body and mind. Thus, according to Kagan, something can only constitute a person's well-being if it makes a difference to their body or mind. The satisfactions of desires that do not impact on a person's body or mind, therefore, do not constitute their well-being. Such desire-satisfaction is too remote. It is only the satisfactions of desires that do impact on a person's body or mind that constitute their well-being. Being genuinely successful, or having desires fulfilled in actual fact, does not benefit *the person* except insofar as it directly impacts them i.e., through the happiness they feel when they perceive their desires to be satisfied.

Kagan further distinguishes between a person and a person's life. Whereas Kagan argues that a person is nothing more than a body and mind, he claims that things beyond a body and mind may constitute a person's life, such as various relational properties (e.g., being in a certain environment, having certain possessions, attributes, friendships, etc.). Thus, it may be one thing for a *person* to be well off, and another thing for a person's *life* to go well. According to Kagan, well-being concerns what is good for a person, not what is good for a person's life. For something to be of benefit to a person, according to Kagan, it must affect the person; it must make a difference in the person. That is, it must make a difference to the person's body or mind. Being genuinely successful, or having desires fulfilled in actual fact, may constitute a person's *life* going well, but does not benefit *the person* except insofar as it directly impacts them.

Kagan's argument provides a clear rationale for restricting the kinds of desire-satisfactions that constitute well-being. According to Kagan, it is only the

---

person's well-being. Griffin's restriction rules out some remote desires (e.g. the prime number of atoms in the universe), but not all of them (e.g. the desire for posthumous fame).

desire-satisfactions that a person is aware of that constitute his or her well-being. Kagan's argument rests on a plausible distinction between a person and a person's life, and the claim that well-being concerns the former. I will take this to be the most plausible rationale for restricting the kinds of desire-satisfaction that constitute well-being in the way that ADH prescribes.

Let us now consider AIH. In contrast to ADH, AIH rejects the pro-attitude requirement. It does not maintain that something must satisfy a person's certain pro-attitude in order to constitute his or her well-being. Rather, according to AIH, pleasure constitutes well-being because certain properties of pleasure have intrinsic prudential value.

Tomas Hurka notes that, due to its *phenomenal quality*, happiness is "the most commonly recognised intrinsic good" (Hurka, 2011: 9). He further notes that, although different theorists refer to this good in different ways, using names such as pleasure, enjoyment, contentment, satisfaction, etc., they all refer to "a good feeling with positive tone" or with a positive "buzz" (Hurka, 2011: 10). It is this phenomenal quality, or *feel*, that has intrinsic prudential value, according to AIH. Happiness is unique in this respect – no other good seems to have a similar kind of phenomenal quality. Thus, according to AIH, pleasure is the only thing of intrinsic prudential value. Other goods, such as knowledge, achievement, virtue, etc., have merely instrumental value, in that they tend to bring about experiences of pleasure (or prevent experiences of pain).<sup>47</sup>

AIH rests on restricting the kinds of goods that constitute well-being in a similar way that ADH rests on restricting the kinds of satisfactions of desires that constitute well-being. Indeed, we might call this the problem of "remote goods" – not all goods contribute towards well-being. Virtue may be morally

---

<sup>47</sup> One objection to AIH is that happiness/pleasure is not in fact unified by a common phenomenal quality. There are many different kinds of pleasant experiences. There are sensory pleasures, like the pleasure of tasting delicious food or receiving a massage. There are the pleasures caused by having our desires are satisfied, like the pleasure of winning a game or getting a promotion. I will not deal with this objection here, in that I think it is ultimately unproblematic for AIH. I agree with Crisp that, despite this variety in the phenomenal qualities of happiness, such experiences all have "a certain common quality – feeling good" (Crisp, 2006: p.109). This does not entail that we can compare the prudential value of different pleasant experience on the basis of this common quality i.e. on how good they feel. There may be different kinds of intrinsic prudential value, whereby each kind is determined by the different kinds of phenomenal qualities of pleasant experiences – we can nonetheless maintain that all these experiences have intrinsic prudential value, and that nothing other than such experiences has intrinsic prudential value.

good, for instance, but may not be good *for* someone. It is because some goods are more prudentially valuable than others that we can sacrifice our well-being for the attainment of certain (non-prudential) goods.<sup>48</sup> Mother Teresa, for example, may have lead a very good life, but not necessarily a life that was good *for her*. It seems right to say that she sacrificed her own well-being in favour of helping the poor. Her efforts to help the poor may have resulted in her cultivating virtue, but such virtue did not result in high levels of well-being. In comparison to other goods, such as enjoyment, intimate relationships, skill-development, etc., virtue does not seem to be very prudentially valuable. In short, some goods are more prudentially valuable than others. Any theory of well-being that rejects the pro-attitude requirement must provide us with an account of which goods are prudentially valuable, or at least more prudentially valuable than others.

AIH provides us with a simple answer to this question: happiness is the only prudential good that constitutes well-being. This clearly accounts for the Mother Teresa case. Being virtuous may have provided Mother Teresa with some amount of pleasure – indeed, at times, it may have resulted in a profound form of happiness – but probably not as much pleasure as she would have received from certain alternative, less admirable, forms of life.

However, is it the case that pleasure is the *only* good that constitutes well-being? What about other goods, such as meaningful achievements, intimate relationships, skill-development, and so on? I think that the most plausible rationale for pleasure being the only good that constitutes well-being is the one provided by Kagan, reviewed above. Recall that Kagan argues that well-being concerns how well a person is doing, and that a person is constituted entirely by a body and mind. Thus, for something to constitute well-being, it must make a difference to a person's body or mind (Kagan, 1992). In other words, for something to constitute well-being, it must impact a person's experiences. Pleasure may be the only thing that constitutes well-being in this respect.

### **2.4.3 The Problem with Prudential Hedonism**

We are now in a position to review Kagan's argument, and thereby the plausibility of both versions of hedonism (ADH and AIH). Perhaps the most

---

<sup>48</sup> Although most well-being theorists accept this view (Smuts, MS), see Rosati (2009) for an alternative view.

obvious counter-argument to accepting the experience requirement is that it goes against our intuitions regarding the constituents of well-being. These intuitions are best illustrated by Robert Nozick's famous "experience machine" thought experiment. Nozick (1974) asks us if we would spend the rest of our lives plugged into an experience machine. The machine will stimulate our brains such that we will experience ourselves as engaging in a lifetime of certain activities, and have no recollection of us having chosen to plug in. Most philosophers of well-being assume that, along with Nozick himself, we will not choose to plug into the machine. The reason for not doing so is that there are things besides our experiences that matter for well-being. Those who plug in would gain in pleasure or enjoyment, but they would also miss out on (a) a great deal of knowledge (since they would have so many false beliefs), (b) *really* being with their friends and family, (c) *really* accomplishing things, etc. (Hurka 2011) In short, Nozick's thought-experiment seems to show us that we care about certain goods beyond the impact that such goods have on our experiential lives.

I think that our intuitions around the experience machine reflect a major problem with prudential hedonism. The problem is that it does not do justice to the *phenomenology* of prudential value. We intrinsically care about our intimate relationships, friendships, career, education, health, etc. These are the things that guide our action – they are the things that we directly pursue in our practical lives. These things guide our action because we care about them; we see them as being good for us. For example, someone who values their children's welfare tends to take pleasure from fulfilling their parental role. Attaining a certain level of welfare for their child does not just guide their action; in addition, it is what seems to be good for them. The problem with hedonism is that it insists that such things are valuable only because of the experiences that they afford. But this is not how things seem *to those who have such experiences*. This is the fundamental tension with hedonism (Barber 2011).

In response to this problem, several theorists have recently challenged our intuitions regarding the constituents of well-being, and have criticized the utility of Nozick's influential thought-experiment in particular (Tännsjö 2007; Crisp 2006a; Kolber, 1994; Kawall 1999; Barber 2011; Baber 2008; Feldman 2012; Hewitt 2010; Weijers 2011; De Brigard 2010; Belshaw 2012; Bramble, MS; Sizer, 2010).

Many of these theorists have argued that the experience machine thought-experiment is invalid, largely due to its outlandish or unrealistic nature (Elster, 2011). Such theorists argue that we cannot trust our intuitions over whether to plug into the machine or not because such intuitions are inevitably influenced by irrelevant factors (i.e. factors other than information about what constitutes our well-being). For instance, Roger Crisp argues that people's choices over whether or not to plug in are likely to be affected by "differing attitudes to risk." (Crisp, 2006: p.635).<sup>49</sup> In a similar vein, Adam Kolber (1994), Joseph Mendola (2006) and Dan Weijers (MS) all note that we have a general fear of the unfamiliar. Perhaps most convincingly, De Brigard (2010) presents empirical evidence in favour of the hypothesis that subjects' choices over whether or not to plug in to the experience machine are largely based on the "status-quo bias". That is, insofar as subjects are averse to abandoning the life they have been experiencing so far, they will tend to choose not to plug in.

Rather than taking a stand on these issues, I believe that we can largely ignore them by considering more realistic versions of Nozick's thought-experiment. Other well-being theorists, such as Griffin (1986) and Sumner (1996), have offered more realistic thought-experiments that seemingly count against the plausibility of prudential hedonism in the same way as the experience machine. Alex Baber (2011) summarizes these kinds of thought-experiments as follows:

"You find out that the 'homemade' cookies you just ate were really baked from store-bought dough; that they have been letting you win at poker; that your lover in fact hates you; that the wilderness you have made your home contains the world's largest call centre, cunningly disguised by artificial trees; that your best friend is paid out of a trust-fund set up by your parents; that the children you have taken hard-won but profound joy from helping to raise are in fact remotely-controlled robots; and so on. In each case you are happy until you discover what a chump you have been, at which point your happiness evaporates." (Baber, 2011: p.272)

In response to these kinds of thought-experiments, people tend to have similar reactions to those that they have in response to the experience machine. That is, people tend to choose to stay in contact with reality rather

---

<sup>49</sup> See also Belshaw (2012).

than being a “chump.” This suggests that we care about being in touch with reality more than we care about being happy and deluded. We care about things beyond our experiential lives.<sup>50</sup>

However, other well-being theorists accept that we care about things beyond our pleasurable experiences, while challenging the implications that this has for the plausibility of prudential hedonism. Such theorists do not think that we need to do justice to the phenomenology of prudential value, in that, what we see as being good for us is not necessarily what *is* good for us. For instance, Harriet Baber (2008) argues that our choices over whether or not to be a chump tell us nothing about well-being unless we assume the truth of the revealed preference-satisfaction view of well-being (i.e. that the things we choose to do make us better off). Similarly, Fred Feldman (2012) notes that “what matters to us” is not the same thing as “what matters”.<sup>51</sup> The point is that the above kinds of thought-experiments are not direct arguments against prudential hedonism. They rely on the assumption that what we care about reflects what in fact has intrinsic prudential value. This assumption may not be true – it could be that the things we *think* intrinsically matter do not in fact have intrinsic prudential value.<sup>52</sup>

But why doubt that our considered preferences over what is good for us reflect what is in fact good for us? The fact that some things matter intrinsically to people seems to be good evidence of their intrinsic prudential

---

<sup>50</sup> Though, some well-being theorists have recently criticized even this conclusion. For instance, Tännsjö (2007) claims that he would not choose reality over being a chump, along with many people who choose to frequently take mind-altering drugs. Similarly, Belshaw (2012) argues that people with extremely poor standards of living (such as those facing an excruciatingly painful end of life) may choose the life of a chump rather than staying in touch with reality.

<sup>51</sup> See also Hewitt (2009) and Kawall (1999).

<sup>52</sup> There may also be things that *do* intrinsically matter that are not good for us simply because there may be things with intrinsic value that have nothing to do with our well-being. This relates to the problem of remote desires discussed above. We may care about the flourishing of a stranger, or some distant future scenario (e.g. the prevention of run-away climate change), and these things may indeed have intrinsic value. Yet, even if these things intrinsically matter, they may not have any intrinsic prudential value.

Smuts (MS) argues that our desire to stay in contact with reality reflects something that intrinsic matters, but does not have intrinsic prudential value. Indeed, he argues that the only things of intrinsic prudential value are our experiences. For reasons mentioned in the previous section, I believe this is inadequate in that it reduces well-being to a notion that is too narrow to be able to play the important role that well-being has in our practical lives.

value. Alternatively put, surely the best explanation that some things matter intrinsically to people is that they have intrinsic prudential value? Not so, according to the hedonist. Our intrinsic values may be best explained by evolutionary or psychological mechanisms, and not the fact that they are things that have intrinsic prudential value. As Chris Heathwood puts it: “It is a common mistake to attribute intrinsic value to highly reliable instrumental values. If we get pleasure every time we appreciate great art, perhaps this causes us to mistakenly judge that appreciating great art is itself of intrinsic value.” (Heathwood, MS: p.10)

According to Crisp, the things we care about are simply those that have reliably provided us with pleasure in the past. In a similar vein, Matthew Silverstein gives the following explanation of our intrinsic desire to be in contact with reality:

“[Our intuitions against prudential hedonism] reflect our desire to remain connected to the real world, to track reality. But...the desire to track reality owes its hold upon us to the role it has played in the creation of happiness. We acquire our powerful attachment to reality after finding again and again that deception almost always ends in suffering. We develop a desire to track reality because, in almost all cases, the connection to reality is conducive to happiness. Our intuitive views about what is prudentially good...owe their existence to happiness.” (Silverstein 2000: p.296).

We have, then, the following two views. According to critiques of prudential hedonism, the theory fails to account for the fact that we care about things beyond our experiential lives. This matters because we see these as being good for us. However, according to hedonists, prudential hedonism can account for the things we care about – we care about these things because they have reliably caused us pleasure (and presumably are likely to continue to cause us pleasure). It is not obvious how to decide which view provides us with the best explanation of why some things matter intrinsically to people. The explanation offered by hedonism certainly *seems* wrong to me – it seems that I intrinsically care about my loved ones, for example, *because they are valuable*. However, the hedonist does not deny that this seems to be the case; they merely contend that, on closer inspection, there is a plausible reason why this seems to be the case, and that is because my loved ones (to continue the example) have repeatedly made me happy. In the absence of any further argument against this explanation, the counterintuitive nature of prudential hedonism does not render it an implausible theory of well-being.

#### 2.4.4 Another Problem with Prudential Hedonism

In the previous section, I argued that the major problem with hedonism is that it does not do justice to the phenomenology of prudential value. We care about things beyond our experiential lives – we see certain things, such as our health and our loved ones as being good for us. In response, however, the hedonist can argue that we should not trust such intuitions. The things we see as being good for us, according to the hedonist, are simply those things that reliably cause pleasure in us; it is the pleasure that we receive from such things that is good for us, not the things themselves.

In this section, I will argue that there is another major problem with hedonism, which is related to this response. Simply put, hedonism does not provide us with an account of well-being as something that we can promote. This is because the things we care about (i.e. the things we see as being good for us) guide our action. When our experiences of pleasure and the things we care about come apart, hedonism is implausible.

In our practical lives, we cannot directly act to attain experiences of pleasure or desire-satisfaction. We can only attain such experiences indirectly. We must act to attain the *objects* or *causes* of pleasure. These objects will tend to be relational properties, rather than states of body or mind. For example, we cannot directly attain the experience of satisfying our desire for an intimate relationship. Instead, we must attain an intimate relationship, which would satisfy our desire. It is the objects or causes of our pro-attitudes that are action-guiding, not the satisfaction of the pro-attitudes themselves. As Tiberius puts it, the satisfaction of pro-attitudes “cannot be produced without attending to their objects and causes. We cannot act to secure ...desire satisfaction directly; rather, we have to bring about the conditions that give rise to them.” (Tiberius, 2007: p.382)

In our practical deliberation and evaluation, we have to focus on things beyond our experiential lives, such as our intimate relationships, friendships, career, education, health, etc. These are the things we care about. Moreover, these are the things that we think of when considering our own well-being or the well-being of others. As a result, well-being plays an important role in our practical lives. For example, we think of our own well-being when considering what career path to go down, or who to be in a long-term relationship with. Likewise, we think of the well-being of others when considering how best to help or punish them.



In restricting well-being to our experiential lives, hedonism cannot account for the fact that well-being is something that we need to be able to promote. The important role that well-being plays in our practical lives may be one of its defining features. Caregivers, developmental organisations and public policy practitioners tend to ask whether certain actions or policies will make people better off. The moral theory of Utilitarianism consists in evaluating the goodness of actions, rules or motives entirely on the basis of whether they maximize people's well-being. In short, we use the notion of well-being in a wide-range of practical deliberation and evaluation. This is not to say that well-being is the only thing that we appeal to in our practical lives – we can sacrifice our well-being, for instance, in favour of things that we deem to be more important (such as moral virtue, beauty, perfection, and so on). Nonetheless, it seems clear that well-being plays a prominent practical role.

Hedonism fails to do justice to the practical importance of well-being. With regards to the different concepts of well-being discussed in chapter one, hedonism reduces well-being to the narrow concept of well-being. This concept only refers to a person's non-relational properties, such as their experiences. As Kagan puts it, it concerns how well a *person* is doing, rather than how well a person's *life* is going (Kagan, 1992). The problem is that this is not how we *use* the notion of well-being. We use the notion of well-being in a wider sense, in a way that includes both our experiences and states of the world.<sup>53</sup>

Interestingly, Kagan explicitly notes this limitation of a narrow concept of well-being. He states the implications of his argument as follows:

---

<sup>53</sup> Note that this does not necessarily mean that well-being concerns how well a person's *life* is going, rather than how well *the person* is doing. We can take issue with Kagan's distinction between a person and a person's life. Kagan argues that a person is nothing more than a body and mind, whereas a person's life is constituted by additional relational properties, such as having certain possessions, attributes, friendships, etc. Yet, we can claim, contra Kagan, that relational properties, as well as a body and mind, also constitute a person. Another way of putting this is that Kagan may make the mistake of equivocating two difference senses of 'affecting' someone. Kagan argues that nothing can affect your well-being without affecting your intrinsic properties (i.e., your experiences). But it seems that something can affect you in the sense of being an event that you have self-interested reasons to care about (Scanlon, 1998). For example, the experience of being subtly mocked (without realizing it) does not affect your body or mind, but it does seem to affect you in another sense. That is, you (and others who care about you) generally have reason to hope that no such event actually happens to you. The notion of well-being more plausibly fits the latter sense of a person, rather than the former.

“If this is right, then the importance of well-being might be less than we often take it to be. In many cases, the pursuit of external personal goods will be far more important than the pursuit of the internal goods that happen to comprise well-being. The more narrowly we understand well-being, the more likely that this is the case. If well-being is limited in its extent, then it may also be limited in its significance.” (Kagan, 1992: p.189)

The problem with this position, however, is that well-being plays an important role in our practical lives for a reason. Well-being guides our action. A plausible substantive theory of well-being should provide us with an account of the things that are good for us that we can pursue in our practical lives. One way of putting this point is that hedonism is objectionably self-effacing. It maintains that the thing that justifies something as having prudential value (i.e. pleasure) should not play a part in our practical deliberation (Williams 2013; Keller 2013).

However, proponents of a more “sophisticated” form of hedonism may challenge these assumptions.<sup>54</sup> The sophisticated hedonist may argue that, even though we cannot pursue experiences of pleasure directly, we can indirectly pursue pleasure through the pursuit of the things that reliably cause pleasure in us. As Peter Railton notes, the sophisticated hedonist’s motivational structure should meet a *counterfactual condition*; according to Railton, “[the sophisticated hedonist] need not always act for the sake of happiness, since he may do various things for their own sake or for the sake of others, but he would not act as he does if it were not compatible with his leading [a life that reliably causes him to experience pleasure].” (Railton, 1984: 145) According to sophisticated hedonism, once we realize that certain things in our lives reliably cause us to experience pleasure then we can pursue those things directly, even though they are merely good for us instrumentally.

Thus, the sophisticated hedonist might claim that it is extremely unlikely for experiences of pleasure to become significantly divorced from the wide range of things that we consider in our practical lives (e.g. our intimate relationships, friendships, professional and academic success, health, etc.). As mentioned in the previous section, hedonists can maintain that the wide range of things

---

<sup>54</sup> I will use the label “sophisticated hedonism” to refer to the theory of prudential hedonism that is justified in the ways outlined in this section. The label is influenced by Peter Railton’s seminal paper on sophisticated consequentialism (Railton 1984).

that we consider in our practical deliberation and evaluation are the things that reliably cause us pleasure – it may be for this reason that we have come to care about them, directly pursue them, and deliberate and evaluate them.

Indeed, the sophisticated hedonist might argue that it is only in extreme cases (such as the case presented by Nozick's experience machine thought-experiment, or Hollywood films such as *The Matrix* or *The Truman Show*<sup>55</sup>) that the objects of our practical deliberation and evaluation reliably come apart from our pleasurable experiences. Thus, in reducing well-being to our experiential lives, hedonism may still be able to adequately guide our action. It is possible that the continued popularity of prudential hedonism, despite numerous critiques throughout the history of moral philosophy, is due to the fact that the ways of living that make us feel good also tend to be ways of living that we desire and aim for: helping others and engaging in meaningful activities, good relationships, and exercising our capacities.<sup>56</sup> In this way, the sophisticated hedonist can seemingly provide us with a view of well-being that accounts for the important role that well-being plays in our practical lives.

However, I do not think that a more sophisticated form of hedonism can account for the things we most care about. This is because the empirical

---

55 Indeed, some theorists have recently argued that even extreme cases, such as the experience machine, cannot replicate the hedonic consequences of being in touch with reality.

For instance, Bramble (MS) argues that the reason continued contact with real people is so essential to our well-being is that any life without it could not be hedonically adequate in the long run. In particular, he argues that a machine could not simulate one's close friends and loved ones in a satisfactory way in the long run: "Only real, conscious selves who understand and have genuine affection and concern for one can know just the right things to say, and the right ways to behave more generally (right down to subtle facial expressions and bodily gestures), in one's company, in enough of the many emotionally complex situations friends confront over the course of a life-long relationship to allow one the full range of highly pleasurable pleasures that are available through such relationships." (Bramble, MS: p.12)

Similarly, Sizer (2010) argues that subjective aspects of happiness are so closely intertwined with and attuned to objective facts about the world that pulling off the kind of a deceit portrayed in the film "*The Truman Show*" is practically impossible. She points out that we have evolved sophisticated emotion detection systems that are exquisitely attuned to subtleties in the actions and affects of others: "Even though we are not consciously aware of them, we are sensitive to automatic emotion microexpressions that can betray deception or insincerity. We perceive and react to these subtleties largely unconsciously, so their effects can be magnified and ramify throughout our interactions and relationships." (Sizer, 2010: p.157)

56 See, for instance, Rawls' "Aristotelian Principle" with regards to how exercising our capacities reliably causes us to feel good (Rawls, 1971).

claims made by the sophisticated hedonist are incorrect. The sophisticated hedonist claims that the things that we care about (i.e. the things that guide our action) *as a matter of fact* tend to reliably cause us pleasure. Thus, hedonism accounts for the important role that well-being plays in our practical lives – it maintains that we should pursue well-being indirectly through the pursuit of the things we care about.

The problem is that these empirical claims are only half right. On the one hand, the sophisticated hedonist is right in claiming that many of the things we value are things that reliably cause us pleasure (even if we do not value them for that reason). Empirical researchers have shown that happiness reliably correlates with conditions such as helping others, and engaging in meaningful activities, good relationships, and exercising our capacities (Haybron 2013). On the other hand, however, the sophisticated hedonist is wrong in claiming that all (or even most) of the things we care about are things that reliably cause us pleasure. For instance, researchers have shown that happiness does not reliably correlate with conditions such as income or education (Kahneman & Deaton 2010). Yet, we tend to care about these conditions – they are the things that guide our action. Hedonism is unable to account for this fact.<sup>57</sup>

In response, the sophisticated hedonist may object that such empirical findings do not concern happiness in the long-term. It may be that, *in the long-run*, conditions such as income or education reliably cause us pleasure. Consider, for instance, going to the dentist. Going to the dentist is unpleasant, but people still do such things because they anticipate that doing so will be better for them than allowing their teeth to decay. They anticipate that some pain now will lead to less pain later. As Daniel Hausman succinctly puts it: “There is such a thing as prudence.” (Hausman 2010: p.330) Prudence often consists in our delaying pleasurable experiences in order to invest in some future positive state. Attaining such things as higher levels of income and educational attainment are clear examples of things that require our prudence. They require us to put in (displeasurable) hard work and effort in order to attain a much greater (pleasurable) reward. Thus, the sophisticated hedonist may claim that the things we care about that require prudence will tend to reliably cause us pleasure in the long-run.

---

<sup>57</sup> For a similar argument, see Angner (forthcoming).

Note that, if this is correct, prudential hedonism may not be inconsistent with the theory-neutral account of well-being. Recall that the theory-neutral account concerns *general* causal connections between well-being ingredients. According to the theory-neutral account, happiness correlates with well-being to the extent that it tends to be caused by other well-being ingredients. Now, the sophisticated hedonist claims that happiness tends to be caused by the things we care about (at least in the long-run). Assuming that the things we care about tend to be well-being ingredients, this suggests that the theory-neutral account and prudential hedonism will tend to be consistent.

Unfortunately, however, I think there are two problems with this response. Firstly, if these empirical claims are true, hedonism may be inadequate for the purposes of measuring well-being. In order to know whether someone is well off, according to hedonism, we would need to measure that person's happiness over a long period of time. Consider, for example, the debate in the subjective well-being literature over whether having children tends to make people happy. Researchers have found that the impact that having children have on people's happiness tends to change over the life course. Very roughly, having children does not tend to make people happier when the children are growing up (i.e. when parents have to do all the work), but does tend to make people happier after they have left home (Hansen 2012; Umberson et al. 2010; Dolan et al. 2008). A similar story can be told about getting an education (Kahneman & Deaton 2010). In order to know whether having children tends to make people happier (and therefore, according to hedonism, better off) overall, researchers would need to measure people's happiness throughout their entire lives, or at least over a very long period of time. This may be practically infeasible. Yet, answering these kinds of questions (i.e. does having children, or getting an education, make me better off?) is what we want to measure well-being for.

The second problem with the sophisticated hedonist's response is that it simply may not be true that the things we care about that require prudence will tend to reliably cause us pleasure in the long-run. Consider, for example, a mountaineer who judges that it would be good for them to partake in an extremely difficult (and dangerous) climb. During the climb we can assume that their level of happiness is low – the climb requires considerable amounts of effort, risk-taking, uncomfortable conditions, and so on. Yet, despite being unhappy, we can assume that the mountaineer is satisfied, as they are predicting that they will successfully climb the mountain. Indeed, upon completion of the climb, we can further assume that the mountaineer is very

happy and very satisfied. However, it is not obvious that the total amount of happiness experienced by the mountaineer throughout the climb (or, indeed, before and after the climb) was very high. The climb may have provided them with a sense of life satisfaction, but it does not seem to have made them happy in the long-run.

It seems that we can tell a similar story about the kinds of conditions that do not tend to reliably cause pleasure in us, such as our level of income or educational attainment (Kahneman & Deaton, 2010). The things we care about that require prudence may not tend to provide us with pleasure in the long-run because the function of pleasure may not be to detect such things. Pleasure may predominantly have the function of being reliably caused by how well we are doing from moment-to-moment, with regards to our short-term goals and objectives. Completing a difficult climb may feel good, as does receiving a promotion or a degree, but getting to that stage may require a considerable amount of hard work. The function of pleasure (and displeasure) may largely be to steer us through such difficult processes, rather than to register the eventual achievement. From this point of view, we should not expect some (or perhaps many) of the things we care about to reliably cause us pleasure in the long-run.<sup>58</sup>

In sum, happiness and the things we care about can come apart. The sophisticated hedonist aims to show that, in practice, this is unlikely. In this section, I have argued that these empirical claims are false. Happiness and our long-term conditions (such as our level of income or educational attainment) may tend to come apart. In such contexts, hedonism maintains that happiness, rather than the things we care about, correlate with well-being. Yet, it is the things we care about that guide our action. Insofar we have reasons to promote our well-being, this renders hedonism an implausible substantive theory of well-being.

#### **2.4.5 What is Right and Wrong about Hedonism**

I think the underlying problem with hedonism is that it seeks to elucidate well-being at the wrong level of explanation. In chapter one, I showed that well-being plays an important role in our practical lives – in deliberating over major life decisions, for example, or in evaluating the help that we give to others. A

---

<sup>58</sup> I will consider this kind of case, as well as the distinction between happiness and life satisfaction, in more detail in chapter five.

large amount of our practical deliberation and evaluation concerns well-being: choosing the right career is good for us; effectively helping the people we care about is good for them; and so on. A plausible theory of well-being, therefore, must provide us with an account of well-being *at this level* – the level of practical deliberation and evaluation. In our practical lives, we consider the things that we care about – our relationships, our professional and academic success, our health, etc. These are the things that we see as having prudential value, and are the things that we can directly pursue. A theory of well-being that strays too far from these things will cease to be a theory about well-being.

In reducing the things we care about to their effects on our experiential lives, hedonism attempts to elucidate well-being at a level of explanation that is too abstract, too far removed from our practical lives and the things we care about. At this level of abstraction, it is not surprising that hedonism is an inadequate theory of well-being for the purposes of measuring well-being. Our practical deliberation and evaluation takes place at the level it does because that is the most effective way of promoting well-being with limited information. We do not need to know whether a particular career path will make us happier in the long-run, for example, in order to assess whether it would be good for us. We merely need to consider various attributes that it has over alternative career paths, such as more autonomy, a greater salary, a more co-operative working community, etc. To a large extent, the effective measurement of well-being will need to also take place at this level. At a greater level of abstraction contextual information is lost, and the measurement of well-being becomes more practically infeasible.

However, the above discussion also helps us to understand what is right about prudential hedonism. To a certain extent, the things we care about (and that make a difference to our well-being) are closely connected to our experiential lives. As mentioned above, the continued popularity of hedonism is perhaps largely due to the fact that a life of pleasure tends to correlate with many aspects of a life that are generally thought to be good for us (such as a life that consists in intimate relationships, good friendships, health, exercising our capacities, altruistic activities, etc.). As Laura Sizer puts it: “the subjective aspects of happiness are so closely intertwined with and attuned to objective facts about the world that...[normally], happiness is a reliable barometer of well-being.” (Sizer, 2010: p.156) In chapter three, I will suggest that we have good reason to treat happiness as an *indicator* of well-being in this respect. Hedonism may be an implausible theory of well-being, but I think it is right to

emphasise that a life of well-being will tend to be closely bound up with a life of happiness.

## 2.5 Adaptive Preferences

In the previous main section, I argued that prudential hedonism is implausible, in that it entails happiness perfectly correlates with well-being. This fails to account for the prudential value of some of the things we most care about. We see the things we intrinsically care about as having intrinsic prudential value, and we directly pursue such things in our practical lives. Prudential hedonism fails to account for the things we care about that *do not* tend to cause happiness, such as one's level of income or educational attainment.

In this section, I will use a similar argument to claim that the opposite kind of substantive theory – one that maintains happiness correlates with well-being to a *lesser* extent than the theory-neutral account – is also implausible. Such a theory fails to account for some of the things we most care about. In particular, it fails to account for the things we care about that *do* tend to cause happiness, such as the quality of one's health and relationships (Kahneman & Deaton 2010).

Why would one think the opposite to that of prudential hedonism, namely that happiness is very weakly (rather than perfectly) correlated with well-being? One reason to question the extent to which happiness correlates with well-being is the phenomenon of *adaptive preferences*. Amartya Sen (1999) and Martha Nussbaum (2000), for instance, on the basis of cases of adaptive preferences, have influentially criticised substantive theories of well-being that maintain happiness (or preference-satisfaction) constitutes well-being. Cases of adaptive preferences consist in circumstances in which people adapt to unfavourable conditions to such an extent that they are still able to be happy or satisfied. Thus, people can be happy despite having extremely poor standards of living. As Sen puts it:

“[P]eople living under tyranny may lack the courage to desire freedom, and may come to terms with the deprivations of liberty, taking whatever pleasure they can in small reliefs, so that in the scale of utility (measured either in terms of mental satisfaction, or in terms of intensities of desire), the deprivations may be muffled and muted...Tyrannies operate not just by violating freedoms, but often by making collaborators out of victims.” (Sen 2004: p.634)



The kinds of cases Sen describes all involve people who lack certain central capabilities (Nussbaum, 2001: p.139). Sen writes of “perennially oppressed minorities in intolerant communities” and “hopelessly subdued housewives in severely sexist cultures” (Sen, 1999: p.62-63). Such people lack central capabilities such as being able to be treated as a dignified human being. I will assume that such central capabilities are important well-being ingredients (that is, we tend to think of them as being good for people, and they tend to be strongly causally connected to many other well-being ingredients). The kinds of cases Sen describes aim to show that, despite the absence of important well-being ingredients, individuals are still able to be happy.

Cases of adaptive preferences suggest that happiness may be very weakly correlated with well-being. It is important to note that this, in itself, would not necessarily be inconsistent with the theory-neutral account. The theory-neutral account merely maintains that happiness tends to be correlated with well-being to the extent that it tends to be caused by (or cause) other well-being ingredients. Cases of adaptive preferences may show that happiness does not tend to be causally related to other well-being ingredients, and thereby is very weakly correlated with well-being. This is entirely consistent with the theory-neutral account developed in this chapter. As Sen claims, happiness may tend to be caused by “small mercies” rather than important well-being ingredients.

As it happens, I think that the kind of cases Sen describes do not convincingly show that happiness fails to be caused by other well-being ingredients. First of all, it is not obvious that the people Sen describes as having successfully adapted to their circumstances are in fact as happy as Sen makes out. As Haybron notes, “while it is easy to imagine people becoming resigned to oppressive circumstances, even registering satisfaction with their lives or showing the world a happy face, it is not so easy to imagine the enslaved, the solitary homeless, and the brownbeater sweatshop laborer leading *emotionally fulfilling* lives.” (Haybron, 2008: p.124)<sup>59</sup>

Secondly, assuming that the people Sen describes are happy, it seems reasonable to assume that their happiness tends to be caused by certain well-

---

<sup>59</sup> Haybron goes on to suggest that adaptation is unlikely to occur for certain kinds of affective states, such as anxiety or stress. He notes that, “if you are living under threat, and things stay that way for a long time, it makes little sense to let your defenses down and cease responding as if you are under threat. So long as they are out to get you, you had best stay prepared.” (Haybron, 2008: p.124)

being ingredients, such as the quality of their relationships and health. Researchers have shown these things to be key determinants of happiness, as well as other potential well-being ingredients, such as meaningful activity, skill development, autonomy, and so on (Diener & Biswas-Diener 2008). Insofar as the people Sen describes are happy, it seems reasonable to assume that it is because their lives contain some of these ingredients, in addition to the small mercies.<sup>60</sup>

Thus, cases of adaptive preferences may not show, as Sen is sometimes taken to claim, that happiness fails to be caused by other well-being ingredients. Happiness tends to be caused by certain well-being ingredients, such as the quality of one's relationships and health (Kahneman & Deaton 2010).

However, one could still argue that cases of adaptive preferences show the following. People may be able to adapt to their circumstances to such an extent that their happiness does not tend to be caused by *the most important* well-being ingredients, such as central capabilities. Even if we assume that happiness does tend to be caused by certain well-being ingredients, such as the quality of one's health and relationships, we can doubt the importance of these things. We could argue that, in contrast to central capabilities, health and relationships are not as important for well-being. Alternatively put, we could argue that, in contrast to the things that do *not* tend to cause happiness, the things that do tend to cause happiness are relatively unimportant.

The upshot of this argument would be that happiness tends to be correlated with well-being to a lesser extent than is maintained by the theory-neutral account. According to the theory-neutral account, happiness is correlated with well-being to the extent that it is causally connected to other well-being ingredients. This is likely to include well-being ingredients such as the quality of one's health and relationships. In contrast, according to the substantive theory of well-being in question, these well-being ingredients may not (significantly) contribute towards well-being.

The problem with this argument is similar to the problem faced by prudential hedonism, discussed above. That is, it does not adequately take into account the things we most care about. In this case, it does not take into account the

---

<sup>60</sup> I will consider the relationship between happiness, well-being and adaptive preferences in further detail in chapter four.

value of the well-being ingredients that tend to cause happiness, such as relationships, health, engagement, autonomy, etc. It claims that these things are relatively unimportant in contrast to the well-being ingredients that do not tend to cause happiness, such as one's level of income and educational attainment. The problem with such a claim is that it fails to do justice to the phenomenology of prudential value. When we care about certain things, we see them as being good for us. Alternatively put, the fact that we care about the well-being ingredients that tend to cause happiness is evidence that these things matter for well-being.

Any substantive theory of well-being that maintains the things we most care are not in fact good for us is implausible. Recall from our discussion of prudential hedonism above that hedonists have a potential reason for doubting the phenomenology of prudential value. According to hedonism, the things we most care about may simply be the things that reliably cause us pleasure. This option, however, is not available to the substantive theory in question. The well-being ingredients that tend to cause happiness are those things that such an account denies have (significant) prudential value. In the absence of any reason to go against our intuitions of what is good for us in this way, such a substantive theory of well-being is implausible.

In sum, I believe that substantive theories of well-being inspired by cases of adaptive preferences – which maintain that happiness correlates with well-being to a *lesser* extent than the theory-neutral account – are implausible. I have argued that such theories are implausible in a similar way to prudential hedonism, in that, they fail to account for some of the things we most care about. In particular, such theories fail to account for the things we care about that *do* tend to cause happiness, such as the quality of one's health and relationships.

## **2.6 The Theory-Neutral Account Revisited**

In the previous two main sections, I have considered substantive theories of well-being that are inconsistent with the theory-neutral account – those that maintain happiness correlates with well-being to either a greater or lesser extent than the theory-neutral account. I have argued that such substantive theories are implausible in that they fail to take into account some of the things we care about. It remains to be seen, however, whether the theory-neutral account of well-being fairs better in this respect. In the final main

section of this chapter, I will show that it does, and is thereby consistent with all plausible substantive theories of well-being.

### **2.6.1 The theory-neutral account and the things we care about**

The theory-neutral account maintains that well-being ingredients tend to be correlated with well-being. Well-being ingredients are goods that (a) have prudential value according to any substantive theory of well-being (i.e. are the objects of platitudes about well-being) and (b) are causally connected to other goods that meet condition (a). So far, I have implicitly assumed that the objects of well-being platitudes coincide with the things that we tend to care about. For example, the claims that income, education, relationships and health contribute towards well-being are all platitudes about well-being. We also care about these things. But do we have any reason for thinking that this will tend to be the case?

I think we do. In formulating substantive theories of well-being, philosophers cannot stray too far away from the things that people tend to care about, such as security, achieving major goals, developing and exercising your capacities, pleasure, and so on (Tiberius 2013b). This is because such theories aim to provide an adequate description of well-being (Sumner, 1996). Providing such a description requires accommodating people's intuitions about the things that are good for them. Without accommodation our intuitions in this way, it is not obvious that philosophers are providing a description of well-being, rather than a description of an altogether different notion.

The result is that the things that are good for people, according to substantive theories of well-being, tend to also be things that ordinary people tend to care about. As Haybron and Tiberius note, this result, "is not surprising given that theories of well-being are standardly defended through the method of reflective equilibrium, which relies heavily on intuitions about cases." (Haybron & Tiberius, MS: p.9) Thus, the objects of platitudes about well-being and the things people care about will tend to coincide. This means that the theory-neutral account will tend to provide us with a plausible understanding of well-being.

Note that the arguments in this chapter support the stronger claim that the theory-neutral account of well-being is consistent with *all* plausible substantive theories of well-being. I have argued that the substantive theories of well-being that are inconsistent with the theory-neutral account (in that they maintain happiness correlates with well-being to either a greater or lesser extent) are implausible. Such theories fail to account for some of the things

we most care about. In contrast, the theory-neutral account of well-being does tend to take into account the things we most care about. The substantive theories of well-being that are consistent with the theory-neutral account (in that they maintain happiness correlates with well-being to the same extent) are plausible theories of well-being in this respect.

### **2.6.2 Is the theory-neutral account a substantive theory of well-being?**

The theory-neutral account is not a substantive theory of well-being, in that it remains neutral over whether any particular well-being ingredient is either constitutively or causally related to well-being. We cannot know which well-being ingredients in fact constitute well-being without a correct substantive theory of well-being.

However, it is worth noting that a substantive theory of well-being has recently been developed, which has much in common with the theory-neutral account. Michael Bishop's Network Theory (NT) is the view that well-being is constituted by positive causal networks of feelings, attitudes, traits and outcomes (Bishop, MS). According to this view, any particular feeling, attitude, trait or outcome contributes towards well-being to the extent that it is part of a positive causal network; that is, to the extent that it tends to be causally connected to other feelings, attitudes, traits or outcomes (what Bishop calls "nodes").

For instance, according to NT, happiness contributes towards well-being to the extent that it is causally connected to certain attitudes, traits or outcomes. Thus, happiness contributes to well-being insofar as it is caused by certain outcomes, such as quality relationships, meaningful activity, health, etc. It also contributes to well-being insofar as it caused certain attitudes or traits, such as being more optimistic, sociable, having more energy, etc. The more these nodes are causally connected to other nodes, the more happiness contributes towards well-being. In short, the extent to which happiness contributes towards well-being is determined by the extent to which it is a part of positive causal networks.

NT is effectively the theory-neutral account of well-being developed in this chapter without the theory-neutral part. The theory-neutral account maintains that any particular well-being ingredient tends to be *correlated* with well-being to the extent that it is causally related to other well-being ingredients. As mentioned above, it remains neutral over whether this correlation is due to a constitutive or causal relationship. In contrast, NT maintains that any

particular node (i.e. feeling, attitude, trait or outcome) tends to *constitute* well-being to the extent that it is causally related to other nodes. For Bishop, the causal networks that bond together different aspects of our lives have prudential value. For the theory-neutral account, these causal relations merely enable us to gain a broad understanding of well-being without having to commit to a particular view about what well-being consists in.

There is, however, one difference between Bishop's substantive theory and the theory-neutral account of well-being developed in this chapter. The aim of developing the theory-neutral account was to gain a broad understanding of well-being that is consistent with most substantive theories of well-being. Thus, I focused on generating a list of goods that have prudential value according to any substantive theory of well-being – the objects of platitudes about well-being. In contrast, Bishop's aim is partly to develop a substantive theory of well-being that accommodates empirical research from the field of positive psychology. Thus, he focuses on generating a list of goods that positive psychologists study – feelings, attitudes, traits and outcomes. In doing so, Bishop may inadvertently miss out some important well-being ingredients from his theory, such as certain resources or capabilities.<sup>61</sup>

## Conclusion

In this chapter, I have argued that we can investigate the correlation between happiness and well-being without a substantive theory of well-being. There are no widely acceptable substantive theories of well-being. Yet, we need a widely acceptable understanding of well-being in order to investigate how successfully we can measure well-being through measuring happiness. I have argued in favour of a theory-neutral account of well-being, which is both broadly informative and consistent with most substantive theories of well-being.

According to the theory-neutral account, we can understand well-being with regard to a number of “well-being ingredients.” Well-being ingredients are the objects of platitudes about well-being that are causally connected to the objects of other well-being platitudes. The theory-neutral account remains neutral over whether any particular well-being ingredient is constitutively or

---

<sup>61</sup> In contrast, in focusing on substantive theories of well-being, the theory-neutral account may miss out on some of the things that contribute towards well-being that philosophical accounts, and people's intuitions about well-being, have failed to acknowledge.

causally related to well-being. For example, insofar as happiness is caused by friendship, it may be that either happiness constitutes well-being and friendship causes well-being, or vice versa. The theory-neutral account merely maintains that both well-being ingredients tend to be correlated with well-being. The extent to which any particular well-being ingredient correlates with well-being is determined by the extent to which it is causally connected to other well-being ingredients. In the remainder of this thesis, I will investigate the extent to which happiness is caused by well-being ingredients, and thereby correlates with well-being.

## Chapter 3

### Happiness as an Indicator of Well-being

#### Abstract

In this chapter, I will argue in favour of the Indicator View – the view that happiness is a defeasible indicator of well-being. According to the Indicator View, it is appropriate for us to take happiness as an indicator of well-being unless we have reason to do otherwise.

In the first main section of the chapter, I will argue that we have reason to believe that happiness tends to be reliably caused by well-being ingredients. This is because the *function* of affective states is to inform and guide our action, which they do partly by detecting evaluative features of our environment that bear on our well-being. It is reasonable to presume that affective states are not generally dysfunctional and thereby that they reliably perform this function. Thus, we have reason to believe that happiness tends to be correlated with well-being; we should, in an important range of cases, treat happiness as an indicator of well-being.

In the second section of the chapter, I will argue against a stronger claim, namely that it is appropriate to take happiness as a general indicator of well-being. We have reason to believe that happiness tends to be reliably caused by things other than well-being ingredients. Affective states partly perform their function by detecting evaluative features of our environment that do not bear on our well-being, such as moral or aesthetic features. In addition, our affective states can be systematically dysfunctional. The Indicator View should be sensitive to the empirical possibility that happiness does not tend to be generally correlated with well-being; we should treat happiness as a defeasible indicator of well-being.

In the third section of the chapter, I will consider the kinds of evidence that are required to defeat the evidence about well-being provided by happiness. I will distinguish between two kinds of defeaters, namely undercutting and rebutting defeaters. I will argue that, in practice, we need evidence of both kinds of defeaters in order to conclude that a particular defeater is present, and thereby not treat happiness as an indicator of well-being.



### 3.1 Introduction

In the previous chapter, I outlined a theory-neutral account of well-being that is both broadly informative and consistent with most substantive theories of well-being. According to the theory-neutral account, happiness correlates with well-being to the extent that it is causally connected to other well-being ingredients. This may either be because happiness is constitutively or causally related to well-being; on this, the account remains neutral. For example, we might find that happiness is caused by certain well-being ingredients (such as the quality of one's health or relationships) but not by others (such as one's level of income or educational attainment). If this were the case, happiness would correlate with certain aspects of well-being (e.g. those aspects that are either constituted by or causally related to health and relationships) but not others (e.g. those aspects that are either constituted by or causally related to income and education). We can therefore, investigate the extent to which happiness correlates with well-being by considering the extent to which happiness is causally connected to other well-being ingredients.

This is the topic of this chapter. I will argue that we have reason to believe that happiness tends to be reliably caused by well-being ingredients. According to the theory-neutral account outlined in the previous chapter, we therefore have reason to believe that happiness tends to be correlated with well-being. Insofar as we have reason to believe that happiness correlates with well-being, we should treat happiness as an indicator of well-being. I will call this the Indicator View.

The Indicator View seems fairly straightforward. It may seem obvious, for instance, to sophisticated hedonists – the main substantive theory of well-being discussed in detail in chapter two. Recall that sophisticated hedonists support their view by claiming that happiness tends to be reliably caused by well-being ingredients, such as our relationships, achievements, health, opportunities, etc. I suggested that such theorists might be right in this respect – happiness and well-being do not tend to come apart; hence the need for outlandish thought experiments to separate the two, such as Robert Nozick's experience machine thought-experiment (or the scenarios presented in Hollywood films such as *The Matrix* and *The Truman Show*).

However, despite the intuitive plausibility of such claims, I argued in chapter two that they require further justification. Indeed, there may be certain kinds

of well-being ingredients that do not tend to reliably cause happiness, such as income and education (Kahneman & Deaton, 2010). According to the theory-neutral account of well-being developed in the previous chapter, if happiness is not causally connected to certain well-being ingredients then we have reason to believe that happiness does not correlate with certain aspects of well-being (i.e. the aspects of well-being that are either causally related to or constituted by those well-being ingredients). In the contexts in which happiness is not reliably caused by certain well-being ingredients, therefore, we have reason to believe that happiness does not correlate with well-being. In such contexts, we should not treat happiness as an indicator of well-being.

Thus, the Indicator View may be less straightforward than it at first seems. On the one hand, the view should reflect the insights of the sophisticated hedonist, namely that happiness and well-being do not tend to come apart. On the other hand, the view should account for the contexts in which happiness and well-being do (perhaps significantly) come apart.

I will argue in favour of this view as follows. Firstly, I will argue that we have reason to believe that happiness tends to be reliably caused by well-being ingredients. This is because the function of affective states is to inform and guide our action, which they do partly by detecting evaluative features of our environment that bear on our well-being. It is reasonable to presume that affective states are not generally dysfunctional and thereby that they reliably perform this function. Thus, we have reason to believe that happiness tends to be correlated with well-being, and should thereby, in an important range of cases, treat happiness as an indicator of well-being.

Secondly, I will argue against a stronger claim (such as those maintained by sophisticated hedonists) namely that it is appropriate to take happiness as a *general* indicator of well-being. We have reason to believe that happiness tends to be reliably caused by things other than well-being ingredients. Affective states partly perform their function by detecting evaluative features of our environment that do not bear on our well-being, such as moral or aesthetic features. In addition, our affective states can be systematically dysfunctional. The Indicator View should be sensitive to the empirical possibility that happiness does not tend to be generally correlated with well-being; we should treat happiness as a defeasible indicator of well-being.

Lastly, I will consider the kinds of evidence that are required to defeat the evidence about well-being provided by happiness. I will distinguish between two kinds of defeaters, namely undercutting and rebutting defeaters. I will

argue that, in practice, we need evidence of both kinds of defeaters in order to conclude that a particular defeater is present, and thereby not treat happiness as an indicator of well-being.

## **3.2 Happiness indicates well-being**

In this section, I will argue that we have reason to believe that happiness tends to be correlated with well-being. Thus, we should, in an important range of cases, treat happiness as an indicator of well-being. We can reasonably infer the causal connections that exist between happiness and well-being ingredients from the *function* of affective states. I will argue that the function of affective states is to inform and guide our action, which they do partly by detecting evaluative features of our environment that bear on our well-being. It is reasonable to presume that affective states are not generally dysfunctional and thereby that they reliably perform this function.

### **3.2.1 The function of affective states**

In chapter one, I claimed that affective states represent certain things as being good or bad. That is, affective states represent evaluative features of our environment. When we feel guilty, for example, we do not merely register the fact that we have transgressed a moral rule. In addition, we register the fact that such a transgression is bad. When we feel satisfied, we both register that we have achieved a goal and that such an achievement is good. Affective states can be viewed as positive or negative insofar as they represent features of our environment that are good or bad respectively. This is the informational function of affective states – to reliably detect the things in our environment that are good or bad.

However, affective states do not merely have the function of being reliably caused by evaluative features of our environment. Although it is true that fear is associated with danger, for example, not every danger in our environment causes fear. We can say the same for sadness with loss, joy with success, anxiety with a potential threat, and so on. Affective states do not have the function of being reliably caused by every loss, success or potential threat that we experience. As Carolyn Price notes, “the loss of a hair or headache is not usually a sad loss...[sadness] is concerned with only *significant* losses or harms.” (Price 2012) The function of affective states entails that they detect significant evaluative features of our environment.

An evaluative feature of our environment is significant insofar as it is *relevant for action*. Affective states do not just provide us with evaluative information; their function is to both *inform* and *guide* us. The significant evaluative features of our environment are those that we should attend to, think about or react towards. We need to avoid/prevent the things in our environment that are bad; and we need to sustain/seek out the things that are good. Affective states should only provide us with information about the features of our environment that we need to guide our behaviour. For example, if we are hungry, it does not help to know the opening times of all the shops nearby – we only need to know about the shops that sell food. Information about shops that sell clothes, electronic products or garden equipment may be useful in others situations, but not in the context of needing to eat something nutritious. In short, affective states have the function of providing us with context-relevant information that enables us to meet our current objectives, goals and concerns.<sup>62</sup>

Some theorists of affect simply assume that the practically relevant evaluative features of our environment are those that are good or bad *for us*, rather than good or bad in some other sense (e.g. aesthetically, morally, epistemically etc.). Paradigmatic affective states, such as sadness, satisfaction, frustration, fear, disgust, interest, joy, attachment, pride, contentment, etc., seem to clearly concern our self-interests. It seems natural, therefore, to claim that the function of affective states entails that they detect prudential features of our environment (and thereby well-being ingredients<sup>63</sup>). For instance, Jesse Prinz states: “An [affective] appraisal is a representation of the relation between an

---

<sup>62</sup> One might object that the function of affective states is not merely to inform and guide action. Affective states may detect facts about mathematics, for example, or facts about what happened in the past. These kinds of states of affairs, although desired, may not be practical in that there is nothing we can do to alter whether or not the state of affairs obtains. In response, I think that such states of affairs may end up being practical in the long-term. However, even if this is not the case, I think that it is reasonable to claim that the *primary* function of affect states is to inform and guide action.

<sup>63</sup> When talking of prudential features, I remain neutral over what constitutes well-being. I take prudential features of our environment to be those things that are causally related to well-being ingredients. In chapter two, I argued that well-being ingredients are goods that are correlated with well-being, either because they are constitutively or causally related to well-being. A list of such goods is likely to include such things as intimate relationships, meaningful achievements, health, leisure, knowledge, etc. In talking of prudential features, I mean to refer to the things that are causally related to these kinds of goods. For example, making progress in one’s goals may partly cause a meaningful achievement, or an improvement in a particular well-being ingredient such as one’s health, knowledge, etc.

organism and its environment that bears on well-being.” (Prinz, 2004: p.51) Similarly, Paul Griffiths argues that basic emotions are “affect programs” – a number of correlated psychophysical reactions – which are “adaptive responses to events that have a particular ecological significance *for* the organism.” (Griffiths, 1997: p.89; emphasis added) Such theorists do not deny that affective states can end up detecting evaluative features of our environment that do not bear on our well-being. For example, when watching a film, we may experience the plight of the main character as if we ourselves were in such a situation, or as if we care about the fictional character as one of our nearest and dearest, both of which are untrue. The theorists mentioned above can account for such cases (as misrepresentations, say) while maintaining that the function of affective states entails that they detect prudential features of our environment.

Unlike such theorists, however, I do not think it is obvious that the practically relevant evaluative features of our environment are always prudential, or even typically prudential. I agree that prudential features of our environment may tend to be relevant for action. However, I believe the extent to which affect states are reliably caused by prudential features as a result of their function requires further investigation.

There are two different reasons why the function of affective states might tend to entail that they detect prudential features of our environment. Firstly, it may be the case that the practically relevant features of our environment in fact tend to be prudential. Secondly, it may be that the function of affective states entails that they are reliably caused by particular kinds of practically relevant features of our environment, and that these kinds of features tend to be prudential. I will briefly consider the first reason, and then consider the second reason in more detail.

The first argument is that the practically relevant features of our environment tend to be prudential. This argument supposes that, although there may be many evaluative features of our environment that are non-prudential (e.g. moral features, aesthetic features, epistemic features, etc.), the ones that provide us with reasons for action tend to be prudential. This may be for a variety of reasons. Perhaps, on consequentialist grounds, good overall outcomes are attained through people acting in their own interests, such as pursuing personal projects, maintaining intimate relationships, looking after their health, and so on. Alternatively, it may be that it is too demanding to respond mostly to non-prudential features of our environment. As agents, we

may need to act on the basis of things we care about and are in our interests (Williams 2013). Or, perhaps we are in a better position to see or appreciate the prudential features of our environment. The evaluative features of our environment relevant for action may simply be those that we manage to appreciate (Keller, 2013).

Although these normative arguments all have some truth to them, they are very controversial. Some ethical theories, for instance, maintain that many of the evaluative features of our environment that are relevant for action are moral, rather than prudential. Such theorists argue that we should devote a significant amount of our personal resources to helping others, being virtuous, and so on. Consider, for example, Singer's famous thought-experiment about a child drowning in a pond. Singer rightly notes that, if we were to walk past a child drowning in a pond (and no-one else were around, we were a competent swimmer, etc.) we should stop and save the child from drowning, even if it means ruining our expensive new pair of shoes. Singer then goes on to note that we are faced with a similar situation everyday with children dying of preventable diseases around the world. Singer argues that, even if helping such children is not an evaluative feature of our *nearby* environment, it is still a feature of our (albeit distant) environment that provides us with reasons for action. We should, according to Singer, be less self-interested and thereby devote more of our personal resources towards helping distant others than we tend to.

I will not attempt to adjudicate between different ethical theories here. Instead, I will focus on the second argument, namely that the function of affective states entails that they are reliably caused by *particular kinds* of practically relevant features of our environment, and that these kinds of features may tend to be prudential. Alternatively put, the argument is that affective states have been *set up* (either through evolution or culture) in such a way that they are reliably caused by the practically relevant features of our environment that bear on our well-being. This does not suppose that the practically relevant features of our environment in fact tend to be prudential. Rather, it supposes that, *historically*, the function of affective states entailed that they were reliably caused by features of our environment that tended to be prudential. This (historical) function of affective states may turn out to be normatively legitimate or illegitimate depending on the outcome of the kinds of ethical considerations noted in the previous two paragraphs (i.e. the general consequences of self-interested action, the demandingness of moral behaviour, and so on). Nonetheless, in the next section, I will argue that the

function that affective states have historically served may have entailed that they were reliably caused by prudential features of our environment.

### 3.2.2 The historical function of affective states

In this section, I will consider two mutually reinforcing arguments in favour of this view that affective states partly perform their function by detecting evaluative features of our environment that bear on our well-being. The first argument involves appealing to the evolutionarily adaptive nature of affective states.<sup>64</sup> Affective states evolved because they enabled us to survive and increase our fitness in certain ways (e.g. to form attachments, gain status, procreate, etc.). The second argument involves appealing to the culturally adaptive nature of affective states. Affective states develop through learning because they enable us to successfully live within a cultural environment (e.g. to be respectable, productive, enjoy particular activities, etc.). If we assume that a person's biological and cultural fitness and well-being ingredients tend to coincide, we have reason to believe that affective states, in an important range of cases, correlate with well-being.

Let us consider the evolutionary argument, first. It seems clear that our affective capacities evolved largely because they enabled us to be better off. Consider, for example, the case of fear. Fear reliably represents the dangers for an organism in its immediate environment, and motivates the organism to avoid those dangers. Fear increases the chances that the organism survives, thereby increasing the probability that it has lots of offspring, spreading its genes. The important point is that, in order to increase an organism's biological fitness, fear must reliably represent dangers *for* an organism.

Other evolved affective states are similar to fear in this respect. Paul Ekman's famous Big Six basic emotions (namely, happiness, sadness, fear, surprise, anger and disgust) can all be viewed as having evolved from representing features of our ancestry environment that were crucial for our biological fitness. Fear, anger and disgust, for example, all arise in contexts where life is potentially at stake (Prinz 2004a). This explains why such affective states may be universal amongst humans (Ekman 1999). Recall Griffiths' notion of

---

<sup>64</sup> Or, more precisely, "basic" affective states (that is, affective states that are, to some extent, genetically pre-determined).

“affect programs” mentioned above: it seems plausible that affective states reliably represent things in an organism’s ecology that bear on its well-being.

The idea is that our affective capacities evolved because such capacities gave us an advantage regarding our biological fitness; in order to be advantageous, our affective states must have been reliably caused by features of our ecology that were important for our survival and reproductive success. These features of our ecology may tend to have an impact on our well-being.

This argument is limited, however, to basic affective states – that is, affective states that have developed as a result of their evolutionarily adaptive nature. Affective states such as love, joy, contentment, pride, admiration, gratitude, wonder, embarrassment, shame, guilt, jealousy, envy, resentment, indignation, etc., may not be basic affective states – they may not have evolved by increasing our biological fitness; rather, such affective states may be largely a product of culture. If this is the case, we cannot claim that such affective states generally reliably represent important features of our ecology on the basis of their adaptive nature. The appeal to the adaptive nature of affective states is restricted to a small class of “basic” affective states.

In response, one could argue that most of our affective states are in fact evolutionary adaptations (and thus largely detect prudential features of our environment). Indeed, Ekman (1999) expanded his basic emotion list to include: amusement, contempt, contentment, embarrassment, excitement, guilt, pride, relief, satisfaction, sensory pleasure, and shame. The problem with this option, however, is that it may simply not be the case that most of our affective states are biologically basic. Is shame biologically basic, for example, or is it an instance of sadness re-calibrated towards one’s self or one’s identity? Similarly, is love basic, or is it a form of attachment coupled with lust? These kinds of questions can be asked of all so-called “basic” affective states (Prinz 2004b). For each kind of affective state, those who defend the claim that it is an evolutionary adaptation must show how that explanation is more plausible than the affective state developing in an alternative way (i.e., through learning). I cannot consider this issue in detail here. However, I think it is unlikely that this can be done for the large majority of affective states.

It is more plausible to presume that many of our affective states have developed through learning, rather than as a result of adaptive evolutionary processes (Prinz, 2004b). This brings us to the second argument in favour of



the view that affective states have historically been set up in such a way that they are reliably caused by prudential features of our environment. The argument involves appealing to the culturally adaptive nature of affective states. Affective states develop through learning because they enable us to successfully live within a cultural environment (e.g. to be respectable, productive, enjoy particular activities, etc.). This is effectively the evolutionary argument applied to cultural (in contrast to biological) evolution. The idea is that (non-basic) affective states have developed insofar as they reliably detected evaluative features of our environment that bear on our “cultural fitness”. Affective states have developed through learning because they have enabled us to successfully function within our cultural environment (e.g. to be respectable, productive, enjoy particular activities, etc.).

For example, guilt is plausibly a culturally developed emotion – a form of separation distress or sadness that has been re-calibrated towards moral transgressions (Prinz, 2004a: p.124-129). When we do something wrong, we feel guilty, and thereby feel motivated to make up for it. Indeed, the anticipation of feeling guilt may often prevent us from harming others in the first place. In doing so, guilt enables us to delay instant gratification and maintain our reputation as a trustworthy individual. In the long-run, having a good reputation is likely to be in our self-interest (Snow 2008; Besser-Jones, MS; Kauppinen, MS).

The idea is that our affective capacities partly developed through learning because such capacities gave us an advantage regarding our cultural fitness; in order to be advantageous, our affective states must have been reliably caused by features of our ecology that were important for our well-functioning within our cultural environment. These features of our environment may tend to have an impact on our well-being.

The appeal to both the culturally and biologically adaptive nature of affective states supports the view that affective states have historically been set up in such a way that they are reliably caused by features of our environment that tend to be prudential. This function gives us reason to believe that happiness tends to be correlated with well-being. Insofar as we have reason to believe that happiness tends to be correlated with well-being, we should, in an important range of cases, treat happiness as an *indicator* of well-being.

The Indicator View, however, is more complicated than this. In the next main section, I will argue against the claim that it is appropriate to take happiness as a *general* indicator of well-being. I will argue that the Indicator View should

be sensitive to the empirical possibility that happiness does not tend to be generally correlated with well-being; we should, therefore, treat happiness as a defeasible indicator of well-being.

### **3.3 Happiness as a defeasible indicator of well-being**

So far, I have argued that it is appropriate to take happiness as an indicator of well-being on the basis of the function of affective states. The function of affective states is to inform and guide action, which they do partly by detecting prudential features of our environment. Those impressed by this argument might be tempted to claim that happiness generally indicates well-being – a claim that has been endorsed by well-being theorists such as Martha Nussbaum (2003) and Laura Sizer (2010). In this section, I will argue that this generality claim is mistaken. I believe that we should maintain that, in an important range of cases, happiness indicates well-being, while leaving room for the empirical possibility that it does not generally do so.

I will argue for this view as follows. Firstly, I will argue that we have reason to believe that happiness tends to be reliably caused by things other than well-being ingredients. Affective states partly perform their function by detecting evaluative features of our environment that do not bear on our well-being, such as moral or aesthetic features (I will call these affective states “non-prudential affective states”). Secondly, I will argue that affective states can be systematically dysfunctional. Lastly, I will claim that the contexts in which affective states are either systematically dysfunctional or reliably detect non-prudential features of our environment are readily identifiable. The upshot is that we can treat happiness as a defeasible indicator of well-being.

#### **3.3.1 Non-prudential affective states**

In this section, I will argue that, in certain contexts, we have reason to believe that affective states may reliably detect evaluative features of our environment that do not bear on our well-being. In such contexts, happiness may not tend to be correlated with well-being.

I argued in the previous main section that the (historical) function of affective states entails that they reliably detect prudential features of our environment. This is because affective states have been set up to detect evaluative features that bear on our biological and cultural interests, which may, in turn, bear on our self-interests. However, both our biological and cultural fitness

can come apart from our self-interests. That is, in certain contexts, what is in our biological or cultural interests may not be good *for us*.

Let us consider what is in our biological interests, first. The problem with the appeal to the evolutionarily adaptive nature of affective states is that biological fitness and well-being can significantly come apart. Features of our environment that are relevant to our biological fitness may not be relevant to our well-being. The appeal to the adaptive nature of affective states assumes that what is in an organism's evolutionary interest is also in its self-interest. But this conclusion does not follow. Natural selection does not have the same "aim" as the organism. What promotes the spreading of your genes may have little to do with what is in your interest.

Having motivations that threaten your self-interest may nonetheless promote the survival and spreading of your genes. For example, it may be an advantage from the point of view of genes that you are inclined to sacrifice yourself for your offspring. But this is extremely unlikely to be an advantage from the point of view of your well-being. In short, survival and the spreading of genes do not necessarily coincide with well-being.

We can say a similar thing about what is in our cultural interests. Cultural selection also does not have the same "aim" as the organism. Particular cultural practices, norms and values may have little to do with what is in your interest. For example, it may be an advantage from the point of view of a particular culture that you are inclined to sacrifice yourself for your nation. But, again, this is extremely unlikely to be an advantage from the point of view of your well-being. Cultural practices, norms and values do not necessarily coincide with well-being.

We can accept that either our biological or cultural fitness can significantly come apart from our well-being, while maintaining that the two things often coincide. For instance, insofar as basic affective states (such as hunger, fear, disgust, satisfaction, stimulation, attachment, etc.) have evolved to help us survive and maintain a state of well-functioning, it is likely that such affective states will be reliably caused by well-being ingredients. Similarly, insofar as non-basic affective states (such as love, pride, shame, embarrassment, etc.) have evolved to help us function well within a cooperative society, it is likely that such affective states will be reliably caused by well-being ingredients. A succinct way of putting this point is that the *ultimate* function of our affective states may be to increase either our biological or cultural fitness, but this is often achieved through the *proximate* function of promoting our well-being.

Indeed, we might think that the impact culture has on the development of affective states partly “corrects for” affective states that have evolved to detect features of our environment that are no longer (or never were) in our self-interests.<sup>65</sup> For example, we may be biological pre-disposed towards violent behaviour (insofar as such behaviour increases our biological fitness). Cultural norms and rules, however, prevent these aspects of our biological nature from dominating our action (Pinker 2012). One instance of this is that children internalize values of kindness and not doing harm at a very early age as a result of instruction and pressure from caregivers (Hoffman 2001). Cultural norms and rules steer us away from our biological pre-dispositions towards ways of life in which we can live well in cooperation with others (Prinz 2007a). This shift may tend to bring our emotions in-line with certain well-being ingredients (e.g. good relationships, healthy communities, etc.).

However, even if our combined biological and cultural fitness tends to coincide with our self-interests, this may not be the case in some contexts. In oppressive cultures, for instance, people internalize norms such as “You are not good enough”, “You do not deserve to be included in the group”, “You do not deserve to participate”, and “Your voice does not matter” (Cudd 2006). These norms clearly do not promote our well-being. This is likely to be the case even if internalizing such norms means that we are able to live a better life from fitting into the oppressive culture that we are apart of.

Again, one could emphasise that this is an extreme example, and that *generally* such cases are unlikely to occur. It may be that our affective dispositions, as a whole, constrain the development of affective states that are reliably caused by things that are not in our self-interests. Indeed, this can be seen in the affective consequences of internalizing oppressive cultural norms. The individual who has internalized the norm “I am not good enough” is likely to feel a lack of competence as a result of not exercising their skills, and boredom as a result of not seeking out challenges (Besser-Jones, MS). The individual who has internalized the norm “I do not deserve to be included in the group” is likely to feel a lack of belongingness as a result of not engaging meaningfully with others (Besser-Jones, forthcoming). The point is that

---

<sup>65</sup> On this view, non-basic affective states are developmentally altered versions of basic affective states (Prinz, 2004b). For example, the emotion of guilt may be a version of sadness that has been “re-calibrated” to reliably detect a loss of moral status (as a result of violating a moral norm, say) in contrast to a loss in general. This framework entails that emotions are neither entirely biologically basic nor entirely socially constructed.

oppressive norms often stand in conflict with other well-being ingredients that *are* reliably caused by our affective states (such as a sense of competence and belongingness). Thus, the internalization of such norms may be unlikely; our self-interests may constrain the extent to which both our biological and cultural fitness can come apart from our well-being.

This is not to deny that, in certain contexts, the things that promote our biological or cultural fitness may be in conflict with well-being ingredients. In the remainder of this section, I will consider certain affective states that have been set up to detect things that may have no impact on our well-being. I will briefly outline three different kinds of affective states, namely *moral*, *empathic* and *aesthetic* emotions. The function of all three kinds of affective states entail that they detect environmental features that are good or bad, but not necessarily good or bad for us.

Moral emotions concern other-related emotions, such as resentment, contempt, indignation, admiration, sympathy and compassion. For example, the compassion we have towards another reliably represents their suffering as something that is bad. It is not obvious that such suffering concerns our own well-being. It may do so in the context of the compassion we feel towards our loved ones: for such people, our own well-being depends partly on how well they are doing. But this does not seem to be the case with the compassion we can feel towards strangers. Tappolet gives the example of seeing a stranger from a distance about to get run over by a speeding truck (Tappolet 2010: p.343). She argues that you are likely to feel fear for that person even though they are a stranger, but that your reaction does not mean that your own well-being momentarily depends on their well-being.

In contrast to moral emotions, empathic emotions concern feeling what another person is feeling. We are able to ‘catch’ other people’s emotional states through simulating a similar kind of emotional state in ourselves. Such emotions do not concern our own well-being, but rather concern another person’s well-being. Likewise, putting ourselves in “another person’s shoes” often causes affective states that represent our own well-being as views from another person’s standpoint – such emotions are likely to have no bearing on our well-being. We may gain insight into how other people’s lives are going for them, but this insight may not bear any direct relationship to our own well-being.

Lastly, aesthetic emotions may not represent features of our environment that concern our well-being. Some aesthetic emotions represent the well-being of

fictional events and characters (or perhaps represent our own well-being as viewed from within the respective fictional world) which is likely to have no bearing on how well our lives are going. Musical emotions may accurately represent certain positive and negative properties, but, again, these properties do not bear on well-being ingredients.<sup>66</sup>

In sum, moral, empathic and aesthetic emotions may reliably represent their respective objects, but those objects may not be related to our self-interests. These kinds of affective states (and perhaps others, such as epistemic, narrative, and transcendental emotions) show that the function of affective states does not solely entail that they detect prudential features of our environment.

Again, as with the example of oppression discussed above, one could argue that moral, empathic and aesthetic emotions do not *generally* play prominent roles in our lives (in comparison to clearly self-interested emotions, such as satisfaction, frustration, attachment, fear, etc.). For example, an artist who devotes all his life to his art, ignoring self-interested pursuits, such as being healthy, having good friendships and engaging in enjoyable activities, may be unlikely to sustain such a lifestyle. We may be able to say a similar thing about the moral saint who spends all of their time and resources helping others without helping herself (people like Mother Teresa may be the exception to the rule). The idea is that, if our actions become significantly divorced from our self-interests then it may be unlikely that we will be able to function reasonably well (both within our biological limits and our particular culture).

I think the extent to which this is the case is largely an empirical question, and I will not attempt to answer it here. It may be that, in general, it is unlikely that our affective states reliably detect things that are divorced from our self-interests. However, it may also be that seemingly exceptional cases, such as the devoted artist or the moral saint, are more common than we tend to think (how about single parents?). The point I want to make here is that the function of affective states does not solely entail that they detect prudential features of our environment.

---

<sup>66</sup> Of course, the pleasure that we receive from aesthetic appreciation may contribute towards our well-being, but this is a more general point about the relationship between pleasure and well-being. The point I want to make here is that the objects of aesthetic appreciation themselves (e.g. the beauty of a sunset, the interest of a piece of art, etc.) do not necessarily contribute towards well-being.

### 3.3.2 Dysfunctional affective states

In this section, I will argue that, in certain contexts, we have reason to believe that affective states may systematically fail to detect evaluative features of our environment that bear on our well-being. In such contexts, as with the kinds of affective states discussed in the previous section, happiness may not tend to be correlated with well-being.

The two cases I will consider concern affective states that have developed as a result of promoting our biological or cultural fitness respectively. I will assume that, in such cases, our biological and cultural fitness tends to coincide with well-being ingredients. The problem is that, even if the affective states developed as a result of reliably detecting prudential features of our environment, they may cease to do so in certain contexts or under certain conditions.

Let us begin by considering affective states that evolved because they reliably detected prudential features of our environment. The problem is that, in our present environment, our evolved affective states may be no longer adaptive. Our basic affective states may no longer be attuned to features of our ecology that are important for our survival and reproductive success. This is because the environment in which such affective states evolved may not be sufficiently similar to our present environment.<sup>67</sup>

Consider, for example, the affective state of hunger. Hunger represents that we are soon-to-be-undernourished and motivates us to avoid that state, namely through seeking out and consuming certain kinds of food (in particular, energy-high foods, such as sugars and fats). In our adaptive environment, it is likely that food was scarce. Thus, the affective state of hunger may have evolved to represent that we are soon-to-be-undernourished *within a food-scarce environment*. In our present environment, food may not be so scarce – indeed, it may be abundant. Thus, our hunger may unreliably represent that we are soon-to-be-undernourished. In our present environment, hunger may often represent that we are soon-to-be-undernourished when in fact there is plenty of food available.

More generally, it is possible that many failures of self-control may largely be accounted for by differences in our adaptive and present environment (Offer

---

<sup>67</sup> In the context of evolutionary psychology, this is often referred to as “mismatch theory.” (Nesse 2004)

2006). In our adaptive environment, resources were scarce, and we had to grab them when they came along. In contrast, in our present environment, resources are abundant, and we must delay gratification in order to accrue resources for the future.<sup>68</sup> The point is that basic affective states evolved to represent features of the environment in which they evolved. This kind of environment is likely to be significantly different from the kind of environment in which we find ourselves now.

In response, one could argue that, although this may be the case in certain contexts, it is unlikely to *typically* be the case. If it were typically the case that our evolved affective states were no longer attuned to important features of our ecology, affective states would not play the role that they tend to play in our practical lives. We rely on our affective states to both inform and guide our behaviour. If we feel sad, for example, we represent the loss of something we care about *and we are motivated* to avoid similar losses in the future. If our affective states consistently misrepresented the world, we would also consistently fail to function in a normal manner.

Indeed, this may be exactly what happens with certain psychological disorders, such as depression and mania. In depression, things are often represented as being much worse than they in fact are. In mania, things are often represented as being much better than they in fact are. Both conditions have damaging effects on our normal functioning. As another example, Damasio (2005) shows that individuals with damage to the affective areas of the brain have radically non-adaptive behaviour. Such individuals are unable to follow a schedule, do not seem to be able to learn from mistakes and are unable to make simple decisions.<sup>69</sup> This suggests that, at least typically, our affective states are attuned to important features of our ecology. Insofar as most people tend to function reasonably well, we can infer that most people's affective states reliably detect prudential features of their environment.

This response is reasonable to the extent that most individuals do not suffer from the serious problems of agency (of the kind that Damasio describes) that

---

<sup>68</sup> As Haybron notes, we may not just have a "sweet tooth", but also a "stuff tooth" or a "status tooth" (Haybron, 2008: p.245)

<sup>69</sup> Damasio emphasizes that the subject's theoretical understanding of their various circumstances are normal, but their sense of their importance is lost, rendering them unable to behave in ways that are in their interest and unable to make sound decisions or, in some cases, to make decisions at all.



would be expected if their affective states were generally dysfunctional. However, the problem with this response is that it ignores the extent to which a typical individual may fail to function well in certain contexts as a result of dysfunctional or inappropriate affective states. I suggest that this may occur to a substantial extent.

This is perhaps best illustrated by affective states that have developed through learning. Such affective states are not dysfunctional as a result of significant mismatches between our evolutionary and present environment. However, they may tend to be dysfunctional as a result of mismatches between our present environment and the environment in which they developed. Our affective dispositions may fail to adapt to changes in our environment. This is in part possible because our affective states are, to a certain extent, cognitively impenetrable (Faucher & Tappolet 2008). The cognitive impenetrability of affective states can render them to be largely immune to change. This is clearly seen with recalcitrant emotions. Upon walking on a glass floor high above the ground we may believe that we are safe, but still feel fear. Similarly, we may have phobias concerning all sorts of things (from spiders, to wide open spaces, to clowns) that we know on a cognitive level are not dangerous or threatening. Fear responses may often detect danger, but in these kinds of cases they do not. Moreover, we cannot always change such responses via our cognitive states, such as our judgements or beliefs. It is possible, therefore, that affective states that developed to successfully inform and guide our behaviour no longer do so within our present environment.

Consider, for example, how our circumstances can change to be incongruent with the circumstances in which our affective responses were learnt. In child development, feelings of separation distress often reliably detect being vulnerable in some way (such as not having enough food). However, as the child develops, the same signs of vulnerability (i.e., a lack of a caregiver) may not tend to be reliably caused by actual vulnerability. Children need to learn through their experiences that such signs no longer concern their security. Healthy development largely consists in become emotionally independent in this way. This is, however, an extremely difficult process, the failure of which is often at the heart of therapies such as psychoanalysis. These kinds of therapies often stress the fact that emotional responses that were effective in the past no longer have the same utility that they once had.

In addition, our affective states may develop to be overly responsive to certain features of our environment. We have a well-documented “negativity bias”, whereby we focus more on potential threats and dangers than on positive values (Kahneman, 2011).<sup>70</sup> We are also prone towards focusing on the short-term consequences of particular events and actions, in contrast to the long-term consequences (Rachlin 2004). Both of these biases may have a huge impact on our behaviour. Our affective states may be more in-tune with short-term threats than our long-term opportunities. Moreover, from an evolutionary or early developmental perspective, these biases make sense. It makes sense to focus on the possibility of negative events (such as going starving or being abandoned) as such events could have been life threatening. And it makes sense to focus on the short-term (such as eating as much food as possible now, rather than saving some for later) as we largely depended on such timeframes. The problem is that our present environment may be largely risk free and may contain many evaluative features related to our future. A negativity bias and a short-term bias may have been advantageous in either our evolutionary or developmental history, but that does not mean that such biases continue to result in the reliable detection of certain well-being ingredients.

As a final example, our personality traits make a significant difference to the affective states that we experience; a person’s personality traits are one of the largest determinants of a person’s level of measured happiness and are, to a certain extent, immune to change (Diener & Biswas-Diener 2008). Two people with different kinds of personality traits (an extrovert and an introvert, say) may have similar circumstances yet detect different evaluative features of those circumstances. In doing so, both individuals lose out on important information about certain well-being ingredients (in the same way that people lose out on certain kinds of information due the negativity and short-term biases). These differences may seem unimportant in certain situations (such as deciding whether to socialize on a particular evening) but may add up over

---

<sup>70</sup> To make things more complicated, psychologists have also documented that people tend to have a “positivity bias”, whereby most moments are experienced as being positive. The interplay between these two biases seems to be as follows. With regards to the negativity bias, negative events are experienced as being bad to a much greater extent than corresponding positive events are experienced as being good. With regards to the positivity bias, most events are experienced as being positive. As Barbara Fredrickson succinctly puts it: “whereas negativity dominates positivity in intensity, positivity dominates negativity in frequency.” (Fredrickson 2013)

time to have a substantial impact on our lives (e.g. having developed or not developed good friendships) (Nettle 2007).

The point is that, even if our affective states inform and guide our action, they may cause us to systematically act in ways that are not in our self-interest. Indeed, we cannot infer from the fact that affective states developed to successfully inform and guide our behaviour that they currently fulfil that purpose. We may spend large amounts of time trying to avoid being vulnerable and trying to avoid other immediate threats, for example. As a result, we may consistently misrepresent certain prudential features of our environment: the ways in which we are socially secure, the fact that we have various long-term opportunities, and so on. In effect, these systematic misrepresentations can be viewed as mental disorders that tend to be pervasive and universally shared. They do not prevent us from leading normal lives, but they may well prevent us from functioning reasonably well. It is this viewpoint that inspires much Buddhist (as well as Stoic) philosophy, whereby our emotions often lead us astray from reality and well functioning. In sum, even if our affective states have been set up (through evolution or learning) to detect evaluative features of our environment, they may no longer reliably do so in our present environment.

To sum up this section, I have argued that, although affective states have the function of informing us and guiding our actions, this does not mean that (in our present environment) they typically succeed in performing this function. Intuitively, there are certain contexts in which we think that affective states fail to reliably detect evaluative features of our environment. We do not use our happiness as a proxy for how well our financial accounts are going, for example. Nor do we use happiness as a proxy for the well-being of mental patients, such as manic depressives. In this section, I have suggested that such contexts may in fact be widespread. There may be many contexts in which we would not use happiness as a proxy for well-being. These contexts may add up; so much so that it would be imprecise and uninformative to claim that happiness is generally functional.

### **3.3.3 Refining the Indicator View**

In the previous two sections, I have argued that, in certain contexts, we have reason to believe that happiness does not tend to be reliably caused by well-being ingredients. In such contexts, affective states may either tend to be systematically dysfunctional or tend to detect non-prudential features of our

environment. These contexts may add up. Thus, it is inappropriate to treat happiness as a general indicator of well-being.

In this section, I will consider how we should refine the Indicator View in light of these contexts. One might think that the existence of such contexts renders the Indicator View untenable. For, how can we tell that information about happiness tells us about our well-being, rather than some other feature of our environment? Perhaps our affective states are largely dysfunctional, or generally detect non-prudential features of our environment?

I do not think that this is a problem for the following reason. For any given context, we can identify whether affective states are likely to be prudentially relevant or not. For example, how people feel during moral or political debates may not significantly bear on their well-being – the function of their emotions in such contexts may entail that they detect moral features of their environment, which do not necessarily coincide with prudential ones. In contrast, how people feel during a recession may significantly bear on their well-being – the function of their emotions in such contexts may entail that they detect prudential features of their environment, such as the likelihood of employment for themselves and their loved ones.

We can say a similar thing about dysfunctional affective states. Findings from psychology and cognitive science can inform us about the kinds of affective states that are likely to be dysfunctional and those that are not. As mentioned in the previous section, there are certain contexts in which we intuitively think that affective states fail to detect evaluative features of our environment. We do not use our happiness as a proxy for how well our financial accounts are going, for example. Nor do we use happiness as a proxy for the well-being of mental patients, such as manic depressives. This is presumably because, within such contexts, our affective reactions suffer from a particular lack of information or certain cognitive biases. In contrast, we do use happiness as a proxy for how well we are doing in our day-to-day activities, such as our progress in work activities that provide us with instant feedback, or the maintenance of our close relationships.

In short, for any given context, we can identify whether affective states are likely to be reliably caused by well-being ingredients or not. We can do this by considering the nature of the context in question. Affective states are representational states with a particular kind of content: sadness represents loss; satisfaction represents the achievement of a goal; fear represents danger; and so on. The content of an affective state tends to be determined

by the context in which it is caused. We tend to experience sadness, for example, in a context in which we have experienced a significant loss. In order to have such an experience, we need to have a reasonable amount of information about our situation, and to not suffer from any particular cognitive biases. We can, then, identify whether any particular affective state is likely to be prudentially relevant by considering the nature of the context in which the affective state was caused.

Once we have identified the kinds of contexts in which we have reason to believe that happiness does not tend to be reliably caused by well-being ingredients, we can treat happiness as an indicator of well-being outside of those contexts. It is for this reason that we should treat happiness as a defeasible indicator of well-being.

In the remainder of this chapter, I will consider the kinds of evidence that are required to defeat the evidence about well-being provided by happiness. I will call these kinds of evidence, “happiness defeaters”.

### **3.4 Happiness defeaters**

In the previous main section, I argued that it is appropriate to treat happiness as a defeasible indicator of well-being. That is, according to the Indicator View, we should treat happiness as an indicator of well-being unless we have reason to do otherwise. In the final section of this chapter, I will consider the kinds of evidence that are required to defeat the evidence about well-being provided by happiness. I will distinguish between two kinds of defeaters, namely undercutting and rebutting defeaters. I will argue that, in practice, evidence of both kinds of defeaters must be present in order to not treat happiness as an indicator of well-being.

#### **3.4.1 Undercutting and rebutting defeaters**

An undercutting defeater undermines the evidential connection between an indicator (in this case happiness) and the thing indicated by it (in this case well-being). In the first main section of the chapter, I argued that we have reason to believe that happiness tends to be correlated with well-being on the basis that the function of affective states entails that they tend to detect prudential features of our environment. In the second main section of the chapter, however, I showed that this evidential connection is undermined by two different kinds of evidence. The first kind of evidence concerns contexts in which the function of affective states entails that they detect non-prudential

(e.g. moral, aesthetic, epistemic, etc.) features of our environment. The second kind of evidence concerns conditions in which affective states are dysfunctional. Both kinds of evidence provide us with reason to believe that, in certain contexts, happiness does not tend to be correlated with well-being.

In contrast to an undercutting defeater, a rebutting defeater is a superior form of evidence that supports contrary beliefs about well-being to those supported by happiness. Thus, rebutting defeaters are distinct from undercutting defeaters in that they do not directly concern the evidential connection between happiness and well-being. Rather, rebutting defeaters provide us with reason to believe that happiness is not correlated with well-being because it is in conflict with other (more reliable) indicators of well-being.

In the remainder of this section, I will consider potential rebutting defeaters. I will argue that well-being ingredients that do not tend to be causally connected to happiness, such as a person's level of income or educational attainment (Kahneman & Deaton 2010), may be rebutting defeaters.

In chapter two, I argued that happiness correlates with well-being to the extent that it is causally connected to other well-being ingredients. For example, happiness tends to be caused by various well-being ingredients, such as good health and quality relationships (Kahneman & Deaton 2010). To this extent, happiness correlates with well-being. This may either be because happiness constitutes well-being (and therefore good health and quality relationships tend to cause well-being) or because happiness is caused by well-being (and therefore good health and quality relationships tend to constitute well-being). Either way, according to the theory-neutral account of well-being outlined in chapter two, happiness correlates with well-being to the extent that it is causally connected to well-being ingredients.

This point is important in the contexts in which two well-being ingredients come into conflict with each other. In such contexts, the well-being ingredient that is causally connected to other well-being ingredients *to a greater extent* is more likely to be correlated with well-being. Consider, for example, the finding that happiness is not causally related to income in the long-term (Easterlin, 2010). It is possible, therefore, that there are rich people who are unhappy and poor people who are happy. How comparatively well off are such people? Are the unhappy rich better off than the happy poor, or vice versa? Assuming that there are no undercutting defeaters present, there are two broad ways in which we could interpret such a context: we may believe that income correlates with well-being to a greater extent than happiness, or

vice versa. In chapter two, I argued that we should justify our beliefs in such contexts on the basis of the causal connections between the well-being ingredients in question and other well-being ingredients. If happiness is causally connected to other well-being ingredients to a greater extent than income is then we should believe that happiness is more likely to be correlated with well-being. We should believe the converse insofar as income is causally connected to other well-being ingredients to a greater extent than happiness is.

Let us assume, for the sake of argument, that income is causally connected to other well-being ingredients to a greater extent than happiness is. If this were the case, in the contexts in which happiness and income conflict, income can be viewed as a rebutting defeater. That is, in such contexts, the evidence provided by income gives us reason to believe that happiness is not correlated with well-being.

In practice, empirical researchers seem to implicitly assume that there are many kinds of evidence about well-being that are rebutting defeaters to the evidence provided by happiness. For instance, Stiglitz et al. (2009), in their influential report for the *Commission on the Measurement of Economic Performance and Social Progress* on measures of well-being beyond GDP, listed the following eight goods as indicators of well-being:

- material living standards
- health
- education
- personal activities including work
- political voice and governance
- social connections and relationships
- environment
- insecurity of an economic as well as physical nature

These kinds of lists seem to rest on the implicit assumption that the goods that make up the list are the well-being ingredients that are most causally connected to other well-being ingredients. Insofar as happiness is not one of the goods on the list, when happiness and any of the above goods come into conflict, the implication is that each good on the list is a rebutting defeater of the evidence provided by happiness.

In sum, rebutting defeaters provide us with reason to believe that happiness is not correlated with well-being because it is in conflict with other (more reliable) indicators of well-being. Rebutting defeaters may include indicators of well-being, such as a person's level of income or educational attainment. In

order to defeat the evidence provided by happiness, these indicators must provide us with evidence of well-being ingredients that are causally connected to other well-being ingredients to a greater extent than happiness is.

### **3.4.2 Evidence of happiness defeaters**

The presence of either an undercutting or rebutting happiness defeater provides us with reason to not treat happiness as an indicator of well-being. However, we do not always know whether or not a happiness defeater is present. In this section, I will consider the evidence of happiness defeaters that is required in order for us to not treat happiness as an indicator of well-being.

Let us consider undercutting defeaters, first. The presence of an undercutting defeater within a particular context means that happiness provides us with no evidence about well-being within that context. However, in practice, we do not tend to have straightforward evidence that an undercutting defeater is present within a particular context. Evidence of undercutting defeaters within a particular context, such as evidence of a lack of information or certain cognitive biases, does not tend to *completely* undermine the evidential connection between happiness and well-being within that context. In order to do that, we need evidence that the undercutting defeaters are widespread and pervasive. Having a cold, for example, may tend to make me focus on more negative events and pay less attention to certain positive aspects of my well-being. Yet, the extent to which this undermines the evidential connection between my happiness and my well-being is not obvious. If something really bad happened to me, such as the loss of a loved one, it is unlikely to go unnoticed because of my temporary illness. We need further information to conclude that an undercutting defeater is present.

We can say a similar thing about rebutting defeaters. The presence of a rebutting defeater with a particular context means that the evidence about well-being provided by happiness within that context is simply outweighed by the evidence provided by the defeater. Again, however, we do not tend to have straightforward evidence that a rebutting defeater is present within a particular context. Evidence of rebutting defeaters within a particular context, such as well-being ingredients that do not correlate with happiness, does not obviously outweigh the evidence about well-being provided by happiness. In order to do that, we would need to know that these superior forms of evidence are not themselves prone to any particular undercutting defeaters, such as



certain cognitive biases or scope effects. We need to know this information in order to conclude that a rebutting defeater is present.

In practice, I think we need evidence of both kinds of defeaters in order to conclude that a particular defeater is present. That is, we need evidence of both rebutting defeaters and undercutting defeaters in order to not treat happiness as an indicator of well-being.

Consider how evidence of rebutting defeaters is needed to confirm the presence of certain undercutting defeaters. For example, we may think that our bias towards the short-term (our “myopia” for the future) undermines the evidence provided by happiness about our long-term well-being. In order to conclude that such an undercutting defeater is present, we need evidence of certain rebutting defeaters. For instance, we might find that happiness does not correlate with our more long-term well-being ingredients, such as our financial security or our meaningful relationships. If this were the case, we can reasonably conclude that evidence of our bias towards the short-term defeats the evidence of well-being provided by happiness.

Alternatively, consider how evidence of undercutting defeaters may confirm the presence of certain rebutting defeaters. For example, we may have evidence that income rebuts the evidence provided by happiness about our well-being (that is, measures of income may correlate with measures of other well-being ingredients to a greater extent than measures of happiness). In order to conclude that such a rebutting defeater is present, we need evidence of certain undercutting defeaters. For instance, we may find that happiness is prone to various cognitive biases that prevent it from being reliably caused by certain important well-being ingredients, such as our achievements of long-term goals or our opportunities for future well-being. If this were the case, we can reasonably conclude that evidence of income defeats the evidence of well-being provided by happiness.

It is in these ways that evidence of both rebutting and undercutting defeaters can enable us to determine the kinds of contexts in which we have reason to believe happiness does and does not tend to correlate with well-being. Identifying these contexts is an exciting topic for both further philosophical and empirical research.

### **3.4.3 Potential evidence of happiness defeaters**

So far, I have claimed that the evidence provided by happiness does not justify propositions about well-being when we have evidence that both kinds

of defeaters are present. What we are justified in believing about well-being in a particular context depends on the total amount of evidence available within that context. When we have evidence of both undercutting and rebutting defeaters, we cannot justify propositions about well-being on the basis of information about happiness.

But what are we justified in believing about well-being when we don't have evidence of such defeaters? Is it simply the case that, without evidence of both kinds of defeaters, we are justified in forming beliefs about well-being on the basis of information about happiness?

I have already argued that we have reason to believe that happiness tends to be correlated with well-being, and therefore that it is appropriate to treat happiness as an indicator of well-being. This suggests that, when we don't have evidence of any defeaters, we have no reason *not* to believe that happiness is correlated with well-being; thus, without evidence of any defeaters, it seems that we are justified in forming beliefs about well-being on the basis of information about happiness.

However, one could be sceptical about this claim. We may have reason to believe that there are other defeaters out there, only ones that we do not have evidence of yet. According to this objection, we can only be sure about justifying of our beliefs about well-being from the evidence provided by happiness when we know that we have considered all or most of the potential evidence of defeaters.

For example, it may be that, as we gain a better understanding of the heuristics and biases that influence our cognition, we may come to realise that there are a large amount of contexts in which our affective states are dysfunctional, and thereby fail to reliably detect well-being ingredients. These heuristics and biases may provide us with reason to believe that, within the contexts in question, happiness does not correlate with well-being. Such defeaters may currently exist, only we do not yet know about them.

Perhaps, then, we cannot treat happiness as an indicator of well-being until we have a greater understanding of possible defeaters? I think such scepticism is unwarranted. We can incorporate the fact that the evidence provided by happiness can be defeated in certain contexts without limiting the extent to which information about happiness justifies beliefs about well-being outside of those contexts. Once we *identify* and *control for* the contexts or conditions in which the evidence provided by happiness can be defeated, it

seems plausible to claim that we can infer information about well-being from information about happiness.

The problem with the sceptical view is that it limits the amount of progress that can be made in the measurement of well-being. Empirical research will always take place under the limitation of imperfect knowledge (Chang 2004); yet, this need not paralyze us from making judgements about well-being from different kinds of evidence about well-being. If we have reason to believe that happiness tends to be reliably caused by well-being ingredients, we have reason to believe that happiness tends to be correlated with well-being. When there is no counter-evidence available, there is at least some kind of weighting in favour of treating happiness as an indicator of well-being. This is required if we are to make progress in the measurement of well-being. Thus, I believe we need evidence against the correlation between happiness and well-being in order to defeat a rational presumption in its favour.

There is, however, an element of truth in the more sceptical view. The general weighting in favour of happiness providing us with information about well-being may be generally weak in light of the potential to discover additional defeaters. Simply put, the evidence provided by happiness should, perhaps, be taken with a pinch of salt. However, this is a different position to the one that maintains there is either no weighting in favour (or even a weighting against) the epistemic merit of evidence provided by happiness in the absence of evidence of defeaters. Such a position is unwarranted. We can see that our affective states generally enable us to function well, except in certain contexts and conditions. Thus, even if we do not know the limits of how well it enables us to function, we can at least set a weighting that is somewhat in favour of the epistemic value of happiness rather against it.

## **Conclusion**

In this chapter, I have argued in favour of the Indicator View – the view that it is appropriate for us to treat happiness as a defeasible indicator of well-being. I have argued that the function of affective states is to inform and guide our action, which they partly do by reliably detecting evaluative features of our environment that bear on our well-being. This provides us with reason to believe that happiness tends to correlate with well-being, and thereby indicates well-being in an important range of cases.

Happiness is a defeasible indicator of well-being because, in certain contexts, we have reason to believe that happiness does not tend to correlate with well-

being. These are the contexts either in which affective states are dysfunctional or in which the function of affective states entails that they detect evaluative features of our environment that do not bear on our well-being (i.e. undercutting defeaters). Alternatively, there may be contexts in which we have direct evidence that happiness does not tend to be causally connected to certain well-being ingredients (i.e. rebutting defeaters).

In the next two chapters, I will refine the Indicator View by further considering the function of affective states. In chapter four, I will show that, although the function of affective states entails that they detect well-being ingredients, things are actually more complicated. I will argue that the function of affective states entails two things in this respect, namely to detect: (a) *unexpected changes* in well-being ingredients and (b) *levels* of well-being ingredients. In chapter five, I will show that, the function of affective states entails that they specifically detect (unexpected) *local changes* in well-being. Happiness, therefore, is appropriately treated as a defeasible indicator of local changes in well-being.

## Chapter 4

### Happiness as an Indicator of Changes in Well-being

#### Abstract

In this chapter, I will refine the Indicator View outlined in chapter three. I will argue that happiness is a defeasible indicator of *changes* in well-being. According to this refined version of the Indicator View, it is appropriate to treat happiness as an indicator of changes in well-being unless we have reason to do otherwise.

Again, the argument in favour of this view is based on function of affective states. The function of affective states is to inform and guide action. I will argue that the practically relevant evaluative features of our environment tend to be changes in well-being ingredients. This provides us with reason to believe that happiness will tend to correlate with changes in well-being. Thus, we can treat happiness as an indicator of changes in well-being.

In addition, I will argue that, in certain contexts, we may have reason to believe that happiness does not correlate with changes in well-being. In such contexts, the practically relevant evaluative features of our environment tend to be levels of (in contrast to changes in) certain well-being ingredients. Within such contexts, we have reason to believe that happiness will tend to correlate with certain aspects of our level of well-being; thus, in such contexts, we should not treat happiness as an indicator of changes in well-being.

Lastly, I will show how this refined view of the function of affective states can enable us to interpret empirical findings from the study of happiness. In particular, I will show that the Indicator View provides us with plausible interpretations of the Easterlin Paradox and the phenomenon of adaptation.

## 4.1 Introduction

In the previous chapter, I argued that the function of happiness partly entails that affective states detect well-being ingredients. In this chapter, I will refine this view. I will argue that the function of affective states entails that they detect *changes* in our well-being ingredients.

I will argue in favour of this refined version of the Indicator View as follows. In the first main section of the chapter, I will review three philosophical theories of the function of affective states, namely the views of Elijah Millgram, Timothy Schroeder and Peter Railton.

In the second main section of the chapter, I will show that, from these three views, we can establish two main claims. First, the function of affective states entails that they detect (unexpected) changes in well-being. This claim is endorsed by both Schroeder's and Railton's views and is consistent with Millgram's view. Second, in certain contexts, the function of affective states entails that they detect levels of certain well-being ingredients. This claim is consistent with all three views. Thus, we have reason to believe that, in an important range of cases, happiness tends to correlate with changes in well-being; in certain contexts, however, we have reason to believe that happiness correlates with levels of well-being.

In the third main section of the chapter, I show how this view can provide us with plausible interpretations of some of the main findings from the study of happiness. In particular, I will show that the Easterlin Paradox and the phenomenon of adaptation can be explained by the fact that happiness indicates changes in well-being. In addition, I will show that other important findings can be explained by the fact that, in certain contexts or conditions, happiness indicates either unexpected changes in well-being or certain aspects of one's level of well-being.

## 4.2 Three views of the function of affective states

In this section I will compare and contrast three philosophical theories of the function of affective states, namely the views outlined by Elijah Millgram (2000), Timothy Schroeder (2001; 2004; 2006) and Peter Railton (2008; MS). I will show that all three views are similar in that they maintain the function of our affective states entails that they tend to detect *changes* in our well-being, in contrast to our absolute level of well-being.

In discussing their respective theories, all three theorists use the terms “pleasure/displeasure”, “positive/negative affect”, “positive/negative emotion” and “happiness/unhappiness” interchangeably. For the purpose of evaluating the significance of their arguments for the claims of this chapter, the differences between these notions do not matter.<sup>71</sup> To keep things simple, I will discuss the claims of each view in terms of pleasure.

#### 4.2.1 Millgram’s View

Millgram argues in favour of the normative claim that pleasure should indicate changes in well-being. His target is the particular utilitarian view that we should maximise well-being and that well-being is constituted by pleasure. According to this version of utilitarianism, we should maximise pleasure. Millgram argues that this view is wrong because it consists in the wrong view of pleasure. According to Millgram, pleasure does not constitute well-being – the function of pleasure is to indicate changes in well-being.

Why does Millgram think that pleasure has the function of indicating changes in well-being? His argument is perhaps best illustrated by the following two examples that he provides:

“Getting admitted to (or rejected by) a prestigious college may make an enormous difference in how well the prospective student’s life is going, while making relatively little difference to how she feels: after a week or so of celebration or dejection, the applicant will be back on an even keel. The sense of accomplishment in bringing home the washing machine is gone by the next morning, although the contribution the appliance will make to one’s well-being is not. (You will very shortly cease to feel much of anything about the fact that you can now just drop your clothes into the machine, instead of spending your afternoons at the laundromat; but, feeling or no feeling, the washing machine continues to wash your clothes and to save you those afternoons.)” (Millgram, 2000: p.119)<sup>72</sup>

---

<sup>71</sup>This is not to say that there are no differences between pleasure, affect and emotion. In chapter one, I argued that affective states (a) represent features of our environment that are practically relevant, (b) represent environmental features as good or bad, and (c) may be either conscious or non-conscious states. In contrast to affective states, pleasures may exclude both components (a) and (c). Pleasures/displeasures may simply represent something as good/bad. In addition, pleasures/displeasures are always conscious states; they are what it feels like to represent something as good/bad.

<sup>72</sup> From now on, until stated otherwise, all references are from Millgram, 2000.

In these kinds of cases, Millgram claims that new circumstances should not make a difference to how we feel in the long-run. If someone cannot get used to their improved circumstances, Millgram claims, we will think that they are not getting something right. According to Millgram's view, experiencing pleasure is like receiving news: "When good news comes along, we become elated; happiness surges. But after a while, even though the news is as good as it ever was, pleasure fades." (p.121) News informs us of *changes* in our circumstances. Once these changes have been incorporated into our lives we should then be concerned with new changes, rather than dwelling on previous ones.

It is worth being clear on why Millgram thinks this is the case. One might think that it is counterintuitive not to be concerned with old changes. After all, as Millgram points out, our improved circumstances have not got worse – they are as good as they were when we first received news of them. So why not continue to take pleasure in old changes, as well as new ones? Millgram thinks that we should get used to our improved circumstances for the following reason. Pleasure, according to Millgram, is a mechanism for directing our priorities. We care about maximizing our well-being, and therefore care about improving our circumstances in some way. Thus, it makes sense to suppose that there's no point in dwelling on the good things that are already at hand. What we need to do is focus on changes in our circumstances and react to those changes in a way that will enable us to improve our well-being.<sup>73</sup>

On Millgram's view, then, pleasure should indicate *changes* in well-being. Like all indicators, pleasure can fail in various ways to indicate changes in well-being accurately. We may have false beliefs about our circumstances, for example (either thinking that our circumstances have changed when they in fact have not, or vice versa). Millgram also notes that we may put pleasure to use in fictional contexts, the same way that we put visual representations of our surroundings to a different use when going to the movies. Presumably, pleasure may also (correctly) indicate changes in our circumstances that do

---

<sup>73</sup> Although Millgram may implicitly assume that we care about maximizing our well-being, this assumption does not matter for organization of this chapter. Pleasure can still be viewed as a mechanism for directing our priorities, even if we do not care about maximizing our well-being. Millgram's view merely rests on the claim that we care about improving our circumstances in some way. Pleasure enables our lives to get better in some respect, through informing us of (and motivating us to react to) changes in our circumstances.



not relate to changes in our well-being, such as empathic pleasures, as well as (fictional and non-fictional) aesthetic pleasures. Although Millgram does not mention this last point, this possibility is consistent with his view.

In providing an account of why pleasure should indicate changes in well-being (rather than absolute levels of well-being) Millgram uses the analogy of hill-climbing. Hill-climbing is a technique used to climb the highest hill in a particular landscape in the event that you can only see the terrain around you (rather than the whole of the landscape). It works by identifying the highest point within a nearby radius, and then standing on that point. This process is repeated until there is no point within the small radius that is higher than the point you're standing on. Millgram claims that hill-climbing is efficient in that it doesn't require significant amounts of information that may be hard to come by, such as information about the overall topography of the landscape. For this reason, he claims that:

“Human beings are built to hill-climb [...] in order to hill-climb effectively, you don't need to know what your altitude is (even though you are hill-climbing in order to gain altitude). All you need to know is: which of a few adjacent points is higher than the others? If your strategy for improving your welfare is to hill-climb in the welfare landscape, you don't actually have to know how well you're doing; it's enough to know whether some change you make is a change for the better or for the worse.” (p.123)

Millgram claims that pleasure is an effective device for pursuing a hill-climbing strategy to improve our well-being. With limited information about the “welfare landscape,” one *should* pursue this strategy with such a device. In sum, Millgram argues that pleasure (as an indicator of changes in well-being) provides us with an effective way of promoting our well-being.<sup>74</sup>

In support of this view, Millgram provides us with several cases in which pleasure does in fact tend to indicate changes in well-being. He cites studies which show that subjects are not significantly happier or less happy a year after winning the lottery or becoming paraplegic respectively (Brickman et al. 1978). Such findings suggest that, although good fortune and disaster tend to

---

<sup>74</sup> As discussed in chapter three, we do not need to accept Millgram's claims that pleasure *should* have this function in order to accept the fact that pleasure has this function in a historical sense. That is, we can claim that pleasure has been set up (either through evolution or learning) to provide us with an effective way of promoting our well-being, even if pleasure should be used in a different way (e.g. to respond to moral concerns).

make us feel pleasure or displeasure, these feelings wear off once we become accustomed to our new situation: once you get used to your new circumstances, they will “not make that great a difference to how you feel” (Millgram, 2000: p.117).

Of course, there are exceptions to this tendency. Millgram admits that sometimes people get depressed and stay that way; that severe trials may permanently affect one’s ability to be happy, even if becoming paraplegic is not enough. On Millgram’s view, however, these are cases in which pleasure fails to perform its function. Depressed individuals, for instance, are dysfunctional, according to Millgram, in that they are unable to detect new changes in their well-being and thereby effectively promote their well-being.

In addition, it is worth noting that there are cases in which pleasure succeeds in performing its function, but that doing so may be an ineffective strategy for promoting well-being. Pleasure may be an effective tool in many circumstances, but it is not invariably the right tool, and doesn’t invariably produce the best outcome (i.e., improve our well-being). Millgram again uses the hill-climbing metaphor to highlight two ways in which this may be the case, which I will briefly consider in turn.

First, the effectiveness of hill-climbing relies on going from point A to point B being of a similar difficulty to going from point B back to point A. That is, if we discover that the highest point we have reached is not in fact the highest point that we can reach, we may need to retrace our steps before being climbing to the highest possible point. In certain contexts, this may be considerably more difficult or even impossible. This analogy can be clearly seen with changes in well-being. For example, although a degree in medicine may be an improvement in one’s well-being, it may not be as good as a degree in law in the long-run (i.e. if one ends up deciding to be a lawyer); unfortunately, however, once one has completed a medicine degree, one may no longer have the time or money to do a law degree.

Second, if all we do is hill-climb then we can get trapped at the top of a relatively low hill. This may be the case even if there are higher hills (just out of view) to climb elsewhere. With regards to well-being, we may reach a dead-end of potential further improvements, and not be able to see any alternative options available to us. For example, we may consistently try and improve our well-being by making improvements to the business that we started up. We may try and improve our marketing strategy, customer relations or overall efficiency. Only being able to focus on making these kinds

of improvements may leave us blind to the fact that starting up a very different kind of business would make us much better off. Simply put, if all we do is focus on changes in well-being, we may fail to see whether the improvements we have made have made us significantly better off overall.

Thus, hill-climbing can be ineffective both because we may be unable to retrace our steps and because we may be unable to see better improvements in our well-being. Although indicating changes in well-being may often be an effective way of promoting our well-being, this may not always be the case. The above two points will be of interest later on, when considering the contexts in which the function of affective states does not entail that they detect changes in well-being.

In sum, Millgram claims that pleasure has the function of indicating changes in well-being. We should feel good when our lives improve. Conversely, we should feel bad when our lives get worse. Indeed, we often do. Pleasure has the job of indicating changes in well-being because responding to such changes tends to be an effective strategy for the promotion of well-being.

#### **4.2.2 Schroeder's View**

Schroeder defends a representational view of pleasure. He maintains that pleasure (displeasure) represents net changes in desire satisfaction (frustration). For the purpose of evaluating the significance of his arguments for the claims of this chapter, the differences between desire-satisfaction and well-being do not matter.<sup>75</sup> Thus, in discussing Schroeder's view, I will use the terms "desire-satisfaction" and "well-being" interchangeably.

After arguing that pleasure/displeasure are representational states, Schroeder claims that pleasure represents net gains in desire satisfaction and displeasure represents net losses: "By and large, getting what we want pleases us, and being pleased is a sign that things are going our way. Similarly, having our desires frustrated is unpleasant, and being displeased is a sign that things are going against our wishes." (Schroeder, 2000: p.513)

---

<sup>75</sup> Again, the notions used by Schroeder reflect the aims of his argument. Schroeder focuses on the notion of desire-satisfaction, rather than the on notion of well-being, primarily because he does not want to commit to any particular theory of value (i.e., whether a desire-satisfaction theory of well-being is true, whether reasons internalism is true, and so on). Thus, rather than claiming that pleasure represents net changes in well-being, Schroeder maintains a value-neutral view.

Schroeder takes net desire satisfaction or frustration to be a function of the satisfaction or frustration of individual desires, with stronger desires being weighted more heavily than weaker desires in the global evaluation. Thus, net desire satisfaction may be positive or negative to a greater or lesser degree.

Schroeder's view differs from and extends Millgram's view in two respects. Firstly, Schroeder introduces the notion of *intensity* of pleasure, whereby the greater net change in desire satisfaction, the more intense the resulting pleasure.<sup>76</sup> With regards to well-being, Schroeder's view would be that the greater the net change in well-being, the more intense the resulting pleasure.<sup>77</sup>

Secondly, Schroeder claims that pleasures track *net* changes in desire-satisfaction. Thus, insofar as pleasure accurately represents the net change in desire-satisfaction, one cannot experience simultaneous pleasures and displeasures.<sup>78</sup> If, for example, our desire for chocolate is frustrated at the same time as our desire for television is satisfied, we will experience either pleasure or displeasure depending on which of the two desires are stronger. If our desire for television is stronger than our desire for chocolate, we will experience pleasure. This is because we have witnessed a net increase in desire-satisfaction. If both our desires for chocolate and television are satisfied, we will experience a still greater amount of pleasure, which indicates a greater net increase in desire-satisfaction.<sup>79</sup>

---

<sup>76</sup> The work of Davis (1981) and Reizenzein (2009) support this view. Both theorists claim that the intensity of pleasure is determined by the sum of (a) the strength of the desire that has been satisfied and (b) the degree of belief that the desire has been satisfied.

<sup>77</sup> It is important to note that the intensity of pleasure is distinct from the level of arousal. Being pleased sometimes involves being stimulated, excited, aroused: one talks about joy or ecstasy; other times, being pleased involves being contented, relaxed, put at ease: one talks about bliss or satiation. Schroeder correctly maintains that we can distinguish between the high/low arousal component of affective states and the pleasurable component of affective states. One can, for instance, compare a moment of joy and a moment of bliss and say which was more pleasant i.e., which pleasure was more intense.

<sup>78</sup> This does not rule out the *possibility* of experiencing simultaneous pleasures and displeasures. Insofar as pleasure *inaccurately* represents the net change in desire-satisfaction, experiences of simultaneous pleasures are possible.

<sup>79</sup> Again, it is useful to distinguish the intensity of pleasure and the level of arousal, here. According to Schroeder, when desire-satisfactions are combined in this way (e.g., the net satisfaction constituted by the satisfaction of our desire for chocolate and television)

Like all representations, pleasure can fail to represent net changes in desire-satisfaction in various ways. Schroeder outlines three of these ways. Firstly, such representations may be, to a certain extent, cognitively impenetrable. The result is that perceptual representations of net changes in desire-satisfaction can override intellectual representations. Schroeder gives the example of a child having an injection, who may feel less pain as a result of looking away from the needle, despite believing (on an intellectual level) that he is being injected by it (Schroeder, 2000: p.517). Secondly, pleasure may fail to pick up on net changes in desire-satisfaction when there are extraneous influences on affect, such as in cases of depression. Lastly, both pleasure and displeasure can increase. In such cases, pleasure says that our desires are, on balance, being satisfied; displeasure says that they are, on balance, more frustrated. Schroeder likens such experiences to cognitive illusions, whereby we have similar experiences of mutually exclusive phenomena.

One last point is worth mentioning concerning a further respect in which Schroeder's view differs from/extends Millgram's view. In later work, Schroeder (2004; 2007) added the important role that expectations play in the elicitation of pleasure. Consider the following example provided by Schroeder:

"If Theresa is accustomed to sleeping on a nice double bed, being forced to sleep on a plastic-coated single bed (say, in a student dorm) will be unpleasant; but if she lives in the dorm, she will become accustomed to the less-nice bed, and stop being caused displeasure by it. Similarly, if she is sure that Montreal will win the Stanley Cup, she will be more displeased by a loss than if that was what she was expecting all along, all else being equal." (Schroeder, 2007: 260)

The role that expectations play in the things that reliably cause pleasure makes sense if pleasure represents net changes in desire-satisfaction. This is because net changes in desire-satisfaction must be net changes relative to *something*. Now, it may be that pleasure represents change relative to total levels of desire-satisfaction, or absolute levels of well-being (as in Millgram's view). However, Schroeder claims that this is less obvious in the case of pleasure. He suggests that pleasures in this sense are more like our sense of

---

there is only one resultant pleasure, which represents the net desire-satisfaction. We can, however, distinguish between the desire-satisfactions on the basis of their sources (i.e., eating chocolate and watching television) and the attendant non-pleasure associated feelings (i.e., arousal vs. non-arousal) that come with each source.

warmth and cold, which represents changes from a particular baseline rather than absolute levels of heat.

Schroeder modified his view so that pleasure represents only *unexpected* net increases in desire-satisfaction. He further stresses that this expectation-related baseline is also, to a certain extent, cognitively impenetrable. That is, it is our gut-level expectations that play the important role. This explains statements such as, “Of course I knew she would be late, but somehow it still surprised me”, or cases whereby one “can’t quite believe” they got the great new job, romantic partner, or piece of financial luck they did (Schroeder, 2007). In these cases, gut-level and intellectual-level expectations come apart, and our gut-level expectations tend to determine the pleasure or displeasure that we feel.

The work of emotion theorist Nico Frijda (1988) supports the importance of gut-level expectations in eliciting pleasure. He notes that, generally speaking, the frame of reference that determines what counts as an emotional event consists of that which is deemed possible: “Those who grieve and mourn have not really taken their leave from the departed person; they still expect him or her at the other end of their arms, bed, or table. Those who feel that they should be able to cope suffer when they cannot cope.” (Frijda, 1988: p.353) The role of expectations also explains how small improvements in well-being can sometimes cause (seemingly disproportionate) intense pleasures, such as the large amount of joy experienced from receiving a gift from a stranger. Such changes are small, but unexpected, and thereby consist in a greater unexpected net change in well-being.

In sum, Schroeder claims that pleasure (affect) represents unexpected net changes in desire-satisfaction (well-being). We tend to feel good as a result of our lives improving in unexpected ways. Conversely, we tend to feel bad due to our lives getting unexpectedly worse.

#### **4.2.3 Railton’s View**

Like Millgram, Railton cites “surprising” findings from the study of happiness that are in need of some kind of explanation. As mentioned in chapter one, this field of research is often referred to as the study of subjective well-being (SWB). SWB researchers have found that very few life conditions make a difference to our long-run SWB. SWB in developed countries has not increased over the past 50 years, despite large increases in the material standard of living. Many chronic illnesses and serious disabilities do not result in long-term declines in SWB. The old tend to have higher levels of SWB than

the middle-aged. Climate, religion, family, friends, social status, etc. tend not to matter for SWB. What does matter is personality, intimate relationships, self-determined activities, keeping physically and mentally active, employment and freedom (Railton, 2008; MS).

These kinds of findings are often explained by the “set point” theory of SWB: despite changes in life conditions, we slide back to our SWB set point. Railton sums up this trend as follows:

“What we desire promises a gain in well-being. We like this idea, and so are motivated to attempt to satisfy the desire (or remove frustrations to it). As the desire is satisfied, we do experience a “shot” of well-being. When habituation causes this well-being to disappear like water in sand, we do not blame satisfying the desire (that felt good), but the fact that we are not now satisfying the desire.” (Railton, MS: p.14)<sup>80</sup>

In offering an explanation for these phenomena, Railton claims that it is only because SWB has a set point structure that it can play the role it should in our psychic economy. In this sense, his view is much like Millgram’s view. In contrast to Millgram’s hill-climbing analogy, Railton illustrates the role of the affective system using the analogy of an autopilot in an aircraft. An autopilot works by recording an “error signal” whenever the aircraft drifts very slightly off course (to one side or another). In response to an error signal, the autopilot turns on or off the appropriate motors (that move the aircraft from one side or another) to remain on course. Once the ship is back on course, the “error signal” ceases to be recorded, and the motors shut down. Railton notes that this system contains the basic elements of regulation:

“A set-point value, a monitor providing positive or negative feedback concerning the state of the system relative to the set-point, an automatic connection between this information and corresponding control of the system state, and continued monitoring to determine whether this control operation is succeeding in restoring the set-point value. Our machines and our bodies are full of such homeostatic regulators.” (p.14)

Railton goes on to claim that our affective system is a more intelligent regulator than this. In a further analogy, he likens our affective system to a delta-meter in a racing sailboat. Delta-meters respond to changes in speed. When the boat is travelling at a constant speed, *fast or slow*, the meter reads zero. But if the boat begins to speed up, the meter gives a positive reading

---

<sup>80</sup> From now on, until stated otherwise, all references are to Railton, MS.

(the greater the increase in speed, the greater the positive reading). And if the boat slows down, the meter gives a negative reading (the greater the decrease in speed, the greater the negative reading).

With regards to well-being, the *position* of the racing sailboat is analogous to a person's absolute *level* of well-being. The *speed* (or velocity) of the boat is analogous to a person's *change* in well-being. And the change in speed (or acceleration/deceleration) of the boat is analogous to a person's rate of change in well-being. Railton uses the analogy of the delta-meter to show that the function of pleasure does not merely entail that it detects changes in well-being. It may also detect a person's rate of change in well-being.

Railton considers this kind of regulation to be more intelligent because it is not wed to any set point. The zero point of the delta meter does not correspond to a certain velocity; rather, it corresponds to the *change in speed* (i.e. acceleration/deceleration) – the first derivative of velocity. Absolute speed loses its significance, so long as the sail trimmer is doing what they can to secure every possible gain in speed at that moment, and avoids every possible way of slowing the boat down. In this way, our affective system “is not keyed to our absolute rate of movement through life (income, accomplishment, etc.), but to whether we are making more or less progress in attaining the goals or desires of the moment.” (p.15)

It is worth being clear on how Railton's position differs from Millgram's view in this respect. The key difference between the two views concerns the respective *baseline* from which pleasure is caused by prudential features of our environment. For a regulatory system to be sensitive to change, it must return to a certain baseline. Consider, for example, winning the lottery and feeling good as a result. Through feeling good, we detect the fact that winning the lottery has increased our level of well-being. However, if winning the lottery continued to make us feel good, we would fail to recognise further changes in our well-being. In order to detect such changes, the pleasure caused by winning the lottery must, at some point, dissipate (ideally, when it ceases to be useful for informing and guiding our action). Our level of pleasure must, at some point, return to a particular baseline.

Now, according to Millgram, this baseline is our absolute *level* of well-being. For Railton (as well as Schroeder), this baseline is our *expected changes* in well-being. According to both Railton and Schroeder, our feelings are determined by our anticipations concerning changes in our well-being. If we anticipate that our level of well-being will stay the same, and our level of well-



being increases (decreases), we will experience pleasure (displeasure). Similarly, if we anticipate that our level of well-being will increase, and our level of well-being increases more (less) than we anticipated, we will feel good (bad). In this way, pleasure only detects the most relevant features of our environment for action. Through anticipating the state of the world, and then matching those anticipations with the actual state of the world, we only need to pay attention to genuinely new information. Expected information is redundant. Thus, according to Railton (and Schroeder), the function of pleasure entails that it detects unexpected prudential features of our environment.

Railton's view explains two "surprising" findings regarding the causes of pleasure. Firstly, although many kinds of circumstances have been shown to make no lasting difference to people's level of happiness (e.g. increases in income, status, educational attainment, etc.) certain kinds of events and conditions do tend to have a lasting impact. For example, experiences such as going on vacation or having a meal at a restaurant tend to make people happier than material purchases such as clothes or computer equipment (Van Boven & Gilovich 2003). As another example, people's level of happiness adapts more quickly to news that they definitely have a serious illness than to news that they might have a serious illness (Frederick & Loewenstein 1999). What these kinds of events and conditions have in common is that they tend to be unexpected. Insofar as going on vacation or having a meal at a restaurant is variable, it is hard to predict the impact on such events on our well-being (Wilson & Gilbert 2008). And insofar as potentially having a serious illness is uncertain, the impact of such a condition is also unpredictable. Railton's view emphasises the fact that unexpected circumstances tend to cause us pleasure/displeasure. It explains why certain kinds of events and conditions tend to have more of a lasting impact on our happiness than others – the impact that such circumstances have on our well-being tends to be unexpected.

Secondly, Railton's view explains why social references have such a pervasive impact on happiness (Frank, 2001; Layard, 2005). For example, unemployment tends to have a lasting (negative) impact on happiness in a healthy economic climate, but not so much of an impact within an economic depression (Graham, 2009). In other words, we feel bad when we are the only one who is doing badly, but we feel okay if we are doing just as badly as everyone else. Again, this phenomenon can be explained with reference to our expectations. According to Railton's view, the function of pleasure entails

that it detects how well we are doing relative to how well we expect to be doing. People around us provide us with useful information about what is possible in our situation. If everyone around us is doing well and we are doing badly, we can expect to be doing better. Thus, our actual circumstances will be worse than our anticipated circumstances, thereby making us feel bad. In contrast, if everyone around us is doing just as badly as we are, this suggests that we are doing everything we can in virtue of the options we face, the resources that we are utilising, and so on. Our actual circumstances will be no worse than our anticipated circumstances, and therefore we will cease to feel bad about them. This is consistent with other findings in social psychology, which suggest that we use the information provided by the actions and circumstances of those around us to form predictions about how we should act.<sup>81</sup>

Railton reviews an array of neurological evidence in showing that our affective baseline is our expected changes in well-being. For instance, he reviews the influential work of Schultz et al. (1997), which suggests that we experience pleasure in accordance with the baseline of our expected changes in our circumstances. An alternative way of viewing this point is that pleasure is reliably caused by changes (“prediction errors”) in our expected changes in our circumstances.

What matters, according to this view, is not necessarily whether our well-being improves, but whether our well-being is improving to a greater or lesser degree than expected. Can we improve our well-being to an even greater extent?<sup>82</sup> Or are we trying to improve our well-being too much, perhaps in an unfeasible manner? We are effectively learning at each moment how much improvement to seek, adjusting this to constantly changing circumstances. Whenever we increase our rate of improvement we are rewarded by an experience of positive affect; whenever we decrease our rate of improvement, we are punished by an experience of negative affect. This learning system is continuously monitoring our rate of progress, and uses this information to attune and re-attune our responses to the world.

---

<sup>81</sup> For a comprehensive review of findings from social psychology concerning the impact of others on people’s behaviour see Doris, 2002.

<sup>82</sup> Note that we may be able to improve our well-being to an even greater extent either by continuing our level of progress in the particular task at hand or through focusing attention and devoting resources to alternative areas of our lives. Carver (2003) argues that pleasure has the function of performing this latter role.

Lastly, it is worth noting how Railton's delta-meter metaphor is also instructive of the ways in which using such an indicator may *not* be an effective way of promoting well-being in the long-run. In the main, even if our delta-meter regularly produces positive readings, it still does not provide us with information about our absolute distance travelled. Suppose, for example, that we have been sailing in strong winds or against the tide. We may regularly find new ways of gaining headway, and thereby increasing our rate of progress. Yet, our ultimate progress may be minimal. Without looking over to the horizon, we will not be able to judge whether other boats have made much more progress (in different winds or tides). We can clearly see this analogy with regards to well-being. Returning to the example used in discussing Millgram's view, we may spend a significant amount of trying to improve our business, regularly increasing our rate of improvement in well-being. Yet, at some point, we can come to realize that setting up a different kind of business would have resulted in a greater change in our well-being overall. This kind of information cannot be provided by an affective system that solely indicates either our change in well-being or rate of change in well-being.

In sum, Railton claims that the function of pleasure entails that it detects unexpected changes in our well-being. According to this view, we tend to feel good when our lives improve to a greater extent than we anticipated. Conversely, we tend to feel bad when our lives improve to a lesser extent than expected.

### **4.3 The Refined Indicator View**

In the previous section I reviewed three accounts of the function of affective states. In the second section of the chapter, I will show that, from these three views, we can establish two main claims. First, in an important range of cases, the function of affective states entails that they detect (unexpected) changes in well-being ingredients. This claim is endorsed by both Schroeder's and Railton's views and is consistent with Millgram's view. Second, in certain contexts, the function of affective states entails that they detect levels of certain well-being ingredients. This claim is consistent with all three views. Thus, in an important range of cases, we have reason to believe that happiness tends to correlate with changes in well-being; in certain contexts, however, we have reason to believe that happiness correlates with certain aspects of one's level of well-being.

### 4.3.1 Affective states and (unexpected) changes in well-being

The first claim that we can draw from the views above is that, in an important range of cases, the function of affective states entails that they detect (unexpected) changes in well-being ingredients. All three accounts maintain that affective states should be reliably caused by changes in well-being ingredients relative to a *baseline*. For Millgram, this baseline is our absolute level of well-being ingredients; for Schroeder and Railton, this baseline is our expected changes in well-being ingredients. I believe that the latter views are right in thinking that, in an important range of cases, affective states should be reliably caused by changes in well-being ingredients relative to the baseline of our expected changes in well-being ingredients.

Consider, for example, observing the value of your stocks and shares decline as a result of a poor marketing policy. Each day your stocks decline, and thereby decrease your level of well-being (with regards to certain well-being ingredients, such as income, reputation, achievement, etc.). After a few days, however, the stocks start to decline at a slower rate; they are still declining, but not as quickly as they were before. This change may result in you feeling happy – you may feel good that things are not as bad as they were before (“Perhaps things are looking up after all?”) even though your level of well-being is still decreasing. Yet, this trend should not be detected by pleasure, according to Millgram’s view. According to Millgram, happiness should be reliably caused by changes in well-being relative to your absolute level of well-being. Thus, observing the value of your stocks and shares decline at a slower rate should still result in you feeling unhappy – you should feel bad, according to Millgram, because your level of well-being is decreasing. On the one hand, this is correct – it is useful to know that your stocks are declining. On the other hand, however, merely detecting whether your level of well-being is decreasing or increasing misses crucial information about your situation. It may be that your stocks have declined at a slower rate because of the changes you made to your marketing policy. Indeed, if you continued to make similar changes, your stocks may cease to decline altogether, and may even start to increase. It is in this way that detecting unexpected changes in well-being provides you with practically relevant information beyond your overall change in well-being. The fact that your stocks and shares are declining at a slower rate suggests that you are on the right track; that you should keep on doing whatever it is that you’re doing. Pleasure should be reliably caused by this kind of information insofar as it should inform and guide action.

As another example, consider Charles Carver's (2003) view of the function of pleasure. Carver argues that pleasure indicates increases in our rate of increases in well-being. According to Carver, pleasure indicates not merely that our well-being is increasing; rather, pleasure indicates that our rate of increases in well-being is increasing. We feel good, according to Carver, when we are improving at an even greater rate than we were previously. Carver argues that the function of detecting such changes is to signal that we are improving *more than we need to*. Such improvements are good for us, but we could be devoting our attention and resources to improving other aspects of our well-being, namely the aspects of our well-being that are more in need of improvement. As a result of indicating increases in our rate of increases in well-being, pleasure can provide us with this kind of information. Again, however, according to Millgram's view, pleasure does not have this function. According to Millgram, pleasure should merely be caused by changes in well-being ingredients. This incorrectly rules out contexts in which we require information about whether certain aspects of our well-being ingredients are improving at a greater rate than they need to.

Schroeder and Railton's views capture cases in which happiness should detect rates of change in well-being ingredients in the following way. When we come to expect constant improvements in well-being ingredients, Schroeder and Railton both maintain that affective states should detect unexpected changes in well-being ingredients.

Schroeder and Railton's views also adequately account for cases in which strong affective states should be indicated by relatively small changes in well-being ingredients (and vice versa). Consider, again, the case of experiencing strong feelings of joy as a result of receiving a gift from a stranger. Receiving this gift is not a great change in well-being, yet it is indicated by a strong affective state. According to Schroeder's and Railton's views, affect strength should not indicate the degree of changes in well-being in themselves. Rather, affect strength should indicate the degree of changes in well-being relative to our expectations. Thus, a relatively small change in well-being ingredients may cause a strong affective state insofar as the change was largely unexpected. Conversely, a relatively large change in well-being ingredients may cause a weak affective state insofar as the change was largely expected.

In sum, both Schroeder and Railton's views discussed above endorse the claim that, in some cases, the function of affective states entails that they

detect unexpected changes in well-being ingredients. In such cases, expected changes do not provide us with practically relevant news about our circumstances. In contrast, unexpected changes do – they tell us that the world was not as we had thought it was. This logic is consistent with Millgram's view. Millgram argues that the function of affective states is to indicate changes in well-being because it is these changes that provide us with practically relevant news about how well we are doing.

Indeed, I think that Millgram's view is insightful in this respect. I suggest that changes in well-being ingredients tend to be, to some extent, unexpected. Thus, Millgram's view (that the function of affective states entails that they detect changes in well-being ingredients) may tend to be similar in practice to both Schroeder and Railton's view (that the function of affective states entails that they detect unexpected changes in well-being ingredients).

The reason for believing this claim is simply that most changes in well-being ingredients tend to be, to a certain extent, unpredictable. Consider, for example, getting into a warm bath, and finding it pleasurable. You know the bath will be warm because you tested the temperature before getting in. You also know what a warm bath feels like from previously having a warm bath. Yet, despite knowing how the warm bath would feel, you still felt pleasure from getting into it. This is because, to a certain extent, the experience was unpredictable. We may know that getting into a warm bath feels relaxing in various ways, but we do not know the *exact* way in which the warm bath will relax us – the particular groups of muscles that will be soothed by the warm water, the exact change in external body temperature, and so on. The exact state of affairs involved in getting into the warm bath is unexpected; thus, such states of affairs tend to result in us feeling pleasure.

This phenomenon is perhaps best illustrated by the practice of tickling. An influential model of agency (called the Forward Model: see Prinz 2007b) suggests that we distinguish between active and passive movements by the presence or absence of anticipatory images that match the experiences that we have when our bodies move. For instance, if I am about to lift my arm in a specific direction, I form an anticipatory image of what that movement would feel like. If the actual feeling of lifting my arm in that direction matches my anticipatory image of that movement, the action will feel active. Conversely, if the actual feelings of that movement do not match my anticipatory image, the action will feel passive. Now, the practice of tickling shows just how exact one's anticipatory image needs to be in order to match one's actual feelings.

Notoriously, you can't tickle yourself. This is explained by the fact that the tickle response requires unexpected movements. Blakemore et al. (1998) developed a robotic hand that a subject can use to tickle herself. The subject controls the hand through his or her own finger movements. If the hand moves in sync with the subject's fingers, tickling does not occur – the movements of the hand were entirely expected. However, if the hand moves with a slight delay, tickling does occur – even though the subjects knew they were about to be tickled, and the movements of the hand were expected, the exact time at which the tickling occurs was unexpected. This level of exactness in our expectations suggests that pleasure can still result from relatively expected states of affairs insofar as those states of affairs are not expected exactly.

I think we can reasonably assume that changes in well-being tend to be partly unexpected in this respect. We expect our food to taste a certain way, but we cannot exactly predict how it will taste. Likewise, we can predict how a certain interaction will go, but we are unable to do so entirely. Our expectations of changes in well-being ingredients may have a considerable effect on the intensity of the affective states reliably caused by such changes. But, it is unlikely that they will tend to null our affective responses altogether. It is for this reason that we can assume that changes in well-being ingredients will tend to be unexpected changes, and thereby reliably cause affective states. The upshot is that we have reason to believe that happiness tends to correlate with changes in well-being.<sup>83</sup>

#### **4.3.2 Affective states and certain aspects of one's level of well-being**

In this section, I will argue in favour of a second kind of claim, in addition to the claims argued for in the previous two sections. It is the claim that, in certain contexts, the function of affective states entails that they detect one's level of certain well-being ingredients. Thus, in such contexts, we have reason to believe that happiness tends to correlate with certain aspects of one's level of well-being.

---

<sup>83</sup> The fact that our expectations have a considerable effect on the intensity of the affective states may make a significant difference to the kinds of circumstances that affect our overall levels of happiness in the long-run. I will deal with this issue in more detail in the final main section of this chapter, in considering findings from the study of happiness. I will argue that expectations shaped by social references play a large role in determining the kinds of external conditions that impact our level of happiness over time.

This claim relates to the contexts discussed above (in discussing both Millgram's and Railton's views) in which the sole focus on changes in well-being ingredients can fail to be an effective strategy for promoting well-being. Recall Railton's delta-meter analogy, and how, in certain contexts, a delta-meter fails to provide you with practically relevant information. When sailing in strong winds or against the tide, you may regularly find new ways of gaining headway, and thereby increasing your rate of progress, but make little progress overall. With regards to well-being, we may spend a significant amount of time and effort making progress in a particular endeavour, only to realize that we would have been better off putting our resources into a different project altogether.

When we take a more "reflective perspective" (Tiberius 2005b) the practically relevant features of our environment will not tend to be (unexpected) changes in well-being ingredients. Rather, we need to take into account our level of certain well-being ingredients – how well we are doing in a more global respect.<sup>84</sup> We need to reflect on how much progress we are making, where it is that we are going, and so on. A more reflective perspective sometimes requires that our levels of certain well-being ingredients reliably cause our affective states. As Valerie Tiberius puts it: "Acknowledging the beauty of our everyday surroundings and appreciating the wonders of nature are good things to do and sometimes an invaluable tonic for the materialistic or accomplishment-focused perspectives that can easily absorb us." (Tiberius, 2005b: p.165) Such reflection (which includes a range of cognitive capacities, including our affective dispositions) can enable us to reprioritize certain well-being ingredients or enable us to more effectively pursue our current priorities. Of course, for the purposes of action, such moments of reflection should ultimately be short-lived. As Tiberius goes on to note: "It is good to stop and smell the roses, but not so good to smell them all the time." (Tiberius, 2005b: p.165) For this reason, our absolute levels of certain well-being ingredients will not tend to be a key determinant of happiness.

The function of affective states, in certain contexts, entails that they detect our level of certain well-being ingredients. This fact can account for some kinds of cases that are not as plausibly captured by Millgram's, Schroeder's or Railton's views. In discussing Schroeder's account, above, I briefly discussed

---

<sup>84</sup> In chapter five, I will consider the difference between local and global aspects of well-being in more detail.



an example provided by Frijda (1988) concerning the suffering of those who grieve. Frijda agrees with Schroeder in claiming that the frame of reference that determines whether an event causes an affective state in us consists in our expectations. Thus, Frijda writes: "Those who grieve and mourn have not really taken their leave from the departed person; they still *expect* him or her at the other end of their arms, bed, or table. Those who feel that they should be able to cope suffer when they cannot cope." (Frijda, 1988: p.353; emphasis added) Frijda claims that those who grieve suffer because they (partially) expect the departed person to be present. I think that this explanation is only partly correct, however. Frijda is right to claim that there is some kind of baseline, or relevant comparison, in play. But I think he is wrong in claiming that this frame of reference consists in our expectations. It seems more plausible to claim grief is determined by the fact that the differences between when the person was alive and now when they're dead are practically relevant. That is, it is the loss itself that is salient, not the unexpected nature of the loss. This seems more plausible than imputing an irrational belief that they haven't departed, such that the grieving person is constantly surprised that they aren't there.

Moreover, there are some cases that cannot be adequately accounted for by (unexpected) changes in well-being ingredients. Consider, for example, acts of gratitude that often cause pleasure in those performing such acts. The practice of gratitude is a large part of almost all major religions and is now a major focus of positive psychologists (Emmons 2007). Positive psychologists advise subjects to consider three things that they are grateful for each day. This simple practice can make a significant difference to subjects' level of happiness (Seligman 2011). Yet, often the things one can be thankful for are neither things that have changed in any way nor are unexpected. Indeed, one may generally be grateful for familiar features of one's life, such as one's health, family, loving relationships, food to eat, house to live in, and so on. Again, it seems that a more plausible explanation of these kinds of cases is that such familiar features of one's life become practically relevant as a result of deliberately focusing one's attention on them.

As one last example, the fact that, in certain contexts, the function of affective states entails that they detect our level of certain well-being ingredients can explain the influence that our level of energy has on our happiness (Gailliot 2012). Simply put, when we have high levels of energy, we tend to feel good;

when we have low levels, we tend to feel bad.<sup>85</sup> I think the most plausible explanation of this phenomenon is that our level of energy dictates the extent to which we should focus on our level of well-being. When we have low levels of energy, we tend to focus on our lack of certain well-being ingredients, for instance. This makes sense insofar as we should focus on maintaining our current level of well-being when we do not have enough energy to increase it.

### 4.3.3 Putting it all together

In the past two sections, I have argued in favour of two main claims regarding the function of affective states. First, in a large range of cases, the function of affective states entails that they detect (unexpected) changes in well-being ingredients. Second, in certain contexts, the function of affective states entails that they detect levels of certain well-being ingredients. Thus, we have reason to believe the following: in an important range of cases, happiness tends to correlate with changes in well-being; however, in certain contexts, happiness correlates with certain aspects of one's level of well-being. In short, when we have reason to take happiness as an indicator of well-being, it is appropriate to take happiness as a defeasible indicator of (unexpected) changes in well-being.

I do not think we can justify the stronger claim, namely that, when happiness correlates with well-being, it *generally* correlates with (unexpected) changes in well-being. I do not think we can justify this claim for the same reasons I argued in chapter three that we cannot justify the claim that happiness generally correlates with well-being. In certain contexts, the function of happiness does not entail that it detects (unexpected) changes in well-being. These contexts may add up, to the extent that happiness does not generally correlate with (unexpected) changes in well-being. A theory of the correlates of happiness, which is based on the function of happiness, should be open to this empirical possibility.

I do think, however, that we can investigate and identify the kinds of contexts in which happiness is likely to correlate with changes in well-being in general or, in particular, unexpected changes in well-being or certain aspects of one's

---

<sup>85</sup> In referring to our level of energy, I do not mean our momentary level of arousal, which is taken to be a component of affective states. Rather, in referring to our level of energy, I mean our current level of available energy, such as our blood glucose level. We can have low energy in this sense, but still have high arousal states (e.g. we can be in need of sugar, yet still run away from a danger). Low energy in this sense is strongly correlated with low mood (Gailliot, 2012).

level of well-being. For instance, some studies have manipulated the practical relevance of different features of the same situation, in order to cause different kinds of affective states.<sup>86</sup> For instance, Hsee et al. (1991) show that if an outcome is framed in such a way that its relation to our *level* of well-being is most relevant for action, this will result in a different kind of affect than if the outcome is framed such that its relation to *rate of change* in well-being is most relevant for action. Hsee and his colleagues suggest that affect simultaneously indicates circumstances related to levels of well-being as well as the first and second derivative of levels of well-being.<sup>87</sup> The relative weights of each of these three dimensions of well-being is determined by the way in which the outcome is framed (i.e., which dimension is most relevant for action for the subject at any given time).

The upshot is that, when presented with information about happiness, we can treat it as potentially providing us with three different kinds of information about well-being. Firstly, it may provide us with information about changes in well-being. Secondly, it may provide us with information about unexpected changes in well-being. Thirdly, it may provide us with information about certain aspects of one's level of well-being. By thinking about the contexts in which happiness is caused, we can identify which of these three dimensions of well-being happiness provides us with information about. In the next section, I will interpret some of the key findings from the study of happiness using this refined version of the Indicator View.

## 4.4. Using the Indicator View to Interpret Empirical Findings

### 4.4.1 The Easterlin Paradox and the Phenomenon of Adaptation

Perhaps the most important finding from the study of subjective well-being<sup>88</sup> is the finding that significant increases in income do not tend to result in lasting

---

<sup>86</sup> Indeed, in a footnote, Millgram (2000) mentions these experimental findings, suggesting that his account, if fully developed, would be consistent with the view presented here.

<sup>87</sup> In addition, Hsee et al. (1994) provide evidence that affect co-varies with the *third* derivative of well-being (i.e., changes in rates of change in well-being) though admits that such patterns may be difficult to detect in our practical lives. It seems very unlikely that any higher derivatives of well-being (i.e. fourth or fifth derivatives) can be practically relevant.

<sup>88</sup> In this section, I will talk of “subjective well-being findings” in conjunction with “happiness findings” as the two terms are often used interchangeably in the empirical literature.

significant increases in subjective well-being (on a population level) – the so-called “Easterlin Paradox”. Interestingly, recent studies have shown that increases in income do tend to result in lasting increases in life satisfaction, but fail to impact people’s happiness (Kahneman & Deaton, 2010). I will discuss the difference between happiness and life satisfaction in detail in chapter five. It is also worth noting that the debate over the validity of the Easterlin Paradox findings rages on (Easterlin, 2010). I will not enter into this debate. For our purposes, I will assume that the Easterlin Paradox findings are valid and show how the Indicator View can plausibly take such findings into account.

The Easterlin Paradox seemingly forces us to answer the following question: If income has little impact on happiness in the long-term, is it the case that happiness fails to indicate the impact that income has on well-being, or is it that income has little impact on well-being? Different theorists answer this question differently, often bringing further evidence to bear on the question. Subjective well-being researchers tend to claim that the Easterlin Paradox shows that income has little impact on well-being. For instance, Schwartz (2005) argues that the benefits of income are counter-balanced by the disbenefits of increased choice – too many choices can make us less capable of making good decisions and being satisfied with the decisions that we make. Offer (2007) argues that the benefits of income are outweighed by the self-control problems that affluence creates – problems such as obesity and a lack of commitment to long-term relationships. Both of these theories are plausible, as are many others. Increases in income tend to come with increased mobility and the loss of social capital (Putnam 1995; Lane 2000), more easily accessible status goods (in contrast to relational goods) (Becchetti et al. 2011; Layard, 2005), increases in materialism and advertising (Kasser 2002; Lewis 2013), income inequality (Wilkinson et al. 2009; Marmot 2004), working hours (Coote & Franklin 2010; Frank 2012), urbanisation and environmental degradation (Andreou 2010; Haybron 2011b), and so on. It is possible that all of these factors can mitigate the benefits that increases in income provide.

In contrast to these arguments, other theorists argue that the Easterlin Paradox can be solved insofar as happiness fails to indicate the impact that income has on well-being. Increases in income are clearly beneficial in

---

Although subjective well-being findings often incorporate both findings about happiness and life satisfaction, these two constructs tend to be strongly correlated.

several ways: resulting in increases in health, life expectancy, standard of living, education, access to technology, economic and political freedoms, etc. These benefits may well outweigh the kinds of costs noted above. If so, significant increases in income tend to result in lasting increases in well-being. The findings of the Easterlin Paradox show that such increases in well-being are not indicated by happiness. Perhaps the most prominent advocate of this kind of argument is the development economist Amartya Sen, mentioned in the discussion of adaptive preferences in chapter two. Recall that Sen argues people with a low standard of living can be happy, despite having a low level of well-being.

What can we make of these arguments? Is it the case that happiness fails to indicate the impact that income has on well-being, or is it that income has little impact on well-being? According to the Indicator View, each side of the argument is largely mistaken. Happiness tends to neither successfully nor unsuccessfully indicate levels of well-being, according to the Indicator View. Rather, happiness tends to indicate *changes* in well-being. This is because changes in our well-being tend to be the practically relevant evaluative features in our environment. Thus, we should not expect significant changes in our external circumstances to have a lasting impact on our happiness. Such changes will only tend to be practically relevant for a short period of time, once they have informed and guided our action. Indeed, increases in income have been shown to result in increases in happiness *in the short-term*. (This is the “paradox” part of the Easterlin Paradox, namely that increases in income tends to result in increases in happiness in the short-term, but not the long-term.) After a while, however, we will experience new changes in our well-being which will be practically relevant for us. The former change will cease to have an impact on our happiness. In this way, we should not expect any significant changes in our well-being to have a *lasting* impact on our happiness.

According to the Indicator View, the findings of the Easterlin Paradox are no surprise. Indeed, the findings do not provide us with a paradox. Increases in income tend to have a short-term impact on happiness because happiness indicates environmental features that inform and guide action. Increases in income tend to contribute towards changes in well-being that inform and guide action. However, such changes are only practically relevant for a short period of time. In the long-term, other changes in well-being occur that will be more relevant for immediate action. The upshot is that significant increases in

income (or, indeed, significant changes in most external circumstances) should not tend to make a lasting impact on happiness.

The Indicator View provides a similar interpretation of the phenomenon of adaptation. The phenomenon of adaptation concerns the way in which people seemingly adapt to their circumstances, even if such circumstances are significantly beneficial or harmful to them. This is strikingly illustrated by cases of “adaptive preferences”, which were reviewed in chapter two. As Sen notes:

“The deprived people tend to come to terms with their deprivation because of the sheer necessity of survival, and they may, as a result, lack the courage to demand any radical change, and may even adjust their desires and expectations to what they unambitiously see as feasible.” (Sen, 1999: p.63)

Cases of adaptive preferences involve individuals who are badly off, yet happy. Again, according to the Indicator View, these findings are no surprise. Changes in stable external circumstances tend to have a short-term impact on happiness because happiness indicates environmental features that inform and guide action. Such circumstances tend to contribute towards changes in well-being that inform and guide action. However, such changes are only practically relevant for a short period of time. In the long-term, other changes in well-being occur that will be more relevant for immediate action. The upshot is that changes in stable external circumstances should not tend to make a lasting impact on happiness.

The Indicator View, therefore, gives support to “set-point theory” – the idea that, after reacting to significant changes in certain life conditions (such as winning the lottery, on the up-side, or losing a limb, on the down side), our level of happiness eventually returns to its original “set-point” (Clark et al. 2008; Headey & Wearing 1989; Lykken 1999). According to the Indicator View, after registering changes in well-being, affective states should recalibrate and return to a baseline, in preparation for registering the next change. This is not to say that this baseline cannot change (more on this below). However, we should expect the baseline to be fairly stable over time, resulting in the kind of set points observed in the subjective well-being literature (Headey 2007; Luhmann et al. 2012).

Indeed, one further prediction of the Indicator View is that people’s set-point should tend to be above-neutral. An effective strategy for action is to desire things that are neither too challenging nor too easy to attain (Millgram, 2000). If we only desired things that were too challenging to attain, we would expect

to experience few consistent increases in our well-being, and thereby a low level of happiness over time. In contrast, if we only desired things that were easily attainable, we should expect to experience very consistent increases in our well-being, and thereby a high level of happiness over time. In actuality, we should expect things to be somewhere in the middle: insofar as we desire things that are neither too challenging nor too easy to attain, we should expect to experience fairly consistent increases in our well-being, and thereby a moderate level of happiness over time. This is indeed what positive psychologists have found (Deiner & Biswas-Deiner, 2008; Railton, 2008).

In sum, according to the Indicator View, happiness does not tend to indicate levels of well-being. Happiness tends to indicate whether our lives are getting better or worse, not whether we are doing well or badly. Thus, major changes in one's level of well-being will not tend to be indicated by one's level of happiness in the long-term.

#### **4.4.2 Happiness promoting strategies**

The Indicator View also makes sense of the main strategies for promoting happiness that have been discovered by positive psychologists. Happiness promoting strategies work by increasing the salience of certain dimensions of a subject's well-being. This may either be done directly or indirectly. Directly, interventions may instruct subjects to focus on certain aspects of their well-being, such as recent improvements in their well-being, or particular aspects of their lives that they are grateful for. Indirectly, interventions may either increase the (a) frequency of a subject's improvements in well-being or (b) unexpectedness of a subject's improvements in well-being. I will consider each strategy in turn.

Consider, first, interventions that directly increase the salience of certain aspects of a subject's level of well-being. Perhaps the best-known happiness promoting strategies in the positive psychology literature are various gratitude exercises. Martin Seligman's "What Went Well" task invites subjects at the end of their day to write down three things that went well during the day. Another task invites subjects to write a "gratitude letter" to their friend, relative, lover, teacher, colleague, etc. (Seligman, 2011). Both tasks consist in particular aspects of a subject's level of well-being becoming salient to them. In the "What Went Well" task, the subject is reminded of three events during the day that increased their well-being; in writing a "gratitude letter", the subject is reminded of the value of a particular relationship in their life.

Let us now consider the other two kinds of positive psychology interventions, namely those that increase (a) the frequency or (b) the unexpectedness of a subject's improvements in well-being. According to the Indicator View, in an important range of cases, the function of affective states entails that they detect changes in well-being ingredients. It is no surprise, then, that many happiness interventions attempt to increase the amount of changes in well-being ingredients experienced by a subject. For instance, Teresa Amabile and Steven Kramer suggest that focusing on increasing the frequency of "small wins" throughout one's day makes one happier than focusing on less-frequent larger successes (Amabile & Kramer 2011). Similarly, Kennon Sheldon and Sonia Lyubomirsky's "Happiness Adaption Prevention" (HAP) model emphasises that positive changes in well-being that are continually appreciated are most likely to produce continued happiness (Sheldon & Lyubomirsky 2012). Put simply, increases in well-being tend to result in temporary happiness – the more increases we experience, the happier we feel.

In certain contexts, according to the Indicator View, the function of affective states entails that they detect unexpected changes in well-being; thus, other happiness promoting strategies aim to increase the unexpectedness of a subject's improvements in well-being. Sheldon and Lyubomirsky's HAP model also emphasises that varied increases in well-being tend to result in more lasting happiness than fixed increases in well-being. This is, presumably, because more varied changes in well-being are less predictable – they are less expected. The more unexpected a change in well-being, the greater the practical relevance of that change. The unexpectedness of varied circumstances can account for the effectiveness of other happiness promoting strategies. Seligman et al.'s "Novel Use of Strengths" exercise consists in subjects identifying their five "signature strengths" and using those strengths in new and different ways every day. In contrast to a similar exercise that consists in subjects using their signature strengths in non-varied ways, the more varied exercise was effective at improving happiness and depression scores for the six months that the exercise took place (Seligman et al. 2005).

Indeed, one of the most effective strategies for promoting happiness involves increasing both the frequency and unexpectedness of a subject's improvements in well-being. These strategies focus on producing a greater sense of "flow" in a subject's everyday activities. Feelings of flow tend to arise when a subject is successfully carrying out an activity that is both skilful and challenging. The activity is not challenging to the extent that the subject does



not make consistent progress. And the subject is not skilled in the activity to the extent that they find it too easy. Rather, the challenge of the activity and the skill of the subject are at just the right levels so that the subject improves frequently. Moreover, because of the level of skill involved in the activity, the improvements tend to be both complex and varied. Imagine, for example, the variation in types of rock and rock-faces experienced by a professional rock-climber. Not surprisingly, cultivating experiences of flow has been shown to be one of the most effective ways of promoting lasting happiness (Csíkszentmihályi 2008). Cultivating such experiences tends to result in certain aspects of one's well-being (namely, the frequent and varied improvements being made in the respective activity) being consistently relevant for action.<sup>89</sup>

#### **4.4.3 The key determinants of lasting happiness**

The conditions and circumstances that make a lasting difference to a subject's happiness also tend to be the three dimensions of a subject's well-being (outlined above) that should be detected by affective states. In accordance with the Indicator View, the key determinants of lasting happiness can be divided into three broad kinds, namely conditions that tend to cause: (a) certain aspects of a subject's level of well-being to be salient; (b) frequent changes in well-being; (c) unexpected changes in well-being. I will consider each of these kinds of conditions in turn.

Firstly, conditions and circumstances that tend to cause certain aspects of a subject's well-being to be salient have a lasting impact on a subject's happiness. This, of course, makes sense according to the Indicator View. For example, gratitude exercises that make one focus on various aspects of one's level of well-being (e.g. one's friends, family, health, financial security, etc.) tend to make us feel happier. However, it is not obvious that there are stable conditions that continually make certain aspects of our well-being salient to us. Recall the evidence outlined above regarding the fleeting happiness and unhappiness of lottery winners and disabled patients respectively. It makes sense for an individual to focus on (unexpected)

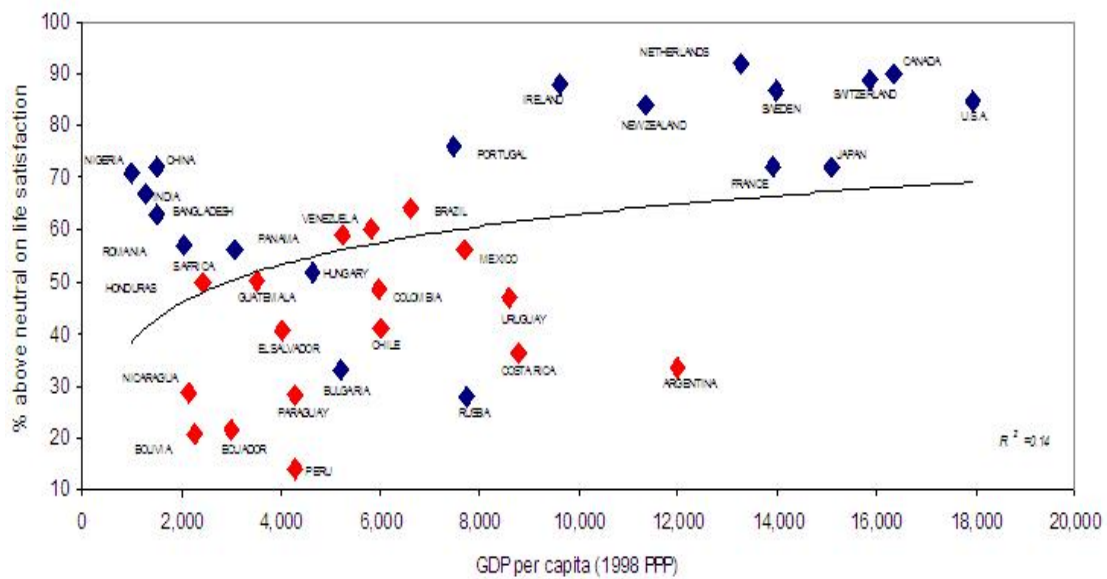
---

<sup>89</sup> Note that many of these happiness-promoting strategies do not necessarily result in directly promoting one's well-being (in addition to promoting one's happiness). However, happiness has been shown to have a variety of beneficial effects on one's attention, cognition and motivation (Fredrickson, 2001; Luyobmirsky et al, 2010). Thus, happiness-promoting strategies may tend to result in indirectly promoting one's well-being. I will consider this issue in more detail in Appendix A.

changes in their well-being in order to effectively promote their well-being. Continuously focusing on stable conditions that relate to certain aspects of one's level of well-being is likely to be practically ineffective.

Nonetheless, there is some evidence that certain conditions make a lasting impact on the salience of certain aspects of our level of well-being. For example, Graham and Pettinato (2002) show that, on average, rich countries are happier than poor countries, but that this trend is negligible after average incomes of \$10,000:

Fig 4.1



The common explanation given for these findings is that income makes a lasting impact on happiness when we are unable to fulfil our basic needs for nutrition, shelter, survival, health, etc. (Layard, 2005; Diener & Biswas-Diener, 2008) When our basic needs are not being met, these aspects of our lives become continuously practically relevant to us. If this interpretation of the findings is true, certain conditions can make a lasting impact on the salience of certain aspects of our level of well-being and thereby make a lasting difference to our happiness.

The second broad kind of conditions and circumstances that have a lasting impact on a subject's happiness are those that tend to cause frequent changes in well-being. Perhaps the most interesting conditions in this respect are those that result in "positive cycles" of positive affect and the condition in

question (Fredrickson, 2001).<sup>90</sup> That is, certain conditions tend to cause increases in well-being that make us temporarily happy, which then tends to cause us to further promote such conditions, which cause increases in well-being that make us momentarily happy, and so on. These kinds of conditions tend to result in upward spirals of improvements in well-being and thereby lasting increases in happiness.

For instance, friendship tends to make a lasting impact on happiness (Kahneman & Deaton, 2008). This may partly be due to the positive cycles of friendship and positive affect (Bishop, MS).<sup>91</sup> Friendship makes us feel good, but feeling good also makes us friendlier. For example, studies show that feeling good makes you more likely to initiate a conversation with a stranger (Isen 1970), offer intimate self-disclosures to others (Cunningham 1988), and judge the people you spend time with as “kind, self-assured, open, tolerant, warm” (Lyubomirsky & Tucker 1998). In general, the friends and family of happier people judge them to be more “socially skilled (e.g., more articulate and well mannered), better public speakers, self-confident, and assertive, and as having more close friends, a strong romantic relationship, and more family support” (Lyubomirsky et al. 2005: p.827, see text for citations). In short, maintaining our friendships makes us happy, and being happy makes us maintain our friendships. The result is that friendships tend to result in consistent increases in well-being.

There is evidence that upward spirals of well-being are caused by other circumstances and conditions, including intimate relationships (Fredrickson, 2001), autonomy and productivity (Côté 1999), coping styles (Fredrickson & Joiner 2002), optimism (Seligman 2006), engagement (Deci & Ryan 2000) and altruism (Thoits & Hewitt 2001). These conditions tend to be key determinants of lasting happiness (Diener & Biswas-Diener, 2008).

The third broad kind of conditions and circumstances that have a lasting impact on a subject’s happiness are those that tend to cause unexpected changes in well-being. I will consider four conditions here, namely

---

<sup>90</sup> I will consider the implications of these positive cycles for the measurement of well-being in more detail in Appendix A.

<sup>91</sup> Of course, there are other explanations of the lasting impact that friendship tends to make on happiness. For example, it may be that those with good friendships also tend to have a positive affective disposition. Nonetheless, I think there is good evidence that positive cycles of affect occur with regards to friendship, and that this can partly explain the relationship between friendship and happiness.

unemployment, age, control and noise. The first two conditions have an impact on happiness due to the expectations produced by social references i.e. information about our own well-being provided by the well-being of others. When one becomes unemployed, one is surrounded by people who are doing better, at least financially. The employed continuously provide the unemployed with information about how well they could be doing financially, thereby increasing one's expectations for higher levels of well-being. These higher expectations are often not met (it is often hard for the unemployed to find employment again) thereby resulting in lasting unhappiness.<sup>92</sup>

A similar mechanism may be at work with regards to the impact that age has on happiness. The relationship between happiness and age tends to follow a U-shaped curve, whereby subjects tend to be happiest when young and old, and unhappiest in their middle-age (Blanchflower & Oswald 2008). This may largely be due to changes in expectations throughout a subject's lifespan. During middle-age, one is surrounded by people who have succeeded in certain aspects of life (career, status, relationships, etc.) which may have the effect of raising one's expectations for achieving such success. Middle-age is often viewed as one's "prime-of-life" (Slote 1983) – the period of one's life in which one has the most amount of time and resources to achieve success, and thereby expects to achieve one's highest level of well-being. In contrast, at old age, one is not expected to achieve great success – staying healthy, relatively autonomous and being a supportive family member will often suffice.<sup>93</sup>

---

<sup>92</sup> The fact that unemployment does not have such a lasting impact on unhappiness within a recession supports this interpretation (Graham, 2010). Within a recession, the unemployed are numerous, which creates the impression that the unemployed may not in fact be able to do better financially.

<sup>93</sup> Three further lines of evidence support this interpretation. The first concerns similar studies done on great apes, which shows that our animal cousins also tend to be unhappiest at middle-age (Weiss et al. 2012). Again, this is not surprising – it is during middle-age in which apes are expected to reach their highest status and reproductive fitness.

The second line of evidence concerns people's value changes over the lifespan (Reed & Carstensen 2012). During middle-age, when people have more time and resources, they tend to value more "extrinsic" values, such as income, status, etc. In contrast, in old-age, when people have less time and resources, people tend to value more "intrinsic" values, such as relationships, leisure, etc. It is reasonable to think that value changes largely occur due to changes in what is possible (i.e. in one's expectations of well-being).

The remaining two conditions that I will consider in this section have an impact on happiness due to expectations that are independent of social references. These conditions are control and noise. A sense of control is one of the largest determinants of happiness/unhappiness (Diener & Biswas-Diener, 2008). This makes perfect sense according to the Indicator View. Negative changes in one's well-being that are unexpected will tend to be more relevant for action than expected negative changes. A lack of control typically comes from experiencing negative changes that one does not expect.<sup>94</sup> The impact of noise on unhappiness is similar in this respect. Background noise is well documented as having a pervasive affect on unhappiness, but it is important to realise that only unpredictable noise (i.e. noise that cannot be expected) has this affect (Railton, MS). One only needs to think about how annoying sporadic bursts of building work outside your office can be to appreciate this affect. This also explains why one-sided mobile phone conversations (which consist in largely unpredictable bursts of conversation) can be so irritable to the third-person observer, in contrast to standard two-sided conversations in the background.

We have seen that, according to the Indicator View, the key determinants of lasting happiness can be divided into three broad kinds, namely conditions that tend to cause: (a) certain aspects of a subject's level of well-being to be salient; (b) frequent changes in well-being; (c) unexpected changes in well-being. An understanding of these three causal factors can enable us to interpret why it is that certain circumstances and conditions do and do not have a lasting impact on happiness. In addition, they can enable us to interpret why particular conditions have varying kinds of impacts on happiness. For example, it is possible that income does not tend to have a

---

The final line of evidence concerns the affect that relative income has on happiness. One study shows that relative income has an impact before the age of 45, but not after (Fitzroy et al. 2013). A plausible explanation of this finding is that, before the age of 45, people tend to expect to attain a similar level of income ("I could achieve that!"); in contrast, after the age of 45, people may tend to not expect to attain much greater levels of income ("I could have/can no longer achieve that"). Thus, relative income has a lasting impact on unhappiness before the age of 45, but not after.

<sup>94</sup> Indeed, when one comes to expect having a lack of control, it tends to cease to cause unhappiness altogether, in a phenomenon referred to as "learned helplessness." Those suffering from learned helplessness no longer respond to negative events out of their control – they seemingly become indifferent to such events. Although learned helplessness is largely considered to be a behaviour phenomenon, it seems reasonable to assume that it also an affective component.

lasting impact on happiness because increases in income do not tend to cause frequent (unexpected) increases in well-being. However, it is also possible that income can have a lasting impact on happiness insofar as one's income is spent in ways that cause frequent (unexpected) increases in well-being (Dunn et al. 2011). Moreover, it is possible that, below a certain level, income may tend to have a lasting impact on happiness; when one's basic needs are not being met, one's poor standard of living may continuously be practically relevant. In this way, the Indicator View can provide us with both plausible and informative interpretations of empirical findings from the study of happiness.

## **Conclusion**

Over this and the previous chapter, I have developed the Indicator View on the basis of the function of affective states. The function of affective states entails that they detect practically relevant evaluative features of our environment. I have argued that a number of different kinds of evaluative features of our environment tend to be practically relevant. In particular, I have discussed how the function of affective states entails that they detect (a) non-prudential features of our environment (e.g. moral, aesthetic or epistemic features); (b) changes in well-being ingredients; (c) unexpected changes in well-being ingredients; and (d) levels of certain well-being ingredients. I believe that these different kinds of evaluative features account for the majority of the evaluative features of our environment that tend to be relevant for action and thereby reliably cause us to experience affective states.

The fact that the function of affective states incorporates these different kinds of evaluative features is reflected in Valerie Tiberius' account of the different perspectives that we can have on our lives. Tiberius notes:

“Because we have many different commitments, and because each commitment is comprised of a pattern of attitudes that can wax and wane, we can be in different perspectives at different times. We can, for example, take a reflective perspective when we are feeling contemplative and we want to think about how our lives are going. Alternatively, when we are absorbed in a project, we take a perspective that focuses on this project and excludes almost everything else. The things that change when our perspective changes are what is most salient to us, which facts are deemed relevant considerations, and what motives are most available.” (Tiberius 2008; p.68)

The claims of this chapter can be viewed as claims about the kind of perspective that we tend to have in our practical lives. I have suggested that, in an important range of cases, this perspective is that of improving our well-being. This perspective requires that affective states reliably detect changes in our well-being. We have good reason, therefore, to treat happiness as a defeasible indicator of changes in well-being.

## Chapter 5

### Happiness as an Indicator of Local Well-being

#### Abstract

In this chapter, I will further refine the Indicator View. I will argue that happiness is a defeasible indicator of *local changes* in well-being (i.e. moment-to-moment changes in our objectives, goals and concerns). I will argue that some of the local changes indicated by happiness have intrinsic prudential value. These changes tend not to be indicated by more global measures of well-being, such as measures of life satisfaction. Yet, I will argue that how well our lives are going on this local level intrinsically matters for our overall well-being. We can, therefore, successfully measure an important aspect of well-being through measuring happiness.

I will argue in favour of this view as follows. In the first section of the chapter, I will argue that we have reason to believe affective states tend to correlate with local changes in well-being; it is local changes that tend to be relevant for action. In contrast, I will suggest that judgements of life satisfaction tend to correlate with global changes in well-being (e.g. long-term changes in our goals, values and projects). Therefore, we should treat happiness as an indicator of local changes in well-being, and life satisfaction as an indicator of global changes in well-being.

In the remainder of the chapter, I will consider the value of the local changes in well-being that tend to be indicated by happiness, and the implications that this has for the measurement of well-being. Some of the local changes in well-being indicated by happiness have instrumental prudential value; some have intrinsic prudential value. In the second section of the chapter, I will consider the local changes in well-being that have instrumental value. Can we successfully measure well-being over time through measuring such changes? I will argue that we cannot. In particular, I will argue that we cannot accurately approximate a person's overall change in well-being over time by aggregating the value of their local changes in well-being (indicated by happiness) over that time.



In the third section of the chapter, I will consider the intrinsic value of the local changes in well-being that tend to be indicated by happiness; these changes constitute what I will call *local well-being*.

In the fourth section of the chapter, I will consider the extent to which local well-being constitutes overall well-being. Some theorists argue that our global (in contrast to local) well-being has a distinct kind of value, in virtue of the fact that it is constituted by meaningful life narratives or concerns. I will argue against this view. I will argue that both local and global well-being constitute overall well-being in the same way; neither kind of well-being has a distinct kind of prudential value. In the fifth section of the chapter, I will suggest that the relative values of people's local and global well-being vary across their lifespan, personality and cultural environment. The upshot is that we can successfully measure changes in an important aspect of well-being through measuring happiness.

## 5.1 Happiness as a defeasible indicator of local changes in well-being

In the first main section of the chapter, I will argue that, in an important range of cases, the evaluative features of our environment that are relevant for action tend to be local changes in our well-being. Local changes in well-being concern how well we are doing from moment to moment: in meeting our current needs, achieving our short-term goals, connecting with those around us, solving a pressing problem, etc. In contrast, global changes in well-being concern how well we are doing in general: in achieving our long-term goals such as our personal projects, lifelong ambitions, achieving the respect of others, making meaningful contributions or accomplishments, etc. Happiness is a fast-responding, sensitive indicator of local changes in well-being that guides our efforts towards success and away from failure or frustration (Railton, MS). In contrast to happiness, I will argue that *life satisfaction* is a broad indicator of global changes in well-being.

### 5.1.1 Affective states and local changes in well-being

In chapter four, I argued that happiness tends to correlate with changes in well-being. The function of happiness, in an important range of cases, entails that it detects (unexpected) changes in well-being ingredients. I argued that this provides us with good reason to treat happiness as a defeasible indicator of changes in well-being.

However, there are different kinds of changes in well-being, each of which are relevant for action in different respects. Two kinds of changes in well-being are especially important, and have recently been highlighted by Peter Railton in his discussion of the function of happiness (Railton 2008; MS). Roughly put, we need to act both on a specific, moment-to-moment level, and on a more general, long-term level. To help gain an understanding of these two modes of action, consider the following vignette provided by Peter Railton:

“You’ve just gotten off the phone with your sister, who told you that, when the visiting nurse dropped in on your aging parent, he realized that the oven had been left on for many hours, carbonizing the food within, which had evidently been completely forgotten. The smell of burnt food filled the house, yet your mother was completely unaware. “We need to do something—we can’t just leave her alone at home anymore,” your sister says, “How soon can you get up there to start working something out?” You feel stricken, your sister has borne the brunt of looking after your mother for the last few years, and it is

clearly your turn to pitch in in earnest...You realize you've finished your teaching for the week and can fly up to visit her. Now you feel a boost of positive energy and rush to call the airline. But they put you on what feels like perpetual hold. Your frustration growing, you cast around for an alternative. "Ah, I can call my travel agent!" Another burst of positive energy. But you realize you don't have his number at the office. You think hard. Maybe you can remember it. Is it 734-677-0900? You brighten. Or maybe it's 734-677-0090? You're impatient—you have to try something. So you try the first, and you get the travel agency's familiar phone tree. Great! You relax your tension just a bit. But when you ask for your agent, she's at lunch. "Damn! Can't waste time while the last seats might be sold. Now what?" You remember your new phone lets you go on the web, though you hate figuring this sort of thing out and never had the time or, in truth, the desire to learn how. Now you are rushed and anxious, but highly motivated, and you make yourself work methodically at it, working your way through the screens, concentrating intensely, trying to figure out what would make sense. You're greatly relieved and a bit proud that you manage to make it all the way onto the web, and start booking a seat." (Railton, MS: p.39)

In this vignette you have received news of several global changes in your well-being: the health and welfare of your parent, the need to help her in old age, and being fair to your sibling. This news instantly changes your current priorities for action, requiring a radical shift in your practical perspective. Yet, Railton notes that you do not dwell on the news itself. Once you have taken in its import, you stay attuned to local changes in your well-being: the more specific, moment-to-moment aspects of your situation. As Railton notes, your affective states supply, "positive energy for the next step, frustration when you can't move forward and associated motivation to look hard for alternatives, as well as some gratification and pride at making headway and taking matters in hand." (Railton, MS: p.40-41)

We need our momentary affective states to be relatively separate from mental states that detect more general, long-term changes in our well-being. More global changes may dictate where we need to go in life, but they do not tend to inform us about how to get there. Actual progress tends to require specific actions, taken from moment-to-moment. The function of affective states may primarily be to inform and guide our action on this level.

Affective states must be continuously sensitive to variations in circumstance, ability, or opportunity. As Railton notes, "In the course of a day, we may need

to take thousands of steps towards meeting our needs, fulfilling our responsibilities, advancing our goals, attending to our concerns, or simply managing not to lose ground.” (Railton, MS: p.41) Our affective states help alert, guide, motivate, and reward us along the way.

Without such information and guidance, we are unlikely to improve our circumstances. This is most notably seen in affective disorders, such as mania and depression. Manic individuals have an overwhelming balance of positive over negative affect, which tends to disrupt the normal regulation of action. Such individuals tend to be impulsive and erratic; their affective states are effectively continuously signalling to them, “you’re doing the right thing, keep at it.” In contrast, depressed individuals have an overwhelming balance of negative over positive affect, which disrupts well-functioning in a different yet equally damaging way. Such individuals tend to be indecisive, ruminative and withdrawn; their affective states are effectively continuously signalling to them, “you’re doing the wrong thing, give this up.” These affective disorders illustrate why affective states have the function of informing and guiding our action on a specific, moment-to-moment level. We need to act on this level in order to function well and make progress in our lives.

I will refer to specific, moment-to-moment changes in well-being as *local* changes in well-being. In contrast, I will refer to more general, long-term changes (such as a loved one being in ill-health) as *global* changes in well-being. Most of our actions take place on a local level, from moment-to-moment. Thus, in informing and guiding our action, our affective states will tend to detect local changes in well-being. This is why, as mentioned in chapter one, our emotions (which tend to detect local changes) will generally dominate our affective state, in contrast to our moods (which tend to detect global changes). The result of this tendency is that happiness will not tend to indicate more global changes in well-being. In short, happiness tends to indicate local, rather than global, changes in well-being.

### **5.1.2 Mood, life satisfaction and global changes in well-being**

In this section, I will suggest that, in contrast to happiness, *life satisfaction* tends to indicate global changes in well-being.

We have seen that happiness indicates local changes in well-being for the following reason: in order to be well-functioning agents, we need to continuously adjust our behaviour in accordance with specific, short-term changes in our circumstances. Local changes in well-being may not be the most important changes in our lives, but they are the ones that generally

dictate our action – it is through such changes that we can achieve more substantial (global) changes. Local changes tend to be instrumental. To return to Railton’s vignette above, you care about booking a plane ticket in order to see your parents and help out your sister. The value of booking a plane ticket is instrumental; what you really care about is helping your parent and being fair to your sibling. Thus, we can view happiness as an indicator of local changes in well-being, which in turn reflects how well we are doing at taking the steps necessary for the achievement of more global changes. Happiness tends to indicate whether or not we are getting to the place we want to go, rather than whether or not we have got there.

This, however, is not the full story of the function of affective states. In addition, we require affective states that inform and guide our action on a more global level. To illustrate the need for a more global indicator, let us again return to Railton’s vignette above. Railton notes that, upon hearing that your mother’s living condition will need to be changed, you may realize that it no longer is enough to count on the occasional visits of nurses or relatives (Railton, MS: p.42). News of her ill-health symbolises that significant changes in your life are about to occur: in both your mother’s well-being and perhaps in many other features of your life. Such changes require a global level of concern for the new situation. In contrast to more local changes, you may not be able to deal with this change through carrying out a few specific tasks within a relatively short period of time. As Railton notes: “During this time you must have some background dissatisfaction or lack of complacency, making possible emergent changes in orientation of thought and effort on a continuing basis, as appropriate.” (Railton, MS: p.42)

This more global state of concern and readiness to act must persist without interfering with the detection of the momentary ups and downs relevant to specific activities and objectives. Thus, what is needed is not a fast-responding, fast-dissipating affective state (i.e. an emotional state). Rather, we need to change our *mood*. Moods set the tone for our life without entirely interfering with out emotional states. As Laura Sizer puts it: “[mood] is more subtle [than emotion], acting as background to our ongoing activities. [It] has more to do with the way one approaches life as a whole than with reactions to particular objects or events.” (Sizer, 2010: p.147)

In contrast to emotional states, moods tend to indicate more global changes in well-being. This does not necessarily mean that moods *represent* more global evaluative features of our environment. I do not wish to take a stand

on the representational nature of moods.<sup>95</sup> However, I do mean to claim that moods have the function of guiding our action on a global level in response to important changes in our environment.

Although moods do not entirely interfere with our capacity to experience momentary affective states, they do impact our propensities to attend, think, feel and act in certain ways. For example, as discussed above, after hearing about a loved one being in ill-health, you may be more anxious and distractible, less satisfied with life (Railton: MS). The effects of moods are persistent and pervasive. Dan Haybron illustrates this fact by noting the effects of four different moods, namely being depressed, elated, anxious and serene: "While in a depressed mood...an individual will likely find little pleasure in what happens, will tend to look on the dark side of things, and may more likely be saddened by negative events. The elated person will exhibit the opposite tendencies. And someone afflicted by anxiety will tend to multiply and exaggerate potential threats, experience greater upset at setbacks, and be more prone to experience fear and perhaps anger. Whereas a more serene individual will tend to take things in stride, see fewer causes for anxiety, worry less about perceived threats, etc." (Haybron 2005: p.17) These examples all illustrate a general feature of moods, namely that they dispose us to experience mood-congruent emotions.

Indeed, the direction of causality between moods and emotions goes both ways. Our mood disposes us to experience mood-congruent emotions, and repeated or intense emotions can contribute to the onset of a mood (Sizer, 2010; Prinz, 2004). For example, if we are continually doing well on a local level, we will tend to feel good, and such feelings may cause us to be in a good mood. Information about how well we are doing from moment-to-moment may have implications for how well we are doing more globally. Thus, although moods do not entirely interfere with our capacity to experience momentary affective states, we should expect moods and emotions to be

---

<sup>95</sup> On the one hand, some theorists argue that moods are simply generalised forms of emotion, and thereby represent general/global evaluative features of our environment, rather than specific/local features (DeLancey 2006; Prinz, 2004). On the other hand, other theorists argue that moods are different kinds of affective states than emotions, in that they can be objectless (Sizer 2000; Siemer 2009). Such theorists argue that moods are fundamentally dispositional states, which influence our attention, cognition and behaviour in various ways without representing evaluative features of our environment. Theorists on both sides of this issue agree that moods can be distinguished from emotions in virtue of the fact that they are more diffuse and pervasive in their effects.

closely related. More precisely, we should expect that positive moods (such as elation or contentment) tend to correlate with positive emotions (such as joy or satisfaction) and negative moods (such as depression or anxiety) tend to correlate with negative emotions (such as sadness or fear). The fact that emotions and moods are causally connected means that how well we are doing on more local and global levels are likely to influence each other.

Global changes in well-being concern how well we are doing on a more global level: the achievement of life goals, material security, personal projects, long-term ambitions, developing intimate relationships with others, etc. I have argued that global changes in well-being do not tend to be indicated by happiness. Rather, in the remainder of this section, I will suggest that such changes tend to be indicated by an altogether different mental construct, namely *life satisfaction*.

Life satisfaction is measured by empirical researchers who ask subjects to judge how satisfied they are with their life (or a particular domain of their life) overall. It seems reasonable to presume that such judgements are formed on the basis of the subject's recently experienced moods (or even current mood). In arguing for this view, Michael Robinson, notes that, "rather than making complex evaluations of current life circumstances, the data in the present investigation suggest that people can simply index their mood states." (Robinson 2000: p.165) That is, moods seem to mediate the relation between global events and life satisfaction: how people judge their lives may depend largely on their moods.<sup>96</sup>

This suggests that we can view life satisfaction as a kind of *meta-mood* – feelings of general satisfaction or dissatisfaction with our lives caused by global changes in our well-being. When we are in a bad mood, we are less satisfied with our lives, and for good reason – our bad mood may reflect the occurrence of a recent global change for the worse in our well-being, such as a loved one being in ill-health. As a result of this change, we may feel that our lives are not good enough – we are dissatisfied. Conversely, when we are in a good mood, we are more satisfied with our lives; we may have just completed a project that we have been working on for some time, or spent

---

<sup>96</sup> This also explains why measures of life satisfaction and happiness tend to be strongly correlated (Diener et al. 2012). Moods tend to cause mood-congruent emotions. Thus, moods influence both our emotions and our life satisfaction. In addition, persistent kinds of emotion can cause the onset of a mood. Insofar as life satisfaction is influenced by our recent moods, therefore, it will also be influenced by our emotions.

time with a good friend. The result is that we may feel that our life is going sufficient well – we are satisfied.

Of course, in certain contexts, life satisfaction may fail to indicate global changes in well-being. Nobert Schwarz and Fritz Strack (1999) have shown that judgements of life satisfaction are often biased by recent affective states caused by seemingly trivial prudential factors, such as the weather or the fortunes of the national football team. In addition, Haybron (2000) argues that life satisfaction may be largely influenced by ethical norms, such as norms of gratitude (“My life is good enough”) or non-complacency (“My life could be better”). For now, it is enough to recognize that, although in certain contexts life satisfaction may not tend to indicate global changes in well-being, it may nonetheless tend to do so outside of such contexts. In short, life satisfaction may be a defeasible indicator of global changes in well-being.

We have, then, the following picture of indicators of well-being based on the mental states that inform and guide action: happiness (largely constituted by emotions) tends to indicate local changes in well-being; life satisfaction (largely influenced by moods) tends to indicate global changes in well-being. There is empirical support for this picture with regards to the different key determinates of happiness and life satisfaction. Conditions that have a consistent immediate impact on our lives, such as our health or the quality of our relationships, tend to correlate with happiness. In contrast, life satisfaction tends to correlate with long-term conditions, such as one’s level of income or educational attainment (Kahneman & Deaton, 2010).

In the next few sections, I will consider the implications that this picture has for the project of measuring well-being through measuring happiness. In the next two sections, I will consider the instrumental and intrinsic value of the local changes in well-being that tend to be indicated by happiness. I will argue that such changes may tend to have intrinsic, not instrumental, prudential value.

## **5.2 The instrumental value of the local changes in well-being indicated by happiness**

Perhaps the most obvious way in which specific, moment-to-moment changes in well-being are valuable is that they have instrumental prudential value. Consider, for example, your average day: you get up, get ready for work, feed yourself, commute to work, check your emails, talk to a colleague, make



varying levels of progress in on-going objectives and other items on your to-do list, have a break, feed yourself, etc. You carry out many (if not most) of your actions and activities in order to achieve some other aim, such as being healthy, completing a project or developing a relationship. And such actions and activities tend to pay-off, otherwise we would stop trying so hard to improve our lives or instead focus on maintaining what we already have. These lines of thought lead us towards the claim that local changes in well-being tend to be valuable largely because they bring about more global changes in well-being.

The instrumental value of local changes in well-being suggests that we can calculate (or at least roughly approximate) how well someone is doing in a more global sense by aggregating their local changes over time. That is, the sum of the local changes in well-being indicated by happiness over time may be a good approximation of a person's overall change in well-being over that time. This method of calculating people's well-being from information about happiness over time has much in common with Daniel Kahneman's notion of Objective Happiness (Kahneman 1999). A person's Objective Happiness is calculated by aggregating his or her momentary levels of affect over time. Kahneman argues that such information tells us about how well that person was doing over that period of time (Kahneman et al. 1997).

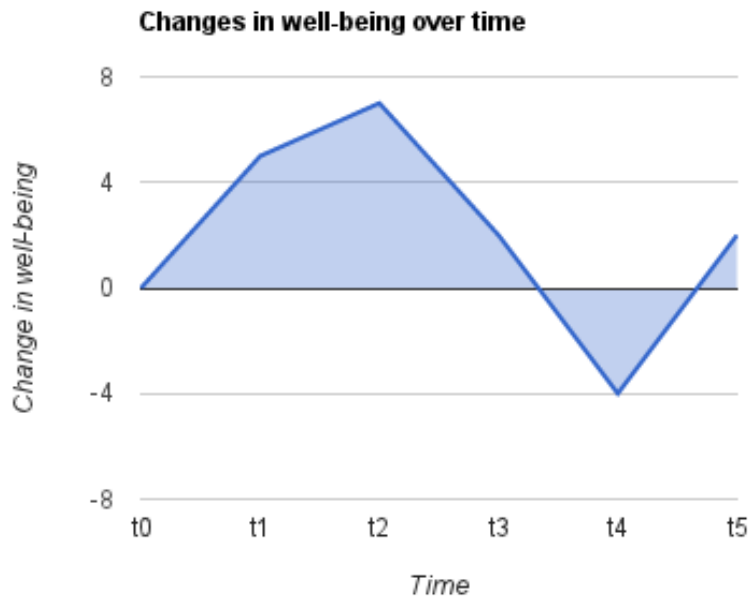
In this section, however, I will argue that this method is implausible for three reasons. The first two reasons reflect the fact that happiness is not a reliable indicator of the instrumental prudential value of local changes in well-being. First, happiness does not indicate the size of local changes in well-being – it merely indicates a person's amount of local changes experienced over time. Second, happiness may indicate “false positives” – local changes in well-being that turn out not to have instrumental value. The third reason is that global changes in well-being over a certain period of time may not be caused by local changes in well-being over that time.

### **5.2.1 Aggregating the size of local changes in well-being over time**

To illustrate the method in question consider the following example: Suppose that at time  $t_0$  my level of well-being is zero. At time  $t_1$  my level of well-being increases by 5. At time  $t_2$  my level of well-being increases by 7. At time  $t_3$  my level of well-being increases by 2. At  $t_4$  my level of well-being decreases by 2. And at  $t_5$  my level of well-being increases by 2. These changes in well-being have been illustrated in the graph below (fig 5.1). In order to calculate the total change in well-being from  $t_0$ - $t_5$ , we need to subtract the area of the

graph below the zero point from the area of the graph above the zero point. The total area corresponds to the total change in well-being over that time.

Fig 5.1



Perhaps the main problem with this method is that affective states do not tend to indicate the *size* of changes in well-being. Positive (negative) affective states indicate increases (decreases) in well-being, but they do not tend to indicate whether they are large or small increases (decreases). The strength of an affective state should be caused by the extent to which a particular change in well-being is relevant for action. The practical relevance of a change in well-being does not tend to coincide with the size of that change.

For example, suppose that, throughout the course of a day, you worked on a big problem and eventually solved it. Let us suppose that solving the big problem consisted in solving ten smaller problems throughout the day, and that solving each of them caused an episode of pleasure. Further suppose that each of the ten smaller problems were entirely of instrumental value – solving them was only valuable insofar as they contributed towards solving the big problem. Now, if affective states indicate the size of changes in well-being, this would imply that the episode of pleasure caused by solving the big problem would have ten times the strength as the episodes of pleasure caused by solving each of the ten smaller problems. This, however, is clearly implausible. The strength of our affective states is not entirely proportional to the value of the things they indicate. Rather, the strength of our affective states is proportional to the extent to which we should act differently. The

reason we don't feel ten times as much pleasure from solving the big problem as we did from solving any of the ten smaller problems is because solving the big problem is not ten times more relevant for action. We may need to act somewhat differently as a result of solving the big problem (let other people know about the solution, think about what problems we need to solve next, whether the solution can be applied to those problems, and so on); but, it is unlikely that such actions will require a substantially greater amount of effort than any of our previous actions throughout the day.

One might object that, in claiming affective states do not tend to indicate the size of changes in well-being, I am confusing the strength of an affective state with its level of arousal. The level of arousal of an affective state is the extent to which an affective state consists in excitable physiological reactions or action tendencies. For example, being elated or stressed is a high-arousal state, whereas being calm or fatigued is a low-arousal state. One might agree that, in the example above, solving the big problem should not cause an affective state with a level of arousal that is ten times that of the affective states caused by solving any of the smaller problems – solving the big problem clearly does not require such high levels of excitable physiological reactions or action tendencies. However, one might argue that, independent of the level of arousal of an affective state, there is a component of an affective state that does tend to indicate the size of changes in well-being.

This objection finds support in James Russell's "circumplex" model of affective states, whereby affective states can be categorised on the basis of their (typical) level of arousal and "hedonic tone" (Russell et al. 1989). One could claim that the hedonic tone of affective states tends to indicate the size of changes in well-being. For example, typical episodes of relaxation and contentment both consist in low arousal affective states with a positive hedonic tone. Yet, a typical episode of contentment has a higher positive hedonic tone than a typical episode of relaxation. This suggests that typical episodes of contentment tend to indicate greater increases in well-being than typical episodes of relaxation. If this is correct, this lends support to the idea that the strength of affective states – conceived of as its level of hedonic tone – tends to indicate the size of changes in well-being.

I agree that the level of arousal and the hedonic tone of an affective state can operate independently of each other. The level of arousal of an affective state has the function of mobilising an effective response to the given situation – one that is either excitable or non-excitable. In contrast, the hedonic tone of

an affective state has the function of representing how relevant the given situation is for (either excitable or non-excitable) action.

However, I still disagree that the strength of affective states tends to indicate the size of changes in well-being. This is because how relevant a given change in well-being is for action is not solely determined by the size of that change. All things being equal, we should expect the strength of affective states to be caused by the size of the changes in well-being. However, all things do not tend to be equal, for at least two reasons.

First, unexpected changes in well-being are of greater practical relevance than expected changes. This is why we can experience so much pleasure from something relatively trivial yet unexpected, such as receiving a small gift from a stranger. Conversely, we can fail to experience much pleasure from something relatively important yet expected, such as finally achieving the goal that we have worked so hard for. In general, if we expect regular increases in well-being then decreases in well-being will be more relevant for action, and vice versa. As Haybron notes: "A Manhattan debutante may consider a day without her cell phone about as bad as it gets, while a crippled resident of Manila's trash dumps is liable to have rather different standards." (Haybron 2007: p.403) Whereas the person in Manhattan may expect her well-being to generally improve throughout the day, the person in Manila may expect the opposite. The result is that infrequent decreases in well-being will be practically relevant for the former individual, and the occasional increase in well-being will be practically relevant for the latter individual. In short, our expectations partly determine both the direction and size of changes that are practically relevant.

Secondly, certain changes in well-being require us to act considerably differently (again, either in an excitable or non-excitable way) whereas others do not. For instance, certain small changes in well-being may require us to act considerably differently (and thereby tend to cause strong affective states), whereas certain large changes may not. I think this is exactly what happens with regard to long-term goals. Long-term goals tend to be unobtainable in themselves. For example, running a marathon off the cuff would be extremely difficult and unhealthy. Such goals are not relevant for action until they are split up into short-term goals: sub-goals (running a half marathon), objectives (going for a run three times a week), tasks (cooking healthy food), and so on. It is these short-term goals that are relevant for action. Success or failure with regard to such goals tends to cause strong

affective states, even though such achievements constitute relatively small changes in well-being. In comparison, success or failure with regard to the long-term goals of running a marathon tends to cause stronger affective states, but not in proportion to the much larger change in well-being that such an achievement constitutes. In general, changes in well-being that have instrumental value are likely to be practically relevant (and thereby cause strong affective states) despite being relatively unimportant.

The upshot of these two points is that the strength of an affective state does not tend to indicate the size of the change in well-being that it indicates. We cannot, therefore, use the method illustrated by the graph above to aggregate local changes in well-being indicated by a person's affective states over time. The method above presumes that we know the following information about local changes in well-being:  $t1=+5$ ;  $t2=+7$ ;  $t3=+2$ ;  $t4=-2$ ; at  $t5=+2$ . We do not know the size of these local changes, however. Instead, we merely know the following:  $t1=+$ ;  $t2=+$ ;  $t3=+$ ;  $t4=-$ ; at  $t5=+$ . That is, we know that four local increases and one local decrease in well-being were experienced over  $t1-t5$ . The information provided by happiness does not accurately tell us additional information about the instrumental value of local changes in well-being.

### **5.2.2 Aggregating instances of local changes in well-being over time**

In this section, I will consider whether a simpler version of the summative method discussed in the previous section can provide us with useful information about how well someone is doing in a more global sense. This simpler version does not consist in aggregating the direction and size of local changes in well-being; it merely aggregates the direction of local changes. It consists in adding up all the increases in well-being indicated by positive affective states, and subtracting all the decreases in well-being indicated by negative affective states over time. I will consider whether this calculation can provide us with an (albeit rough) approximation of someone's overall change in well-being over that time.

I think there are three main problems with this kind of aggregation. Firstly, although the simpler version avoids the flaws of the method outlined above, the assumption that it is better to experience more increases in well-being than decreases in well-being is often false. Our overall change in well-being over time is constituted not merely by the amount of changes in well-being that we experience; it is constituted both by the amount *and size* of changes in our well-being. Consider, for example, two people who both experience

three increases in well-being and three decreases in well-being. The simpler version of the summative method would adjudicate that their overall change in well-being over the time in question is equivalent. Yet, this may be widely inaccurate. One person, for example, may have experienced the death of a loved one, in contrast to minor increases in their well-being (making progress in their hobby, say). Whereas the other person may have experienced the birth of their child, in contrast to minor decrease in their well-being (not being that productive at work, say). The overall change in well-being for the two individuals would be considerably different. Yet, this information is lost by the revised method of aggregation.

Secondly, sometimes (perhaps often) the local changes indicated by happiness do not result in more global changes. That is, such changes may not turn out to be instrumentally valuable; they may not “pay off”. Consider, for example, a person who goes to a good university to study law, only to drop out (after completing most of their course) to become an artist. Let us assume that their time at university did not provide them with any skills that contributed towards their career as an artist, nor did it particularly provide them with the insight that being an artist was their calling in life. Further, let us assume that the student received little intrinsic value from completing parts of their course, despite doing so successfully. In such a case, the progress that the person made throughout their law degree did not turn out to have the instrumental value that it was presumed to have had throughout. We would be wrong to approximate that the local changes in well-being experienced by the person throughout their course contributed towards more global changes in their well-being. The local changes did not pay off.

Thus, local changes may not provide us with much information about a person’s overall change in well-being over time. As an analogy, consider two sailboats, both with delta-meters that provide us with information about whether the boat is getting faster or slower (i.e. we have information about the boat’s changes in speed). Now imagine that one boat is in calm waters, whereas the other one is in choppy seas. Both boatmen are highly skilled, and consistently improve the speed of their respective boats. However, the boat in choppy seas has to battle against the occasional large wave, each of which reduces the boat’s speed to almost zero – after each wave, the boat must start incrementally gaining speed again. Let us further suppose that the boat in calm waters also faces the same amount of set-backs, only such set-backs are caused by minor fluctuations in the wind that have a much smaller effect on the boat’s reduction in speed. In such a scenario, the delta-meters

on both boats register the same amount of local improvements in the speed of their respective boats. Yet, one boat is struggling against major (i.e. global) set-backs, while the other boat faces only minor (i.e. local) set-backs. Over time, the boat in calm waters will have reached a respectable speed, whereas the boat in choppy seas may still be at the same speed at which it started. The point is that, from information about local changes in speed alone, we cannot reliably make judgements about overall changes in speed over time. The same goes for changes in well-being.

This is not to say that all, or even most, local changes are devoid of instrumental prudential value. Presumably, local changes often do pay-off, otherwise we would not invest so much of our time and resources into bringing about them about. Returning to the analogy in the previous paragraph, if the boatman in choppy seas looks to see how much progress he has made and discovers that he has made very little, he may well give up. Thus, we can assume that the local changes in well-being indicated by happiness tend to have instrumental value. The problem is that we cannot know how instrumentally valuable such changes are over time purely from looking at such changes in isolation.

The third problem with this method of aggregation is that even if the local changes indicated by happiness do tend to pay off, we still may not be able to approximate a person's *total* change in well-being over time from their local changes in well-being over that time. The reason is simply that we have no way of knowing about all of a person's global changes in well-being from information about their local changes in well-being. We can only know about global changes caused by local changes. For example, we have no way of knowing whether two people with the same amount of local changes in well-being over a certain period of time experienced the same amount of global changes in their well-being. Even if both of their local changes in well-being paid off and resulted in more global changes, one individual may have experienced a global increase or decrease in their well-being that the other individual did not. One individual may have just received news of a loved one being in ill-health, for example. Yet, we cannot derive this kind of important information from their local changes in well-being.<sup>97</sup>

---

<sup>97</sup> Of course, this information will tend to show up in their happiness to a certain extent, but not in proportion to the importance of the change.

Note that the problem is not merely that people can be either lucky or unlucky – that, through no effort of their own, people can experience global changes in their well-being that are not caused by previous local changes. Although this is true, the main problem is that we can deliberately experience many global changes in our well-being that are not caused as a result of previous local changes. Consider, for example, someone who works very hard at their job throughout the year, achieves a small pay rise, and then meets someone they fall in love with at the Christmas party. The latter change may be much more important than the global change in well-being caused by the consistent local changes in well-being achieved at work throughout the year. The point is that the value of global changes in well-being may not tend to reflect the amount of local changes in well-being that brought about such changes. Thus, we cannot discern the importance of a person's global changes in well-being simply from information about local changes.

I think the above three problems prevent us from being able to roughly approximate overall changes in well-being over time by aggregating the local changes in well-being indicated by happiness over that time. However, this is not to say that happiness tells us nothing about people's well-being. In the next section, I will consider two ways in which the local changes in well-being indicated by happiness may tend to have intrinsic prudential value.

### **5.3 The intrinsic value of the local changes in well-being indicated by happiness**

In the previous section, I argued that, although local changes in well-being tend to have instrumental value, we cannot roughly approximate a person's overall change in well-being over time from such local changes. It may seem, therefore, that local changes in well-being do not provide us with much information about well-being. In this and the next section I will argue that, upon further reflection, this is not the case.

In this section, I will discuss two ways in which local changes in well-being have intrinsic prudential value. First, local changes have intrinsic prudential value insofar as they involve worthwhile activities that we find intrinsically motivating, such as learning and experimentation, developing skills and abilities, and forming and sustaining relationships. Second, local changes may have intrinsic prudential value insofar as they involve activities that constitute our active engagement with things of value. These two kinds of local changes concern what I will refer to as a person's "local well-being". I



will show that happiness reliably indicates these two kinds of local changes, and thus reliably indicates changes in our local well-being.

### **5.3.1 Local changes in well-being and intrinsic motivation**

Let us first consider local changes in well-being concerning worthwhile activities that people are intrinsically motivated towards. Examples include things such as learning and experimentation, developing skills and abilities, and forming and sustaining relationships. These kinds of activities are often partaken in not merely as a means to another (more global) end.

The construct of intrinsic motivation was developed to describe the enjoyment people find in pursuing challenges, in pushing themselves to learn and to try new things (Besser-Jones, MS). It is now widely recognized as an important phenomenon (Deci & Moller 2005; Ryan & Deci 2000; Reiss 2004). Richard Ryan and Edward Deci describe the phenomenon of intrinsic motivation as follows:

“Perhaps no single phenomenon reflects the positive potential of human nature as much as intrinsic motivation, the inherent tendency to seek out novelty and challenges, to extend and exercise one’s capacities, to explore, and to learn. Developmentalists acknowledge that from the time of birth, children, in their healthiest states, are active, inquisitive, curious, and playful, even in the absence of specific rewards. The construct of intrinsic motivation describes this natural inclination toward assimilation, mastery, spontaneous interest, and exploration that is so essential to cognitive and social development and that represents a principal source of enjoyment and vitality throughout life.” (Ryan & Deci, 2000: p.70)

Perhaps the most well-known instance of intrinsic motivation is a specific psychological state coined by Mihalyi Csikszentmihalyi as “flow”. As in all instances of intrinsic motivation, people in flow act solely out of a sense of interest and enjoyment, in contrast to acting for the sake of their ends. While they likely have an end in mind, which is part of their reasons to carry out the activity, thoughts about attaining that end are not motivating them to act. Rather, in the case of flow experiences, people become “completely involved in something to the point of forgetting time, fatigue, and everything else but the activity itself” (Csikszentmihalyi et al. 2005: p.599).

Flow experiences occur when individuals engage in complex and challenging activities that they are competent in. A classic example of a flow experience is rock-climbing. The goal is to reach the summit, and while the rock climber

never loses sight of this goal, the goal is not what motivates them in the moment. What motivates the climber is the challenge of exploring the configuration of the rocks and crevices. They enjoy the activity and lose themselves in it, in a state of “optimal experience”. Similarly, athletes describe being “in the zone”, whereby excellence becomes effortless, crowd and competitors disappearing into a blissful, steady absorption in the moment.

There are many other kinds of activities, beyond rock-climbing and other sports that people are intrinsically motivated towards. Indeed, Antonella Delle Fave and Fausto Massimo point out that there are some cultures dominated by such activities:

“Once part of the mainline farming culture of Europe, the Occitan villagers (in the Italian Alps), cut off from the rest of the world by winter snows, have been left behind as a quaint reminder of a way of life that has long since disappeared elsewhere...When asked if they ever felt the intense concentration, clarity of goals, effortless action characteristic of the flow experience, all the older villagers recognized in it the feeling typical of their everyday working lives. That is how they felt, they reported, when they took the cows to the pastures, when they pruned their orchard, when they sat down to carve a piece of furniture out of wood.” (Delle Fave & Massimo, 1992: p.187)

In order for an activity to generate a state of flow, it must be of a certain kind. Perhaps the most essential factor in stimulating flow appears to be the structure of the specific activity in question: Does it provide a sense of discovery? Does it tap into creative feelings? Does it push the person to a higher-level of performance? (Csikszentmihalyi & Csikszentmihalyi 1992) Without these kinds of features, people partaking in the activity will fail to be interested in it or find it enjoyable; instead, they will be preoccupied with attaining the end of the activity.

The same point holds with respect to other forms of intrinsic motivation. While flow experiences are generated by activities that demand a balance of challenges and skills, more general forms of intrinsic motivation are generated by activities that one finds enjoyable, such as drawing and learning. Such activities enable us, “to take interest in novelty, to actively assimilate, and to creatively apply our skills.” (Ryan & Deci, 2000: p.56)

Intrinsically motivated activities, therefore, are specific types of activities. Not all well-being promoting activities can be intrinsically motivating. For example, one may have a personal project that consists in being a well-known

musician. Playing music may be an intrinsically motivated activity, but other related activities, such as promoting one's gigs, transporting musical equipment, having a part-time job, etc., are unlikely to be. The latter activities do not invoke a specialised skill, and do not present agents with challenges towards which to rise. They are not ones that people have a propensity to find interesting and enjoyable.

Although intrinsically motivated activities are limited in scope, we have reason to believe that happiness tends to indicate local changes in well-being concerning such activities. This is because one of the hallmarks of such activities is that they provide us with instant feedback (Csikszentmihalyi & Csikszentmihalyi, 1990). Success in intrinsically motivated activities is immediate, resulting from the expression of interest or the exercise of a skill. In contrast, success in instrumentally motivated activities is often harder to come by and provides a reward that is ultimately uncertain (i.e. whether or not the given task will be instrumental in bringing about a more global change in well-being). Indeed, an established body of empirical evidence suggests that intrinsically motivated activities are one of the key determinants of happiness (Ryan & Deci, 2000; Ryan & Deci, 2001). This confirms the validity of what John Rawls calls the "Aristotelian Principle", whereby, all other things being equal, the exercise of our capacities makes us happy (Rawls 2005).

However, even if happiness tends to indicate local changes in well-being that concern intrinsically motivated activities, is it the case that such changes have intrinsic prudential value? That is, why think that intrinsically motivated activities are intrinsically valuable?

I believe that, unless we have significant reasons to believe otherwise, it is implausible to deny the intrinsic prudential value of intrinsically motivated activities.<sup>98</sup> When people are asked why they carry out such activities, they will answer that such activities are intrinsically rewarding. The phenomenology of intrinsic motivation is that the activity you are partaking in has intrinsic prudential value. Of course, sometimes these experiences may

---

<sup>98</sup> Note that this argument applies to all theories of well-being, regardless of whether they accept or reject the experience or pro-attitude requirement. For instance, according to theories that accept the pro-attitude requirement, intrinsically motivated activities may have intrinsic prudential value because people care about them (and would presumably continue to do so under ideal conditions). In contrast, according to theories that reject the pro-attitude requirement, intrinsically motivated activities may have intrinsic prudential value because they have certain properties that have intrinsic prudential value, such as the exercise of one's skills or capacities, or their fit with a person's nature.

be mistaken. You could be intrinsically motivated to carry out an activity that is harmful or otherwise lacks value. Similarly, one could argue that typical flow activities tend to be largely self-absorbed – obsessed rock climbers may in fact be wasting their life on trivial pursuits. Yet, we would need strong reasons to doubt the value of most intrinsically motivated activities, such as learning and experimentation, developing hobbies and skills, and sustaining bonds of affection and friendship – most intrinsically motivated activities are worthwhile. In the absence of reasons to believe otherwise, I think we should presume that intrinsically motivated activities tend to have intrinsic prudential value.

In addition, certain accounts of well-being may offer explicit reasons for thinking that intrinsically motivated activities have intrinsic prudential value. I will consider two here. First, intrinsically motivated activities may constitute what we might call “prudential virtue”. This idea relies on an analogy with moral value. It is a hallmark of moral virtue that an activity has higher moral value if it is done automatically, or effortlessly, rather than deliberated over. Julia Annas (2008; 2011) describes the motivational state of the mature virtuous person as a specific form of intrinsic motivation: virtuous activity engages a person in distinct challenges, requires mastery of skills, and, for the mature virtuous person, generates spontaneous interest. By analogy, intrinsically motivated activities may be the hallmark of “prudential virtue.”

The second reason that I will consider is related to Dan Haybron’s “self-fulfilment” account of well-being (Haybron 2008a). Haybron argues that well-being is constituted by the fulfilment of our individual nature, and this involves the fulfilment of our “emotional nature.” I contend that the fulfilment of our emotion nature is largely a matter of carrying out the activities that we find intrinsically motivating. As an illustration of this, consider the following example provided by Haybron:

“Consider a young man, Henry, who has a passion for model trains. Henry has the opportunity to go into business with a profitable model railroad shop at which he knows he would be happy. Yet he decides, after careful reflection, to purchase a farm. He has good reasons for the choice: he imagines—correctly—that he would make a fine farmer, and finds the prospect of working the land highly attractive: he sees an elemental appeal to being outdoors and getting his hands dirty, dealing with matters of human survival, and living in close contact with an independent reality. Nothing made-up about it, and none of the degrading political maneuverings and

double-dealings of the professional world. Finally, he desires the extra money it would bring—it is an unusually profitable venture—and he wants to make his parents, whose own background is agricultural, proud. Henry goes to work on the farm and succeeds admirably. His parents are delighted. He is, in short, successfully carrying out a thoughtfully chosen plan of life. (We can assume that he fulfills his other major aims: marrying a woman he loves, having happy and healthy children, etc.)

...The trouble is, Henry is miserable, and has been since taking over the farm. Though he would prefer to be happy, he thinks happiness overrated: a small, and ultimately dispensable, part of the good life. Life isn't supposed to be fun. Besides, he's pursuing a noble calling: happiness is small potatoes by comparison. (He reads a lot of Tolstoy.) But a few of Henry's old friends know better: he chose the wrong line of work. In spite of his ideals, working the land is not an activity that moves, inspires, or fulfills him. It has the opposite effect. It's not that he hates everything he does; again, he thinks it worthwhile, and it pleases him when his crops do well. But the only time he comes alive is when he indulges in his model railroading hobby. That's what turns him on." (Haybron, 2008a: p.25)

Haybron emphasizes that Henry's actions are not ill-informed in any way – he values the worthwhile lifelong pursuit of working a farm to a much greater extent than he values model railroading. In addition, Henry knows that he would be far happier if we went into the model train business – it's just that he does not highly value his own happiness. Nonetheless, Haybron argues that Henry is not better off in this respect, as he has assigned too little importance to the things that fulfil his emotional nature (i.e. the things that make him happy). Now, we needn't entirely agree with Haybron on this point. That is, it may or may not be the case that the fulfilment of our emotional nature partly constitutes our well-being. Yet, we can agree with Haybron that our emotional nature makes a difference to the things that are good for us. This adds further support to the claim that the activities we are intrinsically motivated by have prudential value.

Lastly, it is worth noting that the local changes in well-being that we experience from our intimate relationships and friendships tend to have intrinsic, as well as instrumental, prudential value. People talk of "maintaining" their relationships through staying in contact, spending time together, confiding in each other, and so on. But, such activities are rarely done for entirely instrument reasons (i.e. to maintain the status of the

relationship). Spending time with friends and loved ones is good in itself, as well as being a good way of maintaining or developing those relationships. It is perhaps for this reason that “relationships” are often cited as one of the main determinants of happiness (Diener & Biswas-Diener, 2008); the time taken maintaining one’s relationships does not merely make one’s life instrumentally better, but also makes an intrinsic difference to the quality of one’s life. These local changes in well-being seem to have a considerable impact on our well-being in themselves.

### **5.3.2 Local changes in well-being and engaging with things of value**

In the previous section, I argued that local changes in well-being have intrinsic prudential value insofar as they involve activities that we find intrinsically motivating. In this section, I want to argue that these and another kind of local changes in well-being may have intrinsic prudential value in an additional respect: such changes constitute our *active engagement* with things of value.

The kinds of local changes I have in mind are those that concern activities we are either *intrinsically* or *autonomously* motivated towards. These changes may have intrinsic prudential value insofar as they constitute our active engagement with things of value. I have already considered activities that we are intrinsically motivated towards, so I will focus on autonomously motivated activities here. Autonomous motivation is a form of “extrinsic” motivation. Whereas intrinsic motivation appeals to interests, extrinsic motivation appeals to *values* (Deci & Moller, 2005: p.591). As we discussed in the previous section, whether or not an activity is interesting on its own (and so is a source of intrinsic motivation) has largely to do with the structure of the activity. If the activity is one that taps into our natural propensities, then it has the potential to prompt a state of intrinsic motivation in the agent. But if it doesn’t have this structure, it is not at all likely to prompt a state of intrinsic motivation; our engagement in the activity will tend to be a form extrinsic motivation.

There are two broad forms of extrinsic motivation, namely “controlled” and “autonomous” forms. A controlled form of extrinsic motivation is one in which people perceive their goals as being *externally* imposed on them, whereas an autonomous form of extrinsic motivation is one in which people perceive their goals as being their own (Ryan & Deci, 2000). You are control motivated insofar as you partake in an activity because you “have” to, rather than because you “want” to; you perceive yourself as being controlled by external features – the threat of punishment or promise of reward, for instance.

Another way of putting this point is that you experience a kind of detachment from your goals: you work to pursue goals that are not of your own making.

The promotion of well-being is like this for many people. We often pursue goals that have been externally imposed on us, be it through society, educators, or religious institutions. We know that we have to “be successful” but do not appreciate *why* such success matter; we know that people will disapprove of us if we are not successful, or that we will feel bad. But these are not autonomous motivations – they are rewards or punishments that are external to us.

An autonomous form of extrinsic motivation is one in which agents *identify* with a particular goal and thereby see it as their own.<sup>99</sup> They may have also *integrated* it into their sense of self through bringing it into harmony with their other goals (Sheldon 2004). The result is that autonomous motivation simply feels different to controlled motivation. When we are autonomously motivated we feel like the “origin” of one’s actions rather than a “pawn” (Deci & Moller, 2005). For example, a child who does their homework because they want to learn, in contrast to a child does their homework because they want to avoid punishment, will feel as if their actions come from within, as opposed to being imposed on them from the outside. Another way of putting this point is that the former child is moved by her values, whereas the latter child is not.

Both autonomously and intrinsically motivated activities may have intrinsic prudential value in that they partly constitute our active engagement with things of value. Such engagement is valuable because we cannot entirely appreciate the things that are good for us from a reflective standpoint. We come to value particular things as a result of our active engagement with them. In contrast to controlled motivation, autonomous and intrinsic forms of motivation are the kinds of motivation that enable us to appreciate the value of the things that we engage with in our lives. As Valerie Tiberius notes, “it is (at least in part) by being a friend, daughter, sibling or parent that we discover what is valuable about these relationships. It is by absorbing ourselves in a hobby or career that we experience the value of accomplishment. It is by losing ourselves in the moment that we experience the value of pleasure, peace of mind, or fun.” (Tiberius, 2005: p.171).

---

<sup>99</sup> This state has much in common with accounts of autonomy that emphasise the importance of identifying with one’s goals, desires, etc. (Dworkin 1988; Frankfurt 1988; Watson 1987)

Autonomously motivated activities, like intrinsically motivated activities, tend to be correlated with happiness (Ryan & Deci, 2000; Ryan & Deci, 2001). Whereas intrinsically motivated activities tend to directly cause happiness, autonomously motivated activities may do so more indirectly. Autonomous forms of motivation tend to result in outcomes that cause happiness, such as a higher degree of motivation, higher performance rates, enhanced coping skill and reduced anxiety (Ryan & Deci, 2000). Thus, the local changes in well-being concerning autonomously motivated activities will tend to be indicated by happiness (albeit, perhaps to a lesser extent than local changes in well-being concerning intrinsically motivated activities).

The upshot is that the local changes in well-being indicated by happiness often have intrinsic prudential value. These kinds of changes concern what I will refer to as a person's "local well-being". We have reason to believe that happiness reliably indicates changes in our local well-being.

The question I will be concerned with in the next section is to what extent does local well-being matter? To what extent can we measure overall well-being through measuring local well-being? In the next section, I will discuss what we might call the "global challenge". It is the view that we can only successfully measure well-being through measuring global well-being (as indicated by life satisfaction), rather than local well-being (as indicated by happiness). Such a view is intuitively appealing and favoured by a number of theorists of well-being. Nonetheless, I will argue that it is false.

## 5.4 The global challenge

In the previous section, I suggested that we can successfully measure changes in well-being through measuring happiness because happiness tends to indicate changes in *local well-being*. That is, happiness tends to indicate local changes in well-being that have intrinsic prudential value.

We can contrast a person's *local well-being* with their *global well-being*. A person's global well-being consists in how well they are doing on a general level – in achieving such life goals as material security, personal projects, long-term ambitions, achieving the respect of others, making meaningful accomplishments, etc. These things certainly seem to be important to our well-being – so much so that we may question the extent to which we can successfully measure well-being through measuring local well-being. How important is local well-being in comparison to global well-being?



I will call this the *global challenge*. The global challenge may consist in two different claims. The first claim is that global well-being is the *only* thing that matters for well-being. According to this claim, changes in local well-being do not contribute towards overall well-being. The second claim is that global well-being matters for well-being in a *distinct way*. According to this claim, both changes in global and local well-being may contribute towards overall well-being, but changes in global well-being may do so in an additional respect.

I will assume that the first claim is false, and therefore will not discuss it in this section. In the previous section, I outlined two ways in which local well-being seems to have intrinsic value.<sup>100</sup> The second claim, however, seems more plausible. Indeed, I think something like it is currently endorsed by a number of well-being theorists.<sup>101</sup> Nonetheless, I will argue that it is also false. I will argue that both global and local well-being contribute towards overall well-being for the same reasons, even if not to the same extent. I will suggest that the extent to which both kinds of well-being contribute towards a person's overall well-being is at least partly determined by factors concerning the person's lifespan, personality and cultural environment.

What are the arguments in favour of the global challenge? I think that there is one way in particular in which global well-being may tend to have a distinct kind of intrinsic value. Changes in global well-being may tend to be *meaningful* changes – they tend to contribute towards the meaningfulness of our lives. Those who find point and meaning in their life are fully invested in their life. As Joseph Raz puts it, “they address themselves to various projects and relationships with energy and commitment.” (Raz, 2004: p.280) As mentioned in chapter one, meaningfulness and well-being are clearly interrelated. The aspects of our lives that contribute towards our well-being tend to be those which (at least would) make us feel that our life is not meaningless. Changes in global well-being may tend to be distinct from local changes in this respect – such changes give our lives meaning.

---

<sup>100</sup> A straightforward way of seeing the falsity of this claim is to compare two lives equal with respect to global well-being, but one life containing a higher level of local well-being. It seems implausible to suggest that the life with a higher level of local well-being does not thereby have a higher level of overall well-being.

<sup>101</sup> See, for instance, Velleman (1991), Raz (2004) and Kauppinen (2012).

I will consider two ways in which changes in global well-being may tend to be uniquely meaningful. Global changes may be meaningful insofar as they involve activities that (a) form a meaningful narrative, or (b) we deeply care about. With regards to (a), I will argue that our engagement in activities that form a meaningful narrative enables us to appreciate the value of both the long-term achievement and actively engaging in things of value. However, this does not mean that either the long-term achievement or the process of engagement is more valuable than short-term achievements or a non-narrative process of engagement. With regards to (b), I will argue that, although we do deeply care about global changes in well-being, this is not unique to global changes – we can come to deeply care about local changes in well-being. The upshot is that changes in both local and global well-being are valuable in the same ways.

#### **5.4.1 Global changes in well-being that are part of a good narrative**

In this section, I will consider whether global changes in well-being tend to have a distinct kind of intrinsic value as a result of being part of a good narrative. For example, overcoming adversity to accomplish one's long-term ambition forms a better narrative than the same accomplishment occurring through luck. There is no doubt that such narratives have aesthetic value. But do they make our lives go better for us? A number of well-being theorists have recently answered this question in the affirmative (Velleman, 1991; Kauppinen, 2012; Portmore 2007; Raibley 2012; Bramble, forthcoming). In this section, I will argue that such theorists are mistaken. Good narratives may enable us to *appreciate* the value of certain long-term achievements and actively engaging in things of value. However, they do not, in themselves, contribute towards well-being.

Before considering the value of narratives, it is worth noting that narratives tend to significantly influence judgements of life satisfaction. For instance, in one study Diener et al. (2001) provided subjects with vignettes of either very positive or very negative lives. Some subjects received vignettes of lives that ended abruptly. Other subjects received vignettes of lives that ended with an extended five-year period, that was still either positive or negative, but only mildly so. Diener et al. found that a positive life was perceived as more satisfying when it ended abruptly on a high note than when it extended for an additional five mildly positive years. In contrast, a terrible life was viewed as less satisfying when the five mildly negative years were added. The authors labelled this the "James Dean Effect", with reference to the fact that we tend

to think people with short intensely positive lives have satisfying lives so long as they end “on a high.”

Further psychological research shows that narratives determine people’s life satisfaction across a wide range of contexts and timescales. For instance, in a seminal study on retrospective evaluations of pain, Daniel Kahneman et al. (1993) found that, under certain circumstances, people are satisfied with more pain to less. Kahneman and his colleagues found that, when subjects were asked how satisfied they were with certain episodes of pain, subjects reported the painful experience of submerging their hand into 14°C ice water for 60 seconds, plus an additional 30 seconds of 15°C ice water, was more satisfying than a shorter trial of 60 seconds of 14°C ice water alone. The addition of the more moderate, though still aversive pain, led to the (seemingly irrational<sup>102</sup>) preference of objectively more pain over less.

This pattern has been shown in retrospective evaluations about a number of disagreeable experiences, including exposure to unpleasantly loud noises, aversive film clips, pressure from a vice, and painful medical procedures (Finn 2010). In each of these cases, the longer episode that ended with less discomfort was judged to be more satisfying than the shorter episode.<sup>103</sup>

The lesson to learn from these studies is that judgements of life satisfaction are significantly influenced by global changes that mark the end of a difficult activity, result in previous local changes “paying-off”, or that end the activity on a “high note.” Such narratives tend to make us feel that our lives are more meaningful, even if they consist in episodes of suffering. Roy Baumeister et al. (2012) summarize these trends as follows:

“Our findings depict the unhappy but meaningful life as seriously involved in difficult undertakings. It was marked by ample worry, stress, argument, and anxiety. People with such lives spend much time thinking about past and future: They expect to do a

---

<sup>102</sup> For commentary on the irrationality of such preferences, see Broome (1996). For an alternative view, see Beardman (2000).

<sup>103</sup> This pattern has also been shown in contexts other than pain and discomfort. For instance, Bridgid Finn (2010) explored whether extending an effortful study experience with a less effortful interval would be preferred to a shorter, unextended interval. She found that the students’ judgements of satisfaction with their study experience were influenced more by the structure of the learning episode, rather than their performance on the test. Like other aversive experiences, effortful study that ended with somewhat easier material was judged to be less difficult, to cause less discomfort, and to be easier to cope with than a study experience that included equally challenging material but that did not include the more moderate material at the end.

lot of deep thinking, they imagine future events, and they reflect on past struggles and challenges. They perceive themselves as having had more unpleasant experiences than others, and in fact 3% of having a meaningful life was due to having had bad things happen to you.” (Baumeister et al, 2012: p. 15)

In the remainder of this section, I will consider whether the “meaningful life”, as described in this quote, has a distinct kind of intrinsic value in contrast to the “happy life.” Is it the case that good narratives have intrinsic prudential value in themselves?<sup>104</sup>

To illustrate the view that they do, consider David Velleman’s description of two possible lives:

“One life begins in the depths but takes an upward trend: a childhood of deprivation, a troubled youth, struggles and setbacks in early adulthood, followed finally by success and satisfaction in middle age and a peaceful retirement. Another life begins at the heights but slides downhill: a blissful childhood and youth, precocious triumphs and rewards in early adulthood, followed by a midlife strewn with disasters that lead to misery in old age. Surely, we can imagine two such lives containing equal sums of momentary well-being. Your retirement is as blessed in one life as your childhood is in the other; your nonage is as blighted in one life as your dotage is in the other.”

(Velleman, 1993: p.331)

Intuitively, it seems that the first life, where things get progressively better, is better for you than the second, where things get progressively worse, and it seems this way even if we are to imagine that both lives contain the same achievements, occurring only at different stages in each life. The first is seemingly a better life, not just in the sense that it makes for a better life story, but also in the sense that it has a distinct kind of intrinsic value. That such lives are better for us has recently come to be known as the “Shape-of-a-Life Phenomenon.” (Feldman 2004: ch.6)

On Velleman’s view, the reason a benefit that comes late in life can have a more profound effect on the value of one’s life is that benefits experienced late in life can *redeem* misfortunes incurred early in life. So a life that gets

---

<sup>104</sup> By this, I do not mean to claim that the narratives themselves have intrinsic prudential value, beyond the events and activities that make up such narratives. Rather, I am interested in the claim that the events and activities that constitute good narratives have a distinct kind of intrinsic prudential value.

progressively better is, in certain cases, better for us than a life that gets progressively worse – it is only in the case of the former that one’s earlier misfortunes are redeemed. In the life that gets progressively better, the earlier trials and tribulations lead to the later successes and thereby redeem themselves. But in a life where the successes precede the misfortunes, the misfortunes could not have served as the foundation for those successes, and so will have been “suffered for naught” (Portmore, 2007).

This all makes sense, but it is not obvious that these kinds of intuitions are best explained by the fact that such narratives have intrinsic prudential value. In the remainder of this section, I will argue that it is more plausible to presume that meaningful prudential events have instrumental prudential value, rather than a distinct kind of intrinsic value.

Firstly, meaningful prudential events tend to result in causally related changes in well-being. Negative events tend to result in further negative changes. For example, the loss of something that matters to me (such as a job I valued, my reputation, or my house keys) tends to cause further bad things to happen. Conversely, positive events tend to result in further positive changes.

Secondly, because of this fact, people may tend to care about meaningful events. We may be unhappy as a result of our hard work not paying-off, not just because we failed to achieve the global change we valued, but also because the meaningful event is likely to cause worse things to happen in the future. Thus, the meaningful event may be bad in three respects: (a) because of the value of the event itself (e.g. losing one’s job); (b) because of the instrumental value of the meaningful event (e.g. being less likely to find another job); and (c) because we care about the meaningful event (e.g. we are unhappy about its repercussions).

Thirdly, meaningful events *signal* the value of several previous local and global changes in well-being (Dorsey, MS). For example, “being a world class tennis player” consists of many discrete events in a person’s life, such as taking up tennis, practicing, entering into tournaments, winning, and so forth. The meaningful event of “being a world class tennis player” narratively unifies these discrete events – it signifies that each event has prudential value. This does not mean that “being a world class tennis player” has value beyond the value of each discrete event that it is composed of, nor does it mean that each of the discrete events has greater value as a result of being part of the more meaningful event. Rather, “being a world class tennis player” has merely

signatory value. It highlights that all the events that make up “being a world class tennis player” have been successfully achieved.

Fourthly, our awareness of narrative relations tends to enable us to leave better lives. As Connie Rosati (MS) puts it, “narratives can affect a person’s good by helping her, as she reflects on events in her life, and as she takes up a view of herself and her life, to see herself and her life in a way that supports her sense of her own worth, that helps to secure her sense of who she is and a sense of direction, and that motivates her to move forward.” Rosati further states that upward trajectories tend to be good for us in that they come out of being, “the controlling authority over ourselves and our lives, of being able to make sense of our lives and to represent them to ourselves as a product, ultimately, of our own autonomous efforts.” Rosati’s point is that when we have experienced success in the face of failure, we will have a sense of ourselves as the controlling authority over our lives and ourselves. This sense of authorship enables us to remain motivated when we fall short of achieving our aims.

Changes in global well-being concerning long-term projects may tend to be especially (instrumentally) valuable in this respect. As Dale Dorsey notes, “long-term projects that help to narratively unify our actions and decisions provide a meaningful structure to our lives as a whole, not just from the outside (though that itself is significant) but also from the inside.” (Dorsey 2011a: p.178) The narrative structure provided by the meaningful pursuit of long-term projects enables us to keep striving in the event of set-backs.

This suggests that narratives enable us to *appreciate* the value of certain long-term projects and actively engaging in the pursuit of such projects. The difference between someone who climbs up the entirety of Mt. Everest and someone who climbs up Mt. Everest from base camp is that the former individual is in a better position to see the value of getting to the top of Mt. Everest more than the latter individual can. Narrative emotions such as pride, joy, hope, self-esteem and elevation – emotions that we can refer to broadly as “feelings of fulfilment” (Kauppinen, 2012) – enable us to grasp the value of such an endeavour.

In this way, I believe that narrative relations are similar to the relations we have with other individuals. I follow Simon Keller (2013) in thinking that our personal relations with others enable us to grasp the intrinsic value of the individuals we have such relationships with. I can see that my friend has intrinsic value in various ways (e.g. the way they listen to me, their sense of

humour, their interest in modern art, etc.). This is partly due to the times I have spent with them over the years, the caring role they have played in my life, the struggles I have witnessed them go through, and so on. Likewise, it may be that people can see that the achievement of their long-term goals has intrinsic value in various ways due to the trials and tribulations that they went through in the process of pursuing them.

This analogy helps us understand the intrinsic value of global changes in well-being that are part of good narratives. Our personal relations enable us to better appreciate the intrinsic value of the individuals who we have relationships with. However, even if we see the value of such individuals more than we see the value of others, this does not make such individuals more valuable. People have equal intrinsic value. I think we can say a similar thing about narratives. Narratives enable us to appreciate the intrinsic value of global changes in well-being that are part of good narratives. However, I do not think that such changes have a distinct kind of intrinsic value in this respect. The kind of local changes in well-being discussed above may provide meaning to our lives, even if they do not form good narratives in the process.

Indeed, Galen Strawson (2004) has influentially argued that not all people may come to appreciate the value of long-term goals and projects as a result of the narratives that they form. He claims that not all people care about their life narrative, and that a person's life narrative is not necessarily valuable. As Strawson writes: "The aspiration to explicit Narrative self-articulation is natural for some—for some, perhaps, it may even be helpful—but in others it is highly unnatural and ruinous." (Strawson, 2004: p.447). He continues: "People can develop and deepen in valuable ways without any sort of explicit, specifically Narrative reflection, just as musicians can improve by practice sessions without recalling those sessions. The business of living well is, for many, a completely non-Narrative project." (Strawson, 2004: p.448). Strawson's argument suggests that narratives may not have instrumental value unless one cares about their life story. Indeed, it may be that, for those who don't care about the narrative of their life, an awareness of narrative relations may tend to make their lives worse.

#### **5.4.2 Global changes in well-being that we deeply care about**

If narratives do not have a distinct kind of intrinsic prudential value, perhaps the objects of global changes in well-being themselves tend to have a distinct kind of intrinsic value. Achieving such life goals as material security, personal

projects, long-term ambitions, achieving the respect or companionship of others, etc., may seem to have a distinct kind of intrinsic value in contrast to achieving momentary objectives and concerns such as developing skills or hobbies, or maintaining relationships.

One way of making this distinction is that global changes in well-being tend to concern things that we *deeply care about*. It is hard to image a meaningful life that does not involve some long-term relationships and projects that we deeply cared about. This idea relates to what Bernard Williams called “personal projects” (Williams, 1993). He described projects as “commitments” with which we are more deeply involved and identified, and which help to constitute our individual character. Williams further claimed that personal projects are what give meaning to our lives.

I believe, however, that this idea goes too far. In that, it is not only the things we deeply care that provide meaning in our lives. As Cheshire Calhoun points out: “Caring is a scalar phenomenon.” (Calhoun 2009: p.632) Once we account for the full spectrum of cares, from deep, moderate to minor cares, it ceases to be obvious that leading a meaningful life depends specifically on the achievement of long-term relationships and projects.

We can agree that the pursuit of what we care about makes our lives meaningful without maintaining that we must “deeply” care about those things. A person’s life might allow plenty of opportunity to scuba dive, do volunteer work, spend time with friends, and take sign-language classes – these are all things they find satisfying and personally expressive, without having to care about them deeply. Looking back at their life, they might well say that they led a meaningful life. Indeed, as Calhoun correctly notes, “the meaningfulness of a life is a function of both how much one cares about particular relationships, activities, and the like as well as how much time one has the good fortune to spend with those things.” (Calhoun, 2009: p.632) A person who has many varied and easily pursued objects of lesser care may end their life having spent more of its days and hours in meaningful activities than their more single-minded, passionate counterpart.

The claim that only global achievements, such as long-term relationships, careers, etc., make life meaningful ignores the meaning provided by more local activities. This is perhaps not surprising insofar as achieving long-term goals, value and projects typically require us to ignore local changes in well-being that have intrinsic value. Consider the average life of a philosopher. While participating in academic philosophy may make their life meaningful in



the long-term, doing so requires that they sacrifice the pursuit of a large number of potentially meaningful local activities in favour of spending their time doing things that merely have instrumental value, such as tracking down footnotes, participating in departmental meetings, and so on.

The upshot of this discussion and the discussion in the previous section is as follows: there are at least two different strategies for putting meaning in life. One may opt for a temporally global strategy, aiming to make one's life about something that is objectively valuable, and to do so by focusing on global changes in well-being. Or one may opt for a temporally local strategy, aiming to pack one's hours and days with objectively meaningful pursuits and interactions, and to do so by not focusing on the sort of global changes in well-being that promise to reduce such opportunities (Calhoun, 2009).

Why is it that people tend to focus so much more on global changes in well-being if such changes do not tend to have a distinct kind of intrinsic value? I suspect the answer is largely pragmatic; that is, long-term achievements tend to have instrumental prudential value. The pursuit of a career, obtaining a degree, being in a committed relationship – these are all long-term activities that tend to provide substantial pay-offs: financial security, recognition, emotional support, etc. Thus, if we want to measure people's opportunities for future well-being, it may make sense to focus on measuring their recent global changes in well-being. If someone's well-being decreases in a global sense (e.g. they lose their job, break up with their long-term partner, etc.) it is likely that such changes will continue to have a negative impact in the future. In contrast, if someone's well-being increases in a local sense (e.g. they spend time learning something of interest or developing a particular skill) we cannot infer, to the same extent, how well they are likely to be doing in the long-term.<sup>105</sup>

## **5.5 The prudential value of both local and global changes in well-being**

In the previous two main sections, I argued that both local and global changes in well-being tend to have intrinsic prudential value; neither kind of changes in well-being have a distinct kind of intrinsic value. The result is that overall well-

---

<sup>105</sup> Though see Appendix A for the merit of measuring people's opportunities for future well-being through measuring their recent levels of happiness.

being is, for the same reasons, constituted by both local and global well-being. In this section, I want to show that the relative values of local and global well-being may vary throughout the life-course, across individuals and across cultures.

Consider, first, how the value of local well-being varies throughout one's life, beginning in childhood. When we think about what is good for children, we partly think of global changes in well-being that have instrumental value, such as a good education. In addition, however, we tend to give considerable weight to their local well-being. Anca Gheaus illustrates this tendency with the following two vignettes:

“One

It has been snowing the entire weekend, and the public transportation stopped working. School closed, and children are happily playing in the snow for hours. Adults struggle to get to work and to carry on with business as usual. This is an urban image that I remember from my childhood, as I assume many other readers will. It is a world of Carefree Childhood and Serious Adulthood, where children and adults lead partially separate lives and enjoy partly different goods.

Two

It has been snowing the entire weekend, and the public transportation stopped working. Schools closed, and children, who have to stay at home, received additional homework by email. They concentrate on independent study and try to make sure they don't fall behind with it. Adults struggle to get to work and to carry on with business as usual ... This is an urban image that one can occasionally see these days. It is a world of Serious Childhood and Serious Adulthood, one in which children and adults lead more similar lives than in world One because children do not have access to some goods which are plausibly intrinsic to the good of childhood.” (Gheaus forthcoming)

These two vignettes are designed to pump our intuitions about what is good for children. I take it that most people will think that world One is better than world Two, even if it is the case that the children in world Two end up being slightly better off as adults than the children in world One. This is presumably because a good childhood includes significant amounts of free time, unstructured play, opportunities for joyful and experimental social interaction, and a sense of being carefree. These things have intrinsic prudential value, in addition to the instrumental value that they have – they are things we would

regret losing if we replaced them with alternative things of the same instrumental value.

People's attitudes tend to be different, however, with regards to adulthood – one's "prime-of-life" (Slote, 1983). As mentioned in chapter four, studies show that our values tend to change over the life-course, focusing on more extrinsic goals in our middle-age (such as career success, financial security, etc.) and more intrinsic goals towards retirement (friendships, family, hobbies) (Reed & Carstensen 2012). Happiness has been shown to take a U-shaped curve over the life-course, whereby we are happiest in our childhood and at old-age, and unhappiest in our middle-age (Blanchflower & Oswald 2008). This is presumably because we focus less on our local well-being in our middle-age. Adulthood is the time of life in which one can focus on long-term achievements that help structure one's life and come to partly form one's identity: either in the form of a particular role – a parent, lawyer, doctor – or in the form of a personal project – someone who is helping to save the rain forest, for example (Betzler 2013). As argued in the previous section, these kinds of achievements may have considerable intrinsic prudential value, even if they are not uniquely valuable in this respect.

It is important to note that not all people's values will tend to change over the life-course in this way. Some individuals may continue to prioritize their local well-being throughout their lives (e.g. the keen learner of new skills, the obsessed rock climber, etc.) whereas others may abandon their local well-being altogether (e.g. the ambitious businessman, the tortured artist, etc.). In other cases, extreme life events may shift one's temporal perspective. It is not uncommon, for example, to hear of cancer patients re-orienting their lives to focus on their local well-being as their prospects of achieving lifelong projects becomes uncertain. Moreover, such shifts in perspective are often not reported as a compromise; rather, they are often seen as providing people with insights into what really matters in life.

Indeed, the extent to which people value their local and global well-being may be largely determined by their particular psychological disposition. For instance, Calhoun notes that people have different spatial dispositions and that such preferences may extend to a temporal perspective:

"Some like open horizons, not knowing what comes next, opportunities for explorations, and novelty. Others prefer more closed horizons, knowing where they are in space and time and finding what comes next familiar and unsurprising. If this is so, one might expect to find some more attracted to

provisional planning and others to commitment. In short, the attraction to commitment may reflect only one style of managing the geography of one's future—a style that involves taking up permanent residences in time and making the future one's home. For creatures who share with their animal kin a sense of comfort in returning to the familiar and a disposition to adopt habitual routines that reduce surprises, one might expect commitment to be widely attractive. But as creatures who also share with their animal kin a vulnerability to boredom and a curiosity about the new, one might also expect that attraction to have its limits." (Calhoun, 2009: p.641)

I believe Calhoun is right to point out the diversity with which people tend to respectively value their local and global well-being. In addition to personality differences, the relative values of local and global well-being may vary considerably depending on one's cultural environment. Recall, for instance, Delle Fave and Massimo's example above of the Occitan villagers who tend to live their everyday working lives in a state of flow (Delle Fave & Massimo, 1992). Such a culture stands in stark contrast with the busy, stressful, achievement-focused lives lead by those in more "Western" cultures. This is presumably because long-term opportunities are available in the latter cultural environment, but may not be in the former.<sup>106</sup>

Now, these facts – about how people's relative preferences for local and global well-being vary according to the life-course, personality and culture – do not necessarily have straightforward implications for the relative values of local and global well-being. It may not be the case, for example, that for people with a disposition to live their life from a local perspective, their local well-being tends to have a greater amount of intrinsic prudential value. The extent to which this is true will depend on the correct account of well-being. A (non-idealized) subjective theory of well-being will maintain that local well-being matters more for those who care more about their local well-being. In contrast, an objective theory of well-being may maintain that those who care more about their local well-being fail to appreciate the value of their global well-being (e.g., the value of various long-term achievements).

---

<sup>106</sup> Indeed, Delle Fave and Massimo note how, as the remote Occitan villages become more accessible to mainland Europe, the younger generation no longer enjoy the traditional forms of life: "Their concentration is disrupted by goals and desires that come from the culture of the plains...For them work is drudgery to be endured only for the money it brings...and since work brings more money in factories, most younger Occitans are settling down to industrial jobs far away from their native valley." (Delle Fave & Massimo, 1992: p.188)

What I think we can say is that both local and global well-being matters, and that this is the case according to all plausible theories of well-being.<sup>107</sup> Insofar as happiness tends to indicate changes in local well-being, therefore, we can successfully measure changes in an important aspect of well-being through measuring happiness. The extent to which we can measure well-being through measuring happiness will depend on the relative values of local and global well-being. In this section, I have suggested that the relative values of local and global well-being may depend on facts about a person's lifespan, personality and cultural environment. The precise way in which this is the case is an exciting topic for future philosophical and empirical research.

## Conclusion

I have argued that we can successfully measure changes in an important aspect of well-being through measuring happiness. In the first section, I showed that happiness tends to indicate local changes in well-being, in contrast to life satisfaction, which tends to indicate global changes in well-being. In the second section, I considered the instrumental value of the local

---

<sup>107</sup> Although I am unable to consider this further here, I believe that a stronger claim may also be plausible, namely that achieving well-being is largely a matter of achieving the appropriate balance between one's local and global well-being. This fits with the idea (discussed in chapter one) that the concept of well-being refers to a kind of life in which one succeeds in all or most of one's important goals relative to one's available time and resources. One may be able to achieve a good life by succeeding in one or more of one's important goals (such as Mother Teresa achieving a good life through succeeding at helping the poor). However, doing so may necessarily result in having to sacrifice many of one's other important goals (such as developing intimate relationships, skilled hobbies, staying healthy, etc.).

It seems reasonable to suggest that, in modern industrial societies, people tend to succeed in achieving a few of their important global goals at the expense of many of their important local goals. Such people may be better off if they shifted their priorities more towards their local well-being. The outcome of such a shift is likely to result in a society of people more like those in community A rather than those in community B as described at the beginning of Haybron, 2008b. Haybron states as follows: "A typical member of A, on a typical day, is in more or less the following condition: at ease, untroubled, slow to anger, quick to laugh, fulfilled, in an expansive and self-assured mood, curious and attentive, alert and in good spirits, and fully at home in her body, with a relaxed, confident posture. A denizen of B, by contrast, is liable to be: stressed, anxious and insecure, spiritually deflated, pinched, and compressed. The differences, let us suppose, owe mainly to differences in the prevailing ways of life in these communities." (Haybron, 2008b: p.3) I suggest that the differences between community A and community B is that the former is focused primarily on the pursuit of local well-being, whereas the latter is primarily focused on the pursuit of global well-being. A community that is good for its citizens may be one in which neither aspect of well-being is pursued at the expense of the other.

changes in well-being indicated by happiness. I argued that we cannot accurately approximate a person's overall change in well-being over time by aggregating their local changes in well-being over that time. In the third section, however, I argued that happiness tends to indicate the local changes in well-being that have intrinsic value. Changes concerning both intrinsically and autonomously motivated activities tend to have intrinsic prudential value.

In the remainder of the chapter, I considered the relative value of local and global well-being. In the fourth section, I argued that global changes in well-being do not tend to have a distinct kind of intrinsic value, neither as a result of (a) being part of good narratives nor (b) concerning things we deeply care about. I argued that these aspects of global changes in well-being tend to enable us to appreciate the value of such changes, rather than providing them with a distinct kind of intrinsic value. Lastly, in the fifth section, I suggested that the relative values of people's local and global well-being may tend to vary across their lifespan, personality and cultural environment.

The upshot is that we can successfully measure changes in an important aspect of well-being through measuring happiness. We cannot, however, successfully measure changes in overall well-being through measuring either happiness or life satisfaction alone. In order to measure a person's overall change in well-being over time, we must measure their changes in both local well-being (indicated by happiness) and global well-being (indicated by life satisfaction) over that time. The fact that happiness and life satisfaction are essential indicators of well-being in this respect is, in many ways, the central conclusion of my investigation.

In the two thesis appendices, I will explore two further ways in which we may be able to measure well-being through measuring happiness. The first is an extension of the Indicator View. It is the view that happiness over time indicates local changes in well-being over that time *and* well-being in the near-future. The second way concerns whether we can infer *levels* of well-being from the local changes in well-being indicated by happiness. I will suggest that we may be able to infer a person's level of well-being from the level of variation in their happiness over time. These two ways add to the extent to which we can successfully measure well-being through measuring happiness.

## Conclusion

In this thesis, I have aimed to clarify the role of happiness research in the measurement of well-being. Happiness is notoriously elusive, yet has recently received a considerable amount of attention in the study of well-being, and from the media and policymakers. Theorists are divided in their interpretations of happiness findings. For example, some theorists interpret the (contentious) finding that happiness does not correlate with income in the long-term as an indication that we should radically rethink the use of GDP in developing and assessing policy. Other theorists interpret the same findings as an indication that there is more to life than happiness; indeed, that happiness may be a relatively trivial affair. In this thesis, I have aimed to provide an account of the relationship between happiness (conceived of as a person's affective state) and an understanding of well-being that is widely acceptable (that is, robust to most plausible substantive theories of well-being). It has the potential, therefore, to unite divided opinion over what happiness research can tell us about well-being.

I have argued that we should treat happiness as a defeasible indicator of local changes in well-being. I called this the Indicator View. Unless we have reason to believe otherwise, happiness tells us about how well we are doing, from moment-to-moment, in our short-term objectives, goals and concerns. In chapter five, I argued that how well our lives are going on this local level intrinsically matters for well-being. It is not all that matters – how well we are doing on a global level (in our long-term projects and values) also intrinsically matters for well-being. The relative values of our local and global well-being may vary across the lifespan, across individuals and across cultures. Yet, at any age, for any individual, in any culture, we can assume that how well we are doing on both levels matter. We can, therefore, successfully measure an important aspect of well-being through measuring happiness.

The Indicator View rests on a theory-neutral account of well-being, outlined in chapter one and developed in chapter two. According to this understanding of well-being, happiness may correlate with well-being either because it is constitutively or causally related to well-being. The extent to which happiness correlates with well-being is determined by its causal connections to well-being ingredients (objects of platitudes about well-being that are causally

connected to other objects of well-being platitudes). The theory-neutral account provided us with a widely acceptable understanding of well-being, from which we could investigate the correlation between happiness and well-being.

In chapters three, four and five, I investigated the extent to which the function of happiness consists in the reliable detection of well-being ingredients. In chapter three, I argued that the function of happiness is to inform and guide action; this partly consists in detecting well-being ingredients. In chapter four, I refined this view. I argued that the function of happiness also partly consists in detecting unexpected changes in well-being ingredients. Lastly, in chapter five, I added that the function of happiness partly consists in detecting local changes in well-being ingredients. We can infer from the function of happiness that happiness will tend to be reliably caused by local changes in well-being ingredients, and thereby tend to correlate with local changes in well-being.

The theory-neutral account also enabled us to identify contexts in which we should not treat happiness as a reliable indicator of local changes in well-being. First of all, in chapter three, I argued that there are contexts in which happiness is systematically *dysfunctional*, and therefore does not tend to be reliably caused by well-being ingredients. In addition, there are contexts in which the function of happiness consists in detecting *non-prudential* features of our environment (such as moral, aesthetic or epistemic features). In chapter four, I argued that there are contexts in which the function of happiness consists in detecting our *level* of certain well-being ingredients or *unexpected* changes in well-being. Further philosophical and empirical research is required to determine these contexts, and thereby the contexts in which information about happiness does not tell us about local changes in well-being. Once these contexts have been determined and account for, we can treat happiness as a reliable indicator of local well-being.



## Bibliography

- Alexandrova, A., 2012a. Values and the Science of Well-Being: A Recipe for Mixing. In *The Oxford Handbook of Philosophy of Social Science*. Oxford University Press.
- Alexandrova, A., 2012b. Well-Being as an Object of Science. *Philosophy of Science*, 79(5): 678–689.
- Alexandrova, A., 2013. Doing Well in the Circumstances. *Journal of Moral Philosophy*, 10 (3): 307-328.
- Almeder, R.F., 2000. *Human Happiness and Morality: A Brief Introduction to Ethics*, Prometheus Books.
- Amabile, T. & Kramer, S., 2011. *The progress principle: Using small wins to ignite joy, engagement, and creativity at work*, Harvard Business Press.
- Andreou, C. 2010. A shallow route to environmentally friendly happiness: why evidence that we are shallow materialists need not be bad news for the environment (alist). *Ethics, Place & Environment*, 13(1), pp.1–10.
- Angner, E. 2008. The Politics of Happiness: Subjective vs. Economic Measures as Measures of Social Well-Being. Social Science Research Network. Available at: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1278163](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1278163) [Accessed January 7, 2014]
- Angner, E., 2010. Are subjective measures of well-being “direct”? *Australasian Journal of Philosophy*, 99999(1), pp.1–16.
- Angner, E., 2012. "Subjective Well-Being: When, and Why, it Matters". Social Science Research Network. Available at: <http://papers.ssrn.com/abstract=2157140> [Accessed January 7, 2014].
- Annas, J., 1995. *The Morality of Happiness*, Oxford University Press.
- Annas, J., 2008. The phenomenology of virtue. *Phenomenology and the Cognitive Sciences*, 7(1), pp.21–34.
- Annas, J., 2011. *Intelligent virtue*, Oxford University Press.
- Arneson, R., 2009. Human flourishing versus desire satisfaction. *Social Philosophy and Policy*, 16(01), pp.113–142.
- Baber, H., 2008. The Experience Machine Deconstructed. *Philosophy in the Contemporary World*, 15(1).

- Barber, A., 2011. Hedonism and the Experience Machine. *Philosophical Papers*, 40(2), pp.257–278.
- Baumeister, R., Vohs, K., Aaker, J., & Garbinsky, E., 2012. "Some key differences between a happy life and a meaningful life." Social Science Research Network. Available at: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2168436](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2168436) [Accessed January 7, 2014].
- Beardman, S., 2000. The choice between current and retrospective evaluations of pain. *Philosophical Psychology*, 13(1), pp.97–110.
- Becchetti, L., Trovato, G. & Londono Bedoya, D.A., 2011. Income, relational goods and happiness. *Applied economics*, 43(3), pp.273–290.
- Belshaw, C., 2012. What's Wrong with the Experience Machine? *European Journal of Philosophy*. Available at: <http://onlinelibrary.wiley.com> [Accessed January 7, 2014]
- Berridge, K.C., 1996. Food reward: brain substrates of wanting and liking. *Neuroscience & Biobehavioral Reviews*, 20(1), pp.1–25.
- Besser-Jones, L. Forthcoming. *Eudaimonic Ethics: A Contemporary Theory of the Philosophy and Psychology of Living Well*. Routledge.
- Betzler, M., 2013. The Normative Significance of Personal Projects. In M. Kühler & N. Jelinek, eds. *Autonomy and the Self*. Philosophical Studies Series. Springer Netherlands, pp. 101–126. Available at: <http://link.springer.com/chapter/10.1007/> [Accessed January 7, 2014].
- Bishop, M. Unpublished manuscript. *The Good Life: The Science and Philosophy of Well-being*.
- Blakemore, S.-J., Wolpert, D.M. & Frith, C.D., 1998. Central cancellation of self-produced tickle sensation. *Nature neuroscience*, 1(7), pp.635–640.
- Blanchflower, D.G. & Oswald, A.J., 2008. Is well-being U-shaped over the life cycle? *Social Science & Medicine*, 66(8), pp.1733–1749.
- Bok, D., 2010. *The Politics of Happiness: What Government Can Learn from the New Research on Well-Being*. Princeton University Press.
- Booth, P. Ed. 2012. *...and the Pursuit of Happiness: Wellbeing and the Role of Government*. Readings 64, pp 245. Available at: <http://www.iea.org.uk/publications/research/and-the-pursuit-of-happiness> [Accessed Jan 7, 2014].
- Van Boven, L. & Gilovich, T., 2003. To do or to have? That is the question. *Journal of personality and social psychology*, 85(6), p.1193.
- Braddock, M., 2010. Constructivist experimental philosophy on well-being and virtue. *Southern Journal of Philosophy*, 48(3), pp.295–323.

- Bradley, B., 2009. *Well-Being and Death*, Oxford University Press.
- Bramble, B. Unpublished manuscript. "Three Cheers for Hedonism about Well-Being."
- Bramble, B. Forthcoming. Whole-Life Welfarism. *American Philosophical Quarterly*.
- Brandt, R.B., 1998. *A Theory of the Good and the Right* 2nd Revised edition., Prometheus Books.
- Brickman, P., Coates, D. & Janoff-Bulman, R., 1978. Lottery winners and accident victims: Is happiness relative? *Journal of personality and social psychology*, 36(8), p.917.
- De Brigard, F., 2010. If you like it, does it matter if it's real? *Philosophical Psychology*, 23(1), pp.43–57.
- Brink, D., 1989. *Moral realism and the foundations of ethics*, Cambridge University Press.
- Broome, J., 1996. More pain or less? *Analysis*, 56(2), pp.116–118.
- Brülde, B., 2007. Happiness theories of the good life. *Journal of Happiness Studies*, 8(1), pp.15–49.
- Calhoun, C., 2009. What Good Is Commitment? *Ethics*, 119(4), pp.613–641.
- Calhoun, C. Unpublished manuscript. *Meaningful Living*.
- Campbell, S. Manuscript. "Phenomenal well-being". Texas A&M University. Available at: <http://repository.tamu.edu/handle/1969.1/3834> [Accessed on Jan 7, 2014]
- Campbell, S. Forthcoming. An Analysis of Prudential Value. *Utilitas*.
- Carver, C., 2003. Pleasure as a sign you can attend to something else: Placing positive feelings within a general model of affect. *Cognition & Emotion*, 17(2), pp.241–261.
- Chang, H., 2004. *Inventing Temperature : Measurement and Scientific Progress: Measurement and Scientific Progress*, Oxford University Press.
- Chen, M. & Bargh, J.A., 1999. Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin*, 25(2), pp.215–224.
- Clark, A.E. et al., 2008. Lags And Leads in Life Satisfaction: a Test of the Baseline Hypothesis. *The Economic Journal*, 118(529), pp.222–243.

- Coote, A, & Franklin, J. 2010. *21 Hours*. Available at: <http://www.neweconomics.org/publications/entry/21-hours> [Accessed on Jan 7, 2014]
- Côté, S., 1999. Affect and performance in organizational settings. *Current Directions in Psychological Science*, 8(2), pp.65–68.
- Crisp, R., 2006a. Hedonism reconsidered. *Philosophy and Phenomenological Research*, 73(3), pp.619–645.
- Crisp, R., 2006b. *Reasons and the good*, Oxford University Press.
- Crisp, R., 2008. Well-being. *Stanford Encyclopedia of Philosophy*. Available at: <http://plato.stanford.edu/entries/well-being> [Accessed at Jan 7, 2014]
- Csikszentmihályi, M., 2008. *Flow: The Psychology of Optimal Experience*, HarperCollins.
- Csikszentmihalyi, M., Abuhamdeh, S. & Nakamura, J., 2005. Handbook of competence and motivation. *The Guilford Press, New York*, pp.598–608.
- Csikszentmihalyi, M. & Csikszentmihalyi, I.S., 1992. *Optimal experience: Psychological studies of flow in consciousness*, Cambridge University Press.
- Cudd, A.E., 2006. *Analyzing oppression*, Oxford University Press.
- Cunningham, M.R., 1988. Does happiness mean friendliness? Induced mood and heterosexual self-disclosure. *Personality and Social Psychology Bulletin*, 14(2), pp.283–297.
- Damasio, A., 2005. *Descartes' Error: Emotion, Reason, and the Human Brain*, Penguin.
- Darwall, S., 2004. *Welfare and Rational Care*, Princeton University Press.
- Davidson, R.J., 2004. Well-being and affective style: neural substrates and biobehavioural correlates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359(1449), p.1395.
- Davis, W., 1981. A theory of happiness. *American Philosophical Quarterly*, pp.111–120.
- Deaton, A., 2010. *Income, aging, health and well-being around the world: Evidence from the Gallup World Poll*, University of Chicago Press.
- Deci, E.L. & Moller, A.C., 2005. The Concept of Competence: A Starting Place for Understanding Intrinsic Motivation and Self-Determined Extrinsic Motivation. In Elliot, A., & Dweck, C., (Eds.) *Handbook of competence and motivation* (pp. 579-597). New York: Guilford Press.

- Deci, E.L. & Ryan, R.M., 2000. The “ what” and “ why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry*, 11(4), pp.227–268.
- DeLancey, C., 2006. Basic Moods. *Philosophical Psychology*, 19(4), pp.527–538.
- Delle Fave, A., & Massimini, F. 1992. Modernisation and the changing context of flow in work and leisure. In Csikszentmihalyi, M, & Csikszentmihalyi, I. *Optimal experience: Psychological studies of flow and consciousness*. Cambridge University Press.
- Diener, E. et al., 1985. The satisfaction with life scale. *Journal of personality assessment*, 49(1), pp.71–75.
- Diener, E. et al., 1999. Subjective well-being: Three decades of progress. *Psychological bulletin*, 125(2), p.276.
- Diener, E. & Suh, E., 1997. Measuring quality of life: Economic, social, and subjective indicators. *Social indicators research*, 40(1-2), pp.189–216.
- Diener, E., Wirtz, D. & Oishi, S., 2001. End effects of rated life quality: the James Dean Effect. *Psychological science*, 12(2), pp.124–128.
- Diener, E. & Biswas-Diener, R., 2008. *Happiness: Unlocking the Mysteries of Psychological Wealth*, Wiley-Blackwell.
- Diener, E. et al., 2009. *Well-Being for Public Policy* 1st ed., Oxford University Press.
- Diener, E., Inglehart, R. & Tay, L., 2012. Theory and validity of life satisfaction scales. *Social Indicators Research*, pp.1–31.
- Dolan, P., Layard, R. & Metcalfe, R., 2011. Measuring subjective well-being for public policy. Available at: <http://eprints.lse.ac.uk/35420> [Accessed on January 7, 2014]
- Dolan, P., Peasgood, T. & White, M., 2008. Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. *Journal of Economic Psychology*, 29(1), pp.94–122.
- Döring, S.A., 2007. Seeing What to Do: Affective Perception and Rational Motivation. *Dialectica*, 61(3), pp.363–394.
- Doris, J. 2002. *Lack of Character: Personality and Moral Behaviour*. Cambridge University Press.
- Dorsey, D., 2011a. First steps in an axiology of goals. *International Journal of Wellbeing*, 1(1).

- Dorsey, D., 2011b. The Hedonist's Dilemma. *Journal of Moral Philosophy*, 8(2), pp.173–196.
- Dorsey, D. Unpublished manuscript. "The Significance of a Life's Shape".
- Dunn, E.W., Gilbert, D.T. & Wilson, T.D., 2011. If money doesn't make you happy, then you probably aren't spending it right. *Journal of Consumer Psychology*, 21(2), p.115.
- Dworkin, G., 1988. *The theory and practice of autonomy*, Cambridge University Press.
- Easterlin, R., 1974. Does economic growth improve the human lot? Some empirical evidence. *Nations and households in economic growth*, 89.
- Easterlin, R., 2001. Income and happiness: Towards a unified theory. *The economic journal*, 111(473), pp.465–484.
- Easterlin, R. et al., 2010. The happiness–income paradox revisited. *Proceedings of the National Academy of Sciences*, 107(52), pp.22463–22468.
- Ekman, P., 1999. Basic emotions. *Handbook of cognition and emotion*, 4, pp.5–60.
- Ekman, P.E. & Davidson, R.J., 1994. *The nature of emotion: Fundamental questions.*, Oxford University Press.
- Elster, J., 2011. How Outlandish Can Imaginary Cases Be? *Journal of Applied Philosophy*, 28(3), pp.241–258.
- Emmons, R., 2007. *Thanks!: How the new science of gratitude can make you happier*, Houghton Mifflin Harcourt.
- Faucher, L. & Tappolet, C., 2008. *The modularity of emotions*, University of Calgary Press.
- Feldman, F., 2002. The good life: A defense of attitudinal hedonism. *Philosophy and Phenomenological Research*, 65(3), pp.604–628.
- Feldman, F., 2004. *Pleasure and the good life: concerning the nature, varieties and plausibility of hedonism*, Oxford University Press.
- Feldman, F., 2010. *What is this thing called happiness?*, Oxford University Press.
- Feldman, F., 2012. What we learn from the experience machine. In Bader, R., & Meadowcroft, J. *The Cambridge Companion to Nozick's Anarchy, State, and Utopia*. Cambridge University Press.

- Finn, B., 2010. Ending on a high note: Adding a better end to effortful study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(6), p.1548.
- Fitzroy, F. et al., 2013. *Testing the Tunnel Effect: Comparison, Age and Happiness in UK and German Panels*, University of St. Andrews. Available at: <http://ideas.repec.org/p/san/wpecon/1304.html> [Accessed January 7, 2014].
- Flanagan, O., 2011. *The Bodhisattva's Brain*, The MIT Press.
- Fletcher, G., 2013. A Fresh Start for the Objective-List Theory of Well-Being. *Utilitas*, 25(2): 206-220.
- Foot, P., 2001. *Natural goodness*, Oxford University Press.
- Frank, R., 2001. *Luxury fever: Why money fails to satisfy in an era of excess*, Simon and Schuster.
- Frank, R., 2012. *The Darwin economy: liberty, competition, and the common good*, Princeton University Press.
- Frankfurt, H.G., 1988. *Freedom of the Will and the Concept of a Person*, Springer.
- Frederick, S. & Loewenstein, G., 1999. Hedonic adaptation. In Kahneman, D., Diener, E., Schwartz, N. (Eds.) *Well-being: The foundations of hedonic psychology*. New York: Russell Sage Foundation.
- Fredrickson, B., 1998. What good are positive emotions?. *Review of general psychology*, 2(3), p.300.
- Fredrickson, B., 2001. The Role of Positive Emotions in Positive Psychology: The Broaden-and-Build Theory of Positive Emotions. *American Psychologist*, 56(3), pp.218–226.
- Fredrickson, B., 2013. Updated thinking on positivity ratios. *American Psychologist*, 68(9).
- Fredrickson, B.L. & Joiner, T., 2002. Positive emotions trigger upward spirals toward emotional well-being. *Psychological science*, 13(2), pp.172–175.
- Fredrickson, B., Tugade, M., Waugh, E., Larkin, G. 2003. What good are positive emotions in crisis? A prospective study of resilience and emotions following the terrorist attacks on the United States on September 11th, 2001. *Journal of personality and Social Psychology*, 84(2), p.365.
- Fredrickson, B.L. & Losada, M.F., 2005. Positive affect and the complex dynamics of human flourishing. *American Psychologist*, 60(7), p.678.

- Frey, B.S., 2010. *Happiness: A Revolution in Economics*, The MIT Press.
- Frijda, N., 1988. The laws of emotion. *American psychologist*, 43(5), p.349.
- Gailliot, M., 2012. Happiness as Surplus or Freely Available Energy. *Psychology*, 3(9), pp.702–712.
- Gheaus, A., forthcoming. The Intrinsic Goods of Childhood and the Good Society. In *The Well-Being of Children in Theory and Practice*. Springer.
- Goldie, P., 2002. *The Emotions: A Philosophical Exploration*, Oxford University Press.
- Graham, C., 2010. *Happiness Around the World: The paradox of happy peasants and miserable millionaires*, Oxford University Press.
- Graham, C. & Pettinato, S. 2002. *Happiness and Hardship: Opportunity and Insecurity in New Market Economics*. The Brookings Institution. Graph available at: <http://www.brookings.edu/research/opinions/2010/01/30-happiness-graham>
- Griffin, J., 1989. *Well-being: its meaning, measurement, and moral importance*, Clarendon Press.
- Griffin, J., 2007. What Do Happiness Studies Study? *Journal of Happiness Studies*, 8(1), pp.139–148.
- Griffin, J., Crisp, R. & Hooker, B., 2000. *Well-being and morality: Essays in honour of James Griffin*, Oxford University Press.
- Griffiths, P.E., 1997. *What Emotions Really Are: The Problem of Psychological Categories* 1st ed., University Of Chicago Press.
- Hansen, T., 2012. Parenthood and Happiness: a Review of Folk Theories Versus Empirical Evidence. *Social Indicators Research*, 108(1): 29–64.
- Hausman, D., 2010. Hedonism and welfare economics. *Economics and Philosophy*, 26(03), pp.321–344.
- Hausman, D., 2012. Why Satisfy Preferences? *Papers on Economics and Evolution*. Available at <http://ideas.repec.org/s/esi/evopap.html> [Accessed on Jan 7, 2014]
- Hawkins, J., 2008. Well-Being, Autonomy, and the Horizon Problem. *Utilitas*, 20(02), pp.143–168.
- Hawkins, J. 2010. "A new theory of well-being." Unpublished paper presented at the annual conference of the *British Society for Ethical Theory*, University of Nottingham, 9th July 2010.



- Haybron, D.M., 2000. Two philosophical problems in the study of happiness. *Journal of Happiness Studies*, 1(2), pp.207–225.
- Haybron, D.M., 2001. Happiness and Pleasure. *Philosophy and Phenomenological Research*, 62(3), pp.501–528.
- Haybron, D.M., 2003. What do we Want from a Theory of Happiness? *Metaphilosophy*, 34(3), pp.305–329.
- Haybron, D.M., 2005. On Being Happy or Unhappy. *Philosophy and Phenomenological Research*, 71(2), pp.287–317.
- Haybron, D.M., 2007. Do we know how happy we are? On some limits of affective introspection and recall. *Nous*, 41(3), pp.394–428.
- Haybron, D.M., 2008a. Happiness, the Self and Human Flourishing. *Utilitas*, 20(01), pp.21–49.
- Haybron, D.M., 2008b. *The pursuit of unhappiness: the elusive psychology of well-being*, Oxford University Press.
- Haybron, D., 2011a. Happiness. *Stanford Encyclopaedia of Philosophy*. Available at: <http://plato.stanford.edu/entries/happiness> [Accessed at Jan 7, 2014]
- Haybron, D., 2011b. Central Park: Nature, context, and human wellbeing. *International Journal of Wellbeing*, 1(2).
- Haybron, D.M., 2013. The Proper Pursuit of Happiness. *Res Philosophica*, 90(3), pp.387–411.
- Haybron, D. Forthcoming. The value of positive emotion: Philosophical doubts and reassurances. In Gruber, J. & Moskowitz. (Eds.) *The dark and light sides of positive emotion*. Oxford University Press.
- Haybron, D. Unpublished Manuscript. "Eudaimonism, Ancient and Modern".
- Haybron, D. & Tiberius, V. Unpublished Manuscript. "The what of well-being policy".
- Headey, B., 2007. The Set-Point Theory of Well-Being: Negative Results and Consequent Revisions. *Social Indicators Research*, 85(3), pp.389–403.
- Headey, B. & Wearing, A., 1989. Personality, life events, and subjective well-being: toward a dynamic equilibrium model. *Journal of Personality and Social psychology*, 57(4), p.731.
- Heathwood, C. 2005. *Desire-satisfaction theories of welfare*. University of Massachusetts Amherst.
- Heathwood, C. 2006. Desire satisfactionism and hedonism. *Philosophical Studies*, 128(3), pp.539–563.

- Heathwood, C. 2010. Welfare. *The Routledge Companion to Ethics*. London and New York: Routledge, pp.645–655.
- Heathwood, C. 2013. Hedonism. In *The International Encyclopedia of Ethics*. Available at: <http://onlinelibrary.wiley.com/book/10.1002/9781444367072> [Accessed Jan 7, 2014]
- Helliwell, J., Layard, R. & Sachs, J., 2012. World happiness report. Available at: <http://eprints.lse.ac.uk/47487> [Accessed Jan 7, 2014]
- Helm, B.W., 2009. Emotions as evaluative feelings. *Emotion Review*, 1(3), pp.248–55.
- Hewitt, S., What do our intuitions about the experience machine really tell us about hedonism? *Philosophical Studies*, pp.1–19.
- Hoffman, M.L., 2001. *Empathy and moral development: Implications for caring and justice*, Cambridge University Press.
- Hsee, C., Salovey, P. & Abelson, R., 1994. The quasi-acceleration relation: Satisfaction as a function of the change of velocity of outcome over time. *Journal of Experimental Social Psychology*, 30, pp.96–111.
- Hsee, C.K., Abelson, R.P. & Salovey, P., 1991. The relative weighting of position and velocity in satisfaction. *Psychological Science*, 2(4), pp.263–266.
- Hurka, T., 2011. *The Best Things in Life: A Guide to What Really Matters*, Oxford University Press.
- Hursthouse, R., 1999. *On virtue ethics*, Oxford University Press.
- Inglehart, R. et al., 2008. Development, freedom, and rising happiness: A global perspective (1981–2007). *Perspectives on Psychological Science*, 3(4), p.264.
- Isen, A.M., 1970. Success, failure, attention, and reaction to others: The warm glow of success. *Journal of Personality and Social Psychology*, 15(4), p.294.
- Isen, A.M. & Daubman, K.A., 1984. The influence of affect on categorization. *Journal of personality and Social Psychology*, 47(6), p.1206.
- Isen, A.M., Daubman, K.A. & Nowicki, G.P., 1987. Positive affect facilitates creative problem solving. *Journal of personality and social psychology*, 52(6), p.1122.
- Kagan, S., 1992. The Limits of Well-Being. *Social Philosophy and Policy*, 9(02), pp.169–189.

- Kagan, S., 2009. Well-being as Enjoying the Good. *Philosophical Perspectives*, 23(1), pp.253–272.
- Kahneman, D., 1999. Objective happiness. In Kahneman, D., Diener, E., Schwartz, N. (Eds.) *Well-being: The foundations of hedonic psychology*. New York: Russell Sage Foundation.
- Kahneman, D., 2011. *Thinking, Fast and Slow*, Penguin.
- Kahneman, D. et al., 1993. When more pain is preferred to less: Adding a better end. *Psychological Science*, 4(6), pp.401–405.
- Kahneman, D., Wakker, P.P. & Sarin, R., 1997. Back to Bentham? Explorations of Experienced Utility. *The Quarterly Journal of Economics*, 112(2), pp.375–405.
- Kahneman, D. et al., 2004. Toward national well-being accounts. *The American Economic Review*, 94(2), pp.429–434.
- Kahneman, D. & Krueger, A.B., 2006. Developments in the measurement of subjective well-being. *The journal of economic perspectives*, 20(1), pp.3–24.
- Kahneman, D. & Deaton, A., 2010. High income improves evaluation of life but not emotional well-being. *Proceedings of the National Academy of Sciences*, 107(38), pp.16489–16493.
- Kasser, T., 2002. *The high price of materialism*, The MIT Press.
- Kauppinen, A., 2012. Meaningfulness and Time. *Philosophy and Phenomenological Research*, 84(2), pp.345–377.
- Kauppinen, A. 2010. Virtue and Happiness. Unpublished paper presented at a public lecture given at Trinity College Dublin, 14th December 2010.
- Kawall, J., 1999. The experience machine and mental state theories of well-being. *The Journal of Value Inquiry*, 33(3), pp.381–387.
- Keller, S., 2013. *Partiality*, Princeton University Press.
- Keller, S., 2004. Welfare and the Achievement of Goals. *Philosophical Studies*, 121(1), pp.27–41.
- Keller, S., 2009. Welfare as Success. *Noûs*, 43(4), pp.656–683.
- Keyes, C.L.M., Shmotkin, D. & Ryff, C.D., 2002. Optimizing well-being: the empirical encounter of two traditions. *Journal of personality and social psychology*, 82(6), p.1007.
- Kim, R. Unpublished manuscript. "What Is This Thing Called Well-Being".

- Kolber, A. 1994. Mental Statism and the Experience Machine. *Bard Journal of Social Sciences*, Vol. 3, pp. 10-17.
- Kraut, R., 2009. *What Is Good and Why: The Ethics of Well-Being*, Harvard University Press.
- Kriegel, U., 2011. Towards a New Feeling Theory of Emotion. *European Journal of Philosophy*, 20(3).
- Lane, R.E., 2000. *The loss of happiness in market democracies*, Yale University Press.
- Lauinger, W., 2013. The Missing-Desires Objection to Hybrid Theories of Well-Being. *Southern Journal of Philosophy*, 51(2), pp.270–295.
- Layard, R., 2005. *Happiness: Lessons from a New Science*, Penguin.
- Lewis, J., 2013. *Beyond consumer capitalism: Media and the limits to imagination*. Polity.
- Luhmann, M. et al., 2012. Subjective well-being and adaptation to life events: A meta-analysis. *Journal of Personality and Social Psychology*, 102(3), p.592.
- Lukas, M. et al., 2010. Desire Satisfactionism and the Problem of Irrelevant Desires. *Journal of Ethics & Social Philosophy*, 4(2).
- Lykken, D., 1999. *Happiness: What studies on twins show us about nature, nurture, and the happiness set-point.*, Golden Books.
- Lyubomirsky, S., King, L. & Diener, E., 2005. The Benefits of Frequent Positive Affect: Does Happiness Lead to Success?. *Psychological bulletin*, 131(6), p.803.
- Lyubomirsky, S. & Tucker, K.L., 1998. Implications of individual differences in subjective happiness for perceiving, interpreting, and thinking about life events. *Motivation and emotion*, 22(2), pp.155–186.
- Marmot, M., 2004. Status syndrome. *Significance*, 1(4), pp.150–154.
- Mauss, I.B. et al., 2011. Don't hide your happiness! Positive emotion dissociation, social connectedness, and psychological functioning. *Journal of personality and social psychology*, 100(4), p.738.
- McMahon, D.M., 2006. *Happiness: A History*, Grove Press.
- Mendola, J., 2006. Intuitive Hedonism. *Philosophical Studies*, 128(2), pp.441–477.
- Mill, J.S., 2002. *Utilitarianism (Second Edition)*, Hackett Publishing.

- Millgram, E., 2000. What's the Use of Utility? *Philosophy & Public Affairs*, 29(2), pp.113–136.
- Morris, S., 2011. In defense of the hedonistic account of happiness. *Philosophical Psychology*, 24(2), pp.261–281.
- Murphy, M.C., 1999. The simple desire-fulfillment theory. *Noûs*, 33(2), pp.247–272.
- Nesse, R.M., 2004. Natural selection and the elusiveness of happiness. *Philosophical Transactions-Royal Society of London Series B Biological Sciences*, pp.1333–1348.
- Nettle, D., 2007. *What makes you the way you are*, Oxford University Press.
- Noddings, N., 2004. *Happiness and Education* New Ed., Cambridge University Press.
- Nozick, R., 1974. *Anarchy, state, and utopia*, Basic Books.
- Nussbaum, M., 1987. *Nature, function, and capability: Aristotle on political distribution*, World Institute for Development Economics Research of the United Nations University Helsinki, Finland.
- Nussbaum, M., 2001. *Women and human development: The capabilities approach*, Cambridge University Press.
- Nussbaum, M., 2003. *Upheavals of Thought: The Intelligence of Emotions*, Cambridge University Press.
- Nyholm, S., 2007. Moral Judgments and Happiness. *Unpublished manuscript. University of Michigan*.
- Offer, A., 2006. *The Challenge of Affluence: Self-Control and Well-Being in the United States and Britain since 1950*, Oxford University Press.
- Olsaretti, S., 2006. *Preferences and Well-Being*, Cambridge University Press.
- Oswald, A.J. & Wu, S., 2010. Objective confirmation of subjective measures of human well-being: Evidence from the USA. *Science*, 327(5965), p.576.
- Oswald, A. 2006. The Hippies Were Right all Along about Happiness. *Financial Times*, 19th January 2006.
- Overvold, M.C., 1980. Self-interest and the Concept of Self-Sacrifice. *Canadian Journal of Philosophy*, pp.105–118.
- Parfit, D., 1984. *Reasons and persons*, Oxford University Press.
- Phillips, J., Misenheimer, L. & Knobe, J., forthcoming. The Ordinary Concept of Happiness (and Others Like It). *Emotion Review*, 71, pp.929–937.

- Pinker, S., 2012. *The Better Angels of Our Nature: Why Violence Has Declined*, Penguin Group (USA) Incorporated.
- Portmore, D.W., 2007. Welfare, Achievement, and Self-Sacrifice. *Journal of Ethics & Social Philosophy*, 2(2).
- Price, C., 2012. The Problem of Emotional Significance. *Acta Analytica*: 1–18.
- Prinz, 2007. All consciousness is perceptual. In B. P. McLaughlin & J. D. Cohen, eds. *Contemporary Debates in Philosophy of Mind*. Blackwell.
- Prinz, 2004a. *Gut reactions: A perceptual theory of emotion*, Oxford University Press.
- Prinz, 2004b. Which emotions are basic. *Emotion, evolution, and rationality*, pp.69–87.
- Prinz, J.J., 2007. *The emotional construction of morals*, Oxford University Press.
- Putnam, R.D., 1995. Bowling alone: America's declining social capital. *Journal of democracy*, 6, pp.65–65.
- Rachels, S. & Rachels, J., 2010. *The elements of moral philosophy*, New York: McGraw-Hill Higher Education.
- Rachlin, H., 2004. *The science of self-control*, Harvard Univ Pr.
- Raibley, J., 2010. Well-Being and the Priority of Values. *Social theory and practice*, 36(4), pp.593–620.
- Raibley, J.R., 2012. Welfare over Time and the Case for Holism. *Philosophical Papers*, 41(2), pp.239–265.
- Railton, P., 1984. Alienation, consequentialism, and the demands of morality. *Philosophy and Public Affairs*, 13(2), pp.134–171.
- Railton, P., 2003. *Facts, Values, and Norms: Essays toward a Morality of Consequence*, Cambridge University Press.
- Railton, P., 2008. The Problem of Well-Being: Respect, Equality, and the Self.
- Railton, P. Unpublished Manuscript. "Subjective Well-being as Information and Guidance".
- Rawls, J., 2005. *A Theory of Justice* reissue., Harvard University Press.
- Raz, J., 1988. *The Morality of Freedom* New Ed., Clarendon Press.
- Raz, J., 2004. The Role of Well-Being. *Philosophical Perspectives*, 18(1), pp.269–294.

- Reed, A.E. & Carstensen, L.L., 2012. The theory behind the age-related positivity effect. *Frontiers in Emotion Science*, 3, p.339.
- Reisenzein, R., 2009. Emotional experience in the computational belief–desire theory of emotion. *Emotion Review*, 1(3), pp.214–222.
- Reiss, S., 2004. Multifaceted nature of intrinsic motivation: The theory of 16 basic desires. *Review of General Psychology*, 8(3), p.179.
- Roberts, R.C., 2003. *Emotions: an essay in aid of moral psychology*, Cambridge University Press.
- Robinson M. D., 2000. The Reactive and Prospective Functions of Mood: Its Role in Linking Daily Experiences and Cognitive Well-being. *Cognition and Emotion*, 14(2), pp.145–176.
- Rosati, C. 2006. Personal Good. In *Metaethics After Moore*. Oxford University Press.
- Rosati, C. 1995. Persons, perspectives, and full information accounts of the good. *Ethics*, 105(2), pp.296–325.
- Rosati, C. 2009. Self-Interest and Self-Sacrifice. In *Proceedings of the Aristotelian Society (Hardback)*. Wiley Online Library, pp. 311–325.
- Rosati, C. Narrative and Personal Good. Available at: <http://onthehuman.org/2010/01/narrative-and-personal-good> [Access Jan 7, 2014]
- Russell, J.A., Lewicka, M. & Niit, T., 1989. A cross-cultural study of a circumplex model of affect. *Journal of personality and social psychology*, 57(5), p.848.
- Ryan, R.M. & Deci, E.L., 2001. On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual review of psychology*, 52(1), pp.141–166.
- Ryan, R.M. & Deci, E.L., 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), pp.68–78.
- Sacks, D.W., Stevenson, B. & Wolfers, J., 2010. Subjective Well-Being, Income, Economic Development and Growth. *National Bureau of Economic Research Working Paper Series*, No. 16441. Available at: <http://www.nber.org/papers/w16441> [Accessed November 8, 2010].
- Sarch, A.F., 2009. On the objectivity of welfare. *Unpublished Manuscript*. University of Massachusetts Amherst.
- Scanlon, T., 1998. *What we owe to each other*, Harvard University Press.

- Schroeder, T., 2006. An Unexpected Pleasure. *The modularity of emotions*, 32, pp.255–272.
- Schroeder, T., 2001. Pleasure, displeasure, and representation. *Canadian Journal of Philosophy*, 31(4), pp.507–530.
- Schroeder, T., 2004. *Three faces of desire*, Oxford University Press.
- Schultz, W., Dayan, P. & Montague, P.R., 1997. A neural substrate of prediction and reward. *Science*, 275(5306), pp.1593–1599.
- Schwartz, B., 2005. *The Paradox of Choice: Why More is Less* New edition., HarperCollins.
- Schwarz, N. & Strack, F., 1999. Reports of subjective well-being: Judgmental processes and their methodological implications. *Well-being: The foundations of hedonic psychology*, pp.61–84.
- Seligman, M., 2006. *Learned optimism: how to change your mind and your life*, New York: Vintage Books.
- Seligman, M., 2011. *Flourish: A Visionary New Understanding of Happiness and Well-being* Reprint., Free Press.
- Seligman, M. et al., 2005. Positive psychology progress: empirical validation of interventions. *American psychologist*, 60(5), p.410.
- Sen, A., 1999. *Development as freedom*, Oxford University Press.
- Sen, A., 2004. *Rationality and Freedom*, Harvard University Press.
- Sheldon, K., 2004. *Optimal human being: An integrated multi-level perspective*, Psychology Press.
- Sheldon, K. & Lyubomirsky, S., 2012. The Challenge of Staying Happier: Testing the Hedonic Adaptation Prevention Model. *Personality and Social Psychology Bulletin*.
- Shriver, A., 2014. The Asymmetrical Contributions of Pleasure and Pain to Subjective Well-Being. *Review of Philosophy and Psychology*: 1–19.
- Siemer, M., 2009. Mood experience: Implications of a dispositional theory of moods. *Emotion Review*, 1(3), p.256.
- Silverstein, M., 2000. In defense of happiness: A response to the experience machine. *Social Theory and Practice*, 26(2), pp.279–300.
- Sizer, L., 2000. Towards a computational theory of mood. *The British Journal for the Philosophy of Science*, 51(4), p.743.
- Sizer, L., 2010. Good and Good For You: An Affect Theory of Happiness<sup>1</sup>. *Philosophy and Phenomenological Research*, 80(1), pp.133–163.



- Slote, M., 1983. *Goods and virtues*, Clarendon Press Oxford.
- Smuts, A. Unpublished manuscript. "A Life Worth Living".
- Snow, N.E., 2008. Virtue and flourishing. *Journal of Social Philosophy*, 39(2), pp.225–245.
- Stevenson, B. & Wolfers, J., 2008. *Economic growth and subjective well-being: Reassessing the Easterlin Paradox*, National Bureau of Economic Research.
- Stiglitz, J.E., Sen, A. & Fitoussi, J.-P., 2009. Report by the commission on the measurement of economic performance and social progress.
- Strawson, G., 2004. Against narrativity. *Ratio*, 17(4), pp.428–452.
- Sumner, L.W., 1996. *Welfare, happiness, and ethics*, Oxford Univ Pr on Demand.
- Tännsjö, T., 2007. Narrow hedonism. *Journal of Happiness Studies*, 8(1), pp.79–98.
- Tappolet, C., 2010. Emotion, motivation, and action: The case of fear. *The Oxford handbook of the philosophy of emotion*, pp.325–345.
- Tatarkiewicz, W., 1976. *Analysis of Happiness*, Kluwer Academic Publishers.
- Taylor, T.E., 2012. *Knowing what is good for you: a theory of prudential value and well-being*, New York: Palgrave Macmillan.
- Taylor, T. Forthcoming. Towards Consensus on Well-being. In Søraker, J. Wong, P., Boer, J., De Rijdt, J., & Brey, P. (Eds.) *Well-Being in Contemporary Society*. Berlin: Singer.
- Telfer, E., 1980. *Happiness*, Palgrave Macmillan.
- Thoits, P.A. & Hewitt, L.N., 2001. Volunteer work and well-being. *Journal of Health and Social Behavior*, pp.115–131.
- Tiberius, V., 2004. Cultural differences and philosophical accounts of well-being. *Journal of Happiness Studies*, 5(3), pp.293–314.
- Tiberius, V., 2005a. Value Commitments and the Balanced Life. *Utilitas*, 17(01), pp.24–45.
- Tiberius, V., 2005b. Wisdom and perspective. *The Journal of philosophy*, 102(4), pp.163–182.
- Tiberius, V., 2007. Substance and procedure in theories of prudential value. *Australasian Journal of Philosophy*, 85(3), pp.373–391.

- Tiberius, V., 2008. *The reflective life: living wisely with our limits*, Oxford University Press, USA.
- Tiberius, V., 2013a. Maximization and the Good. In H. Brockmann & J. Delhey, eds. *Human Happiness and the Pursuit of Maximization*. Happiness Studies Book Series. Springer, pp. 55–67. Available at: <http://link.springer.com/chapter/> [Accessed January 7, 2014].
- Tiberius, V., 2013b. Recipes for a good life: Eudaimonism and the contribution of philosophy. *The best within us: Positive psychology perspectives on eudaimonia*, pp.19–38.
- Umberson, D., Pudrovska, T. & Reczek, C., 2010. Parenthood, childlessness, and well-being: A life course perspective. *Journal of Marriage and Family*, 72(3), pp.612–629.
- Veenhoven, R. & Hagerty, M., 2006. Rising happiness in nations 1946–2004: A reply to Easterlin. *Social Indicators Research*, 79(3), pp.421–436.
- Velleman, D. 1991. Well-being and time. *Pacific Philosophical Quarterly*, 72(1), pp.48–77.
- Velleman, D. 1993. Well-Being and Time. Reprinted in Fischer, J. M. (Ed.) *The Metaphysics of Death*. 329-3. Stanford: Stanford University Press. The article originally appeared in *Pacific Philosophical Quarterly*, 72 (1991): 48-77.
- Watson, G., 1987. Free action and free will. *Mind*, 96(382), pp.145–172.
- Weijers, D., 2011. Intuitive Biases in Judgments about Thought Experiments: The Experience Machine Revisited. *Philosophical Writings*, (No. 50 & No. 51, Summer & Autumn).
- Weijers, D., & Schouten, V. Forthcoming. As assessment of recent responses to the experience machine objection to hedonism. *Journal of Value Inquiry*.
- Weiss, A. et al., 2012. Evidence for a midlife crisis in great apes consistent with the U-shape in human well-being. *Proceedings of the National Academy of Sciences*. Available at: <http://www.pnas.org/content/early/2012/11/14/1212592109> [Accessed Jan 7, 2014].
- White, N., 2006. *A Brief History of Happiness*, Wiley-Blackwell.
- Whiting, D., 2011. The Feeling Theory of Emotion and the Object-Directed Emotions. *European Journal of Philosophy*, 19(2), pp.281–303.
- Wilkinson, R.G. et al., 2009. *The spirit level: why more equal societies almost always do better*, Allen Lane London.
- Williams, B., 2013. *Ethics and the Limits of Philosophy*, Routledge.

- Wilson, T.D. & Gilbert, D.T., 2008. Explaining Away: A Model of Affective Adaptation. *Perspectives on Psychological Science*, 3(5), pp.370–386.
- Winkielman, P. & Berridge, K.C., 2004. Unconscious emotion. *Current Directions in Psychological Science*, 13(3), pp.120–123.
- Wolff, J. & De-Shalit, A., 2007. *Disadvantage*, Oxford University Press, USA.
- Woodard, C., 2013. Classifying theories of welfare. *Philosophical Studies*, 165(3), pp.787–803.
- Wren-Lewis, S. 2013. Well-being as a Primary Good: Towards a Legitimate Role for Well-being in Public Policy. *Philosophy and Public Policy Quarterly*, 31 (2).
- Wren-Lewis, S. Forthcoming. Towards a Widely Acceptable Framework for the Study of Personal Well-being. In Søraker, J., Wong, P., Boer, J., De Rijdt, J., & Brey, P. (Eds.) *Well-Being in Contemporary Society*. Berlin: Singer.

## **Appendix A: Happiness as an Indicator of Changes in Well-being in the Near-Future**

In chapter three, I argued in favour of the Indicator View – the view that happiness is a defeasible indicator of well-being. In chapter four, I refined the Indicator View, adding that happiness is a defeasible indicator of local changes in well-being. In this section, I want to add one more claim to the Indicator view, namely that happiness tends to indicate near-future well-being. As with chapters three and four, this claim can be justified on the basis of the function of affective states. I will argue that, in order to inform and guide our action, trends in our affective states over time orientate us towards future opportunities to promote well-being, the result of which is that happiness tends to indicate our well-being in the near-future.

The function of happiness is both to *inform* and *guide* our action. So far, in discussing the Indicator View, we have mostly considered the informational function of happiness. Affective states should detect evaluative features of our environment that are relevant for action – it is for this reason that they should tend to be caused by local changes in well-being. But, affective states also have various attentional, cognitive and motivational effects that have the function of guiding our action in light of this information. Consider, for example, what happens during an episode of fear caused by a huge wolf leaping out in front of you (Tappolet, 2010). First of all, we see and hear the wolf and implicitly evaluate the situation as threatening – this is the informational component of our fear. This information then results in a number of cognitive and physiological changes that guides our action: our heart races, breathing becomes strained, mouth opens and eyes widen; we feel a pang in our chest and adrenaline pump through our body; we may closely attend to the movement of the wolf and possible exit strategies; we may have urge towards fight or flight; and we may think that we're done for unless we act soon. These attentional, cognitive and motivational effects guide us towards specific actions that will hopefully enable us to survive the dangerous situation.

In general, positive affective states motivate us to sustain or seek out the things that they represent as being good, and negative affective states

motivate us to avoid or prevent the things that they represent as being bad. Thus, both positive and negative affective states tend to guide our behaviour in beneficial ways. A person's affective states, on a very simple level, indicate that their well-being will improve in the near-future to some extent, in that they motivate the promotion of their well-being in certain ways.

However, this fact about affective states alone does not provide us with much information about a person's near-future well-being. We know that, however well or badly a person is doing, their affective states will motivate them to improve their situation. But this is true for all individuals, happy and unhappy. The fact that affective states motivate us to promote our well-being does not enable us to distinguish between two individuals on the basis of how well they are likely to be doing in the near-future.

In order to see how affective states may indicate well-being in the near-future, we need to look at *trends* of either positive or negative affective states over time. For the purposes of this discussion, I will focus on trends that consist in positive affective states over time. Positive affective states over time indicate that our well-being has consistently improved over that time, and thereby guided us to seek out similar increases in well-being in the future. The important point is that, *in addition*, such trends impart information about the opportunities for well-being in our environment and guide us towards taking them. Positive affective states over time may simply be caused by a string of good fortune, but are more likely to be caused by an environment that is full of opportunities for well-being. It is presumably because a person lives within a favourable environment (with regard to improving their well-being) that they have consistently improved their well-being over a certain period of time.

These kinds of trends in our affective states tend to cause moods. Moods indicate opportunities for well-being (or, in the case of negative moods, ill-being). Like emotions, moods involve various physiological and cognitive changes that guide us towards the promotion of our well-being. However, in contrast to emotions, which guide our action in specific ways in reaction to our immediate situation, moods guide our behaviour through orientating us towards the indicated opportunities for well-being in our environment. Consider, for example, the difference in function between anger (an emotion) and irritability (a mood). While the function of anger or fear is to enable us to deal effectively with an offence or threat that has actually occurred, the function of irritability or apprehension is to orientate us towards an

environment in which there is an increased probability that an offence or a threat will occur.

When we experience negative moods (caused by negative affective states over time) such as irritability, it is likely that we are in an unusually challenging or hostile environment. Our mood-state acts as a state of vigilance, and indicates a change in the probability that a certain type of situation will occur. In contrast, positive moods have been shown to have almost the opposite effect – they act as a state of abstracted, broadened attention (Fredrickson 1998; 2001). Subjects in positive moods tend to take in a wider range of information and categorise more broadly than those in negative moods; they do not focus on the small details, but instead perceive ‘the bigger picture’ (Isen & Daubman 1984). When people are in positive moods their brains are more flexibly organised and they tend to make more creative and unusual associations between ideas (Isen et al. 1987). This makes sense insofar as positive moods indicate an environment with abundant opportunities – one in which we need to broaden our mindset in order to take full advantage of.

In sum, happiness over time does not merely indicate consistent improvement in well-being over time, but also indicates an environment full of opportunities for well-being. Such positive trends tend to result in positive moods, which orientate us towards taking such opportunities, thereby further promoting our well-being. It is in this way that happiness over time may tend to indicate both local changes in well-being over that time and our well-being in the near-future.

Unfortunately, I am unable to consider this extended version of the Indicator View in more detail here. I will, however, end this section by outlining some empirical evidence in favour of the view. Firstly, people who experience positive affect more frequently than others tend to be more resilient (Fredrickson et al. 2003), resourceful (Lyubomirsky, King & Diener 2005), socially connected (Mauss et al. 2011) and more likely to function at optimal levels (Fredrickson & Losada 2005). Barbara Fredrickson argues that these benefits occur as a result of the kind of ‘broadened mindset’ involved in experiencing a positive mood. According to Fredrickson, this mindset is the recipe for discovery, discovery of knowledge, new alliances, and new skills. Over time, positive affect spurs the development of resources that contribute towards our well-being.

In a similar vein, Charles Carver argues that positive affect over time has the function of signalling that we are doing much better than expected in a

particular respect, and that we should therefore devote more resources to other areas of our lives. Thus, according to Carver, “positive affect induces an openness to considering possibilities other than the goal currently being pursued.” (Carver, 2003: p.255) Over time, this is also likely to spur the development of resources that contribute towards our well-being.

This broadened, or more balanced, mindset caused by positive affective states over time may indicate our well-being in the near-future in an additional sense. Such a mindset may enable us to appreciate our selves and our lives. This is perhaps best illustrated by the lack of appreciation of one’s self and one’s life that is associated with the mental condition of depression. Depressed individuals tend to lack a sense of self-worth and find no value in any personal activities. They are unable to appreciate the value of themselves and their lives, and thereby lack the motivation to improve their situation (Hawkins, MS). Thus, Haybron writes: “Depression strips pleasure from life, diminishing our functioning in myriad respects, and makes our projects, plans and relationships go worse. Severe anxiety similarly undermines our well-being in many ways. Whereas happiness tends to have the opposite effects.” (Haybron, forthcoming: p.11)

In sum, we have reason to believe that happiness over time both indicates local changes in well-being over that time and changes in well-being in the near-future. This extended version of the Indicator View may fully account for the ways in which people tend to use happiness as a proxy for well-being in their practical lives. Indeed, Haybron suggests that we sometimes make a judgement about how well off someone is on the basis of a few observations of their moods (Haybron, forthcoming). For example, he notes that, “if a person’s intimates find her consistently to be relaxed and in good spirits, resilient in the face of frustrations and other information, they will reasonably infer that she is happy, and doing well.” (Haybron, forthcoming: p.11) Such judgements are informative not just because someone’s affective states indicate how well they have been doing, but also because they tend to indicate how well someone is likely to continue to be doing in the future.

In the second Appendix of this thesis, I will consider one way in which happiness may be informative about a person’s well-being that is *not* reflected in the way we tend to use happiness as an informal proxy for well-being. The idea is that we may be able to infer high levels of well-being from low levels of variance in our affective states over time, and vice versa. We do not tend to use information about happiness in this way because it requires us to observe

patterns in our affective states over time that do not tend to be epistemically accessible. However, the formal measurement of happiness over time may make it possible to observe such patterns.



## **Appendix B: Happiness as an Indicator of Levels of Well-being**

In this thesis, I have argued that happiness (defeasibly) indicates local changes in well-being over time and in the near-future. In this final section, I will consider whether we may be able to infer a person's *level* of well-being from this information. Firstly, I will consider one method that does not work, namely inferring high levels of well-being from a high levels of positive affective states over time, and vice versa. Secondly, I will argue that a more sophisticated version of this method may work, namely inferring high levels of well-being from a low level of variance in our affective states over time, and vice versa.

### **B.1 The Progress Method**

The Progress Method assumes that high levels of positive affective states over time correlates with high levels of well-being. This method may seem intuitive at first. Consider, for example, the difference in well-being between a millionaire and a peasant – I will presume that the millionaire has a high level of well-being, whereas the peasant has a low level of well-being. It seems that the millionaire is in a much better position to improve their well-being than the peasant is. This is, presumably, because the millionaire has various resources and capabilities that the peasant lacks. These resources and capabilities enable the millionaire to improve their well-being in ways that are not available to the peasant.

It seems, therefore, that because the millionaire has a high level of well-being, they are likely to experience a greater amount of progress than the peasant. If this relationship between levels of well-being and improvements in well-being tends to hold, we can (partly) infer a person's level of well-being from their changes in well-being. That is, we can assume that people with a high level of well-being has a high level of progress over time. This progress will tend to be indicated by a high level of positive affective states over that time.

The problem with this method, however, is that it ignores the fact that a person who is badly off may be able to improve their well-being more than someone who is well off. Simply put, it is possible that the worse off a person

is, the more they will be able to improve their well-being – there are simply more improvements for them to make.

Where we stand on this issue is largely determined by whether levels of well-being are relatively finite or infinite. Consider the following analogy with knowledge. If there is an infinite amount of knowledge then we might think that attaining a greater amount of knowledge simply opens up the possibility of attaining even more knowledge. In contrast, if there is a relatively finite amount of knowledge then we might think that attaining a greater amount of knowledge reduces one's chances of attaining much more knowledge. A similar situation relates to people's well-being. If well-being is relatively finite, it may be that peasants have a greater chance of making improvements in their well-being than millionaires do (for, how much more can millionaires do to make their lives better?). Alternatively, if well-being is infinite, it may be that millionaires are more likely to make improvements in their well-being than peasants are (for, just think of the ways in which millionaires could improve their lives).

I am not sure where to stand on this issue, and it seems that there are reasonable arguments for both positions. Unfortunately, appealing to empirical evidence does not help much either. Recall, from chapter four, studies that show that, on average, rich countries are happier than poor countries, but that this trend is negligible after average incomes of \$10,000 (Graham & Pettinato, 2002). This suggests that people in poor countries are less able to consistently improve their well-being, and thereby be happy, than people in rich countries (i.e. countries with average incomes of \$10,000 or higher). However, in contrast to this evidence is the existence of “happy peasants” – individuals who are happy (and presumably, therefore, consistently improving their lives) despite having low levels of well-being.

In sum, it is not obvious whether someone's level of well-being tends to result in them experiencing happiness or unhappiness over time. This does not mean to say, however, that a person's level of well-being will tend to have no effect on their pattern of local changes in well-being experienced over time. In the next section, I will consider one way in which a person's level of well-being may influence their happiness over time in this way.

## **B.2 The Steady Progress Method**

In this section, I will focus on one potential difference between people with high and low levels of well-being, which draws on the work of Jonathan Wolff and Avner De-Shalit (2007). Wolff and De-Shalit are concerned with the topic

of disadvantage. They explain why disadvantage concerns not only a lack of opportunity and low functioning (i.e. the failure to meet basic needs in areas such as being able to live, to enjoy bodily health, and to have attachments) but also increased risks and insecurity about future functioning. Certain resources and capabilities are largely valuable, according to Wolff and De-Shalit, because they help us attain *secure* functionings (i.e., secure beings and doings). For example, having the opportunity to irrigate one's land enables a subsistence farmer to grow food and remain well nourished even in times of drought. Without this capability, the farmer may have been able to grow food and attain nourishment, but these functionings would not have been secure. In times of drought, these functionings would have been lost. Thus, Wolff and De-Shalit claim that a person is largely disadvantaged insofar as they do not have the capabilities that enable them to attain secure functionings.

The important point about Wolff and De-Shalit's analysis of disadvantage for our purposes here is that certain resources and capabilities do not merely enable us to make progress in our well-being. In addition, they enable us to make *secure* well-being improvements. Thus, even if it is the case that people with high and low levels of well-being make the same amount of improvements in well-being, people with high levels of well-being may tend to make more secure improvements than people with low levels of well-being.

It is worth briefly reviewing the two different kinds of ways, as outlined by Wolff and De-Shalit, in which improvements in well-being can fail to be secure. Most straightforwardly, certain kinds of improvements in well-being can be risky in themselves. For example, even if people in America and Ethiopia are both able to eat, being well nourished in Ethiopia is far less secure (due to drought, famine, etc.) than in America. Perhaps more problematic, however, are the kinds of improvements of well-being that are at risk due to their relations with other aspects of one's well-being. Wolff and De-Shalit categorize these kinds of risks into (a) Cross Category Risks, and (b) Inverse Cross Category Risks. Cross Category Risks concern functionings that *depend* on other aspects of one's well-being. For example, anyone relying on their income to buy food will find that risks to employment generate risks to nutrition among other things. In contrast, Inverse Cross Category Risks concern functionings that are not dependent, but are *affected*, by other aspects of one's well-being. For example, steps taken to secure nutrition may have the effect of putting other aspects of one's well-being at risk, such as life and bodily health. People may put their lives in grave danger

to secure food for themselves and their families. Inverse Cross Category Risks, therefore, show that a lack of certain capabilities can have far ranging effects on the security of improvements in well-being. Not being able to attain nourishment, for example, may threaten many, if not *all*, of a person's improvements in well-being. Wolff and De-Shalit illustrate this kind of effect with the following example:

“It is reported that, due to an economic crisis in Israel, during the year 2002–3, one in five families in Israel became unable to guarantee proper nutrition for their children. The report uses the language of ‘insecure capability to have proper nutrition’. This does not necessarily entail hunger...but it means that these families bought cheap food rich in carbohydrates (e.g. rice, pasta) rather than healthier, but more expensive food rich in proteins, iron, and calcium. Among these poorer populations the average daily consumption of calcium was 55 per cent, and protein 65 per cent, of those who could afford proper nutrition. Moreover, among those families whose diet contains higher proportions of carbohydrates, the prevalence of diabetes is more than twice as frequent than in families that enjoy proper nutrition. What is even more striking is that according to this report, in order to try and secure their functioning of nutrition, 24 per cent of families in Israel gave up other ordinary basic needs such as medical supplies and continuous electricity and many of them choose not to pay their mortgage debts. In other words securing the functioning which came under threat (i.e. proper nutrition) made other functionings (e.g. to enjoy warmth in the winter) insecure.” (Wolff & De-Shalit, 2007: p.72)

The different kinds of risks outlined by Wolff and De-Shalit all tend to have the same upshot: People with *low levels* of well-being will have *less secure improvements* in well-being. For example, a peasant may be quickly climbing out of poverty only to be struck by illness, which lands him back in deprivation. In contrast, a millionaire is not likely to lose their wealth, even if they suffer from a chronic illness. The millionaire's improvements in well-being are likely to be steady, whereas the peasant's improvements are not.

This suggests that we may be able to infer a person's level of well-being from their *level of variance* in affective states over time. Someone with a high level of well-being may not tend to experience either a greater or large amount of positive affective states over time than someone with a low level of well-being. However, someone with a high level of well-being may tend to experience a *steady* increase in their well-being over time, and thereby a low level of

variance in their affective states over that time. In contrast, someone with a low level of well-being is more likely to suffer *set-backs* in improving their well-being. Such set-backs are unlikely to be expected and thereby will tend to cause negative affective states. This will tend to result in a high level of variance in their affective states over that time. That is, someone with low levels of well-being may tend to consistently improve their well-being, but such progress is likely to be more precarious than steady.

I believe that this method is plausible, and is well supported by the evidence presented by Wolff and De-Shalit (2007). The idea is that we may be able to (partly) infer a person's level of well-being from their level of variance in their local changes in well-being (indicated by happiness). If a person's happiness over time has a high level of variance, we may be able to infer that they are badly off. Conversely, if a person's happiness over time has a low level of variance, we may be able to infer that they are well off. Of course, this method can be tested. We can observe whether a low level of variance in happiness correlates with having central capabilities and functionings, for example. I have hoped to show in this section that we have reason to believe in the value of such research.