

The Value of Truth

Josef Holden

A thesis submitted to Victoria University of Wellington
in fulfilment of the requirements for the degree of
Master of Arts in Philosophy



————— 2015 —————

Abstract

People often want truth, and it often seems worth wanting. This has led many to claim that truth is *valuable* (VT). This essay argues that there are good reasons to reject VT.

After dealing with preliminary issues, Chapter 1 discusses the *instrumental* value of truth. I argue that, though some instrumentally valuable beliefs are true, there is little reason to think that these beliefs are valuable *because* they are true. Chapter 2 and 3 are concerned with the claim that truth is *intrinsically* valuable (CVT). Chapter 2 examines a serious difficult facing CVT. This is the problem of trivial truths. Though this problem is often discussed, its power is rarely appreciated. I argue that the two most prominent responses to the problem fail. Chapter 3 poses a different question: Do people *in fact* consider truth to be valuable? A few notable exceptions aside, it is generally accepted that they do. Further, a number of arguments for CVT rest on this assumption. I argue it is very doubtful that people value truth intrinsically. If this is correct, the arguments for CVT that rest on this claim collapse.

Acknowledgements

I would not have been able to complete this work without the assistance and support of a number of people.

My first thanks must go to my supervisor Stuart Brock for his guidance and support. I greatly appreciate the time and care he took commenting on and discussing various versions of this work. My thesis has benefited greatly from his advice.

Further thanks are due to the Victoria University Philosophy department and, in particular, the postgraduate community. As well as providing valuable feedback, the process would have been far more difficult without Friday night drinks to look forward to. I would like to single out Matthew Macdonald and Matthew O’Leary. You could not ask for better office mates.

Finally, I would like to thank Amelia Jeffares. As well as providing invaluable editorial advice – and reading over the entire draft – she has been a constant source of support. I could not have done it without her.

Table of Contents

Abstract	i
Acknowledgements	ii
Table Of Contents.....	iii
Chapter 1: Introduction, Clarifications, And The Instrumental Value Of Truth 1	
Introduction.....	1
Clarifications.....	7
On The Instrumental Value Of Truth.....	16
Concluding Remarks.....	34
Chapter 2: The Problem Of Trivial Truths And Two Unsuccessful Solutions ... 36	
Introduction.....	36
Advantages Of CVT And The Costs Of Denial.....	40
A Potential Concern About CVT	41
The Problem Of Trivial Truths	42
Truth As A <i>Prima Facie</i> Value.....	44
Against The PFA	47
Restricted Versions Of CVT.....	55
Against RVT	59
Concluding Remarks On The Problem Of Trivial Truths And CVT.....	76
Chapter 3: Do We Value Truth For Its Own Sake?..... 77	
Introduction.....	77
PC And Its Uses	77
My Argument And A Concern	86
First Negative Argument: Tempting Examples That Prove Nothing	88
Second Negative Argument: A Failed General Argument	90
Against PC	110
Final Thoughts On PC And Its Relationship To CVT.....	120
Conclusion	123
References.....	125

Chapter 1: Introduction, Clarifications, and the Instrumental Value of Truth

Introduction

In the *Laws*, Plato (1973, 730c) claims that ‘Of all things good, truth holds first place among the gods and men alike.’ The idea that truth is *valuable* continues to hold considerable sway. I shall refer to this claim, in its generic form, as:

VT: Truth is valuable.¹

VT is ambiguous in various ways.² It is particularly significant for our purposes that VT does not designate the *kind* of value being attributed to truth. Two possibilities will be considered in this essay. First, it could be claimed that truth is *intrinsically* valuable. According to this interpretation, truth is worth having in itself, or for its own sake. Call this claim:

CVT: Truth is intrinsically valuable.

It could also be claimed that truth is *instrumentally* valuable. On this interpretation, truth is valuable in virtue of leading to other things of value. Call this claim:

LVT: Truth is instrumentally valuable.³

¹ In calling this claim ‘VT’, I follow Horwich (2006, 347).

² Further clarification of VT, and other concepts employed in this essay, can be found in ‘Clarifications’ below.

³ These acronyms are not ideal. Unfortunately, both ‘intrinsic’ and ‘instrumental’ start with an ‘I’, so the most obvious options are unavailable. To deal with this, I have used the last letter of each kind of value.

This essay argues that there are good reasons to be sceptical of LVT, and that CVT should be rejected. As this is a broad topic, it is necessary to focus the discussion in various ways. Unavoidably, this means that interesting questions will not be considered. Such omissions are hopefully compensated for by the benefits of dealing with particular issues in greater depth.

The essay is composed of three chapters. The core arguments are found in chapters 2 and 3. Both are concerned with CVT. Chapter 2 argues that CVT is implausible. To make this manageable, I concentrate on one of the strongest considerations against CVT. This is the *problem of trivial truths*. I am far from the first to discuss this issue. My central ambition is to demonstrate that two prominent responses to the problem fail. If this is correct, then CVT still has a problem – and a serious one at that.

Chapter 3 turns to a different issue: Do we in fact *value* truth for its own sake? Like CVT, the claim that we *do* has a long history. Aristotle (1958, I 1), for instance, opens the *Metaphysics* with the famous line ‘All men by nature desire to know’. Like CVT, this claim also continues to be widely endorsed. As this is a psychological claim – it is about what we *value*, not what is *valuable*⁴ – call it:

PC: People in fact, or actually do, value truth for its own sake.

PC is worth examining for two reasons. Most simply, it is an interesting claim in its own right. More importantly for our purposes, PC is often taken to support CVT. I argue that PC should be rejected and that, as a result, a number of

⁴ This distinction is discussed further below. See page 11.

common arguments for CVT fail. Further, I argue that rejecting PC erodes much of the motivation for endorsing CVT.

Chapter 2 and 3 are largely independent of one another. Certain issues and concepts do recur in both, however. The most obvious example is the problem of trivial truths. This is more widely recognised as a problem for CVT, but I argue that it also undermines PC. This reflects my general belief that this problem leaves little in its wake. Others have agreed with this assessment, but it is not the dominant view. Most who address the problem assume that it can be avoided with minor theoretical adjustments. At minimum, I hope this essay demonstrates that this is not the case.

Though the core arguments are found in chapter 2 and 3, this chapter is more than a mere introduction. It has two additional aims. First, my use of certain concepts will be clarified. This will save space later and allow my arguments to flow more smoothly. The second aim of this chapter is more substantial. Aside from intrinsic value, VT is often defended by appealing to truth's *instrumental* value. I argue that there are strong reasons to be sceptical of LVT. Besides being of independent interest, this argument also serves as a non-pragmatic justification for focusing primarily on CVT.

Brief Survey

A final introductory task is worth undertaking. This is to motivate my claim about the popularity of VT. This serves to curb any concerns that I am just beating a dead horse. My survey is by no means exhaustive. As well as seeming unnecessary, such a survey would be redundant. VT's popularity, and particularly

its contemporary popularity, will be made apparent throughout the essay. My hope here is just to give a clear sense of the *perpetual* popularity of VT.

Note first that VT did not begin with Plato. The pre-Socratic philosopher Heraclitus (1979, 112), for example, claims that ‘Thinking well is the greatest excellence and wisdom: to act and speak what is true, perceiving things according to their nature.’ Though it did not originate with Plato, few philosophers have advanced VT with as much force, passion, or fanaticism. This is evident in Plato’s claim, in the above quote, that truth is more valuable than anything else. To cite just one additional example, Plato’s Socrates (1973, 526d-e) says: ‘I renounce the honours sought by most men, and pursuing the truth, I shall endeavour both to live and, when death comes, to die, as good a man as I possibly can be. And I exhort all other men ... to enlist in this contest which I maintain excels all other contests.’

Moving forward, strong endorsements of VT can also be found in Christian and Medieval philosophy. To take a paradigmatic example, Augustine (1887, XX.41) writes: ‘If anyone should propose to himself to love truth, not only truth which is contemplation but likewise that which is in true propositions... I know not whether he could be said wisely to err in anything.’ Falsity, on the other hand, is ‘hideous and repulsive just in proportion as it appears fair and plausible when we utter or assent to it’ (2009, X.23).

Turning to Modern philosophy, Francis Bacon (1986, 62) writes: ‘But howsoever these things are thus in men’s depraved judgments and affections, yet truth, which doth only judge itself, teacheth that the inquiry of truth, which is the love-making or wooing of it, the knowledge of truth, which is the presence of it, and the belief

of truth, which is the enjoying of it, is the sovereign good of human nature’.

Descartes (1996, 42), often considered the father of Modern philosophy, justifies his method by claiming that it will allow him to ‘avoid ever going wrong’ and ‘to arrive at the truth’. John Stuart Mill’s influential defence of free-expression clearly presupposes VT. A central component of the argument, after all, is that censorship is bad because it stops the truth getting out. Mill (1982, 41-2) writes:

First, if any opinion is compelled to silence, that opinion may, for aught we can certainly know, be true.... Secondly, though the silenced opinion be an error, it may, and very commonly does, contain a portion of truth; and since the general or prevailing opinion on any subject is rarely or never the whole truth, it is only by the collision of adverse opinions that the remainder of the truth has any chance of being supplied.

Frege famously claimed that truth is the aim of all sciences, and considered truth to be in the same evaluative class as *good* and *beautiful*. He (1967: 17) writes: ‘The word “true” indicates the aim of logic as does “beautiful” that of aesthetics or “good” that of ethics.... To discover truth is the task of all sciences; it falls to logic to discern the laws of truth.’

Many others have – in a similar spirit to Frege – conceived of truth as the goal of *philosophy*. As Lloyd Strickland (2013, 1079-81) writes:

Philosophy, as it is understood and practiced in the West, is and has been generally considered to be the search for truth. Throughout the long history of the discipline some of its most celebrated practitioners have explicitly described philosophy this way, e.g. Aristotle (1984, II,

1570), Spinoza (2007, 184) and Berkeley (2008, 68), while others have elected to characterize it as the search for knowledge or wisdom, where both ‘knowledge’ and ‘wisdom’ are synonyms for ‘truth’, or certain kinds of truth at least, e.g. Hobbes (1839, I, 3) and Descartes (1985, I, 186)... [S]o popular has this conception been, and so popular is it still, that it might appropriately be termed the Standard Conception of Philosophy.

Taking truth as a goal or seeking truth does not entail *valuing* truth. Nonetheless, it strongly suggests that these philosophers would endorse VT. It would be somewhat odd, and at least mildly perverse, to self-consciously commit your life to seeking something you believe has no value. For most of us at least, this would be like choosing to commit our lives to playing snakes and ladders, or to collecting strands of hair.⁵

VT persists to the present day. Paul Horwich (2006, 347), in *The Value of Truth*, follows Frege and claims: ‘It is generally taken for granted that truth (along with, for example, justice and beauty) is a good thing, and I have no quarrel with this assumption. The aims of the present paper are to justify it and defend it, to try to explain why it is correct...’ Michael Lynch (2004, 13), in *True to Life: Why Truth Matters*, writes: ‘In believing, we operate under the norm of truth: *Other things being equal, it is good to believe a proposition when and only when it is true...*

⁵ Many of us probably accidentally end up living such lives. What is odd is *choosing* to do so. Further, even if truth is not valuable, philosophy may still be valuable for other reasons. It is the combination of believing that truth is the goal of philosophy and that truth has no value, and then still choosing to spend your life doing philosophy that is strange. It is strange enough, I think, that we can assume that at least most philosophers who think that truth is the goal of philosophy also believe that truth is valuable.

[Not] only is truth a norm of belief, true belief is a value; it is good, other things being equal, to believe all and only what is true, or the truth as such'. Finally, Jonathan Kvanvig (2003, 40-1), in *The Value of Knowledge and the Pursuit of Understanding*, endorses CVT when he writes:

[T]he value of truth is ... intrinsic to truth itself. Perhaps truth is remote in a certain sense, so that attaining it is too much to ask and we will have to settle for something less. Still... finding the truth remains the ideal and is valuable in itself. We might settle for empirical adequacy because we have no route to truth other than through empirical adequacy, but what we are after, and legitimately so, is the truth.

This is not an exhaustive survey, but it gives some indication of the prevalence – both historical and contemporary – of VT. In addition, these quotes illustrate that there is something *attractive* about VT. Most possible claims regarding what has value are never live options – the value of concrete, for instance – but the value of truth, like justice and beauty, has both a deep appeal and significant antecedent plausibility. It is, and has always been, a live option. Whatever the source of this attraction, and whether or not it can be justified, the attraction itself is beyond doubt. My purpose, of course, is to argue that it *cannot* be justified.

Clarifications

I now turn to clarifying my use of certain terms that recur in this essay. I begin with a stylistic point. This concerns my use of the words 'truth' and 'true'.

As may be clear, I am concerned with the value of true *belief*. This is the standard concern. The reason for this is simply that true belief is more plausibly valuable

than anything else truth-related. Consider true propositions, for instance. There is a fact of the matter right now about how the universe began – some proposition about this *is* true. But it hardly seems plausible that just *this bare fact* – that there is a truth about this – is valuable. What people seem to value – and what may be valuable – is possessing this truth. That is, *believing* this truth. Sosa (2001, 49) expresses this point well:

Even if intelligent life had never evolved or otherwise existed, Venus would still have orbited the sun, so it would still have been true that Venus orbited the Sun. It is not the being thus true of what is true that we value.... In loving the truth, then, what we value in not the being true of the truths. What we value in pursuing truth is rather our grasping it, our having it. What does this mean? Only through believing it does one relevantly have a truth: We have the truth that snow is white by believing that snow is white. In pursuing the truth what we want is (at least) true beliefs.

For ease of expression, I sometimes say the value of *truth*. This should be read as shorthand for the value of *true belief*.

Truth, Belief, and Value

Given the concerns of this essay, three concepts that call for clarification are ‘truth’, ‘belief’, and ‘value’. As it would take at least a thesis to give useful or illuminating explanations of these concepts, I shall not attempt to. Aside from some brief remarks below, I primarily follow the standard approach when addressing VT by leaving these concepts intuitive and taking them as primitives.

I assume that truth involves – though is not necessarily exhausted by – something like the equivalence schema: p is true if p . It is true that snow is white if snow is white. Though truth is undoubtedly complex, it seems clear to me that this, or at least something like it, is part of the intuitive concept of truth. If you reject this assumption, not much of this essay will survive the disagreement. If it *should* be rejected, then most writing on this topic may as well be thrown in the trash.⁶

As with truth, my understanding of the concept *belief* is intuitive. It lines up well with Schwitzgebel's (2014, 1) characterization:

Contemporary analytic philosophers... generally use the term 'belief' to refer to the attitude we have, roughly, whenever we take something to be the case or regard it as true. To believe something, in this sense, needn't involve actively reflecting on it: Of the vast number of things ordinary adults believe, only a few can be at the fore of the mind at any single time.... Many of the things we believe, in the relevant sense, are quite mundane: that we have heads, that it's the 21st century, that a coffee mug is on the desk.

In short, then, I assume that to believe p is, roughly, to take p to be true. And that p is true, roughly, if and only if p . I do *not* assume that these are the best accounts of these concepts. My ambition is *only* to use 'truth' and 'belief' in an intuitive sense, not to give any sort of analysis of these concepts.

My use of 'value' and 'valuable' is slightly more delicate. I am less confident that I have a clear idea of what the intuitive conception of value is.

⁶ See Glanzberg (2014) for a good overview of debates concerning the nature of truth.

Further, certain claims that I make would be rejected by defenders of widely held – and plausible – theories.

For example, some of my arguments assume that value is objective – at least in the minimal sense that a person’s judgement about what is intrinsically valuable can be *wrong*, and that what makes it wrong is something *external* to this person. It is not merely a mistake about her psychology. Though this assumption is compatible with many theories of value, it would be rejected by others. This includes theories that claim that intrinsic value depends entirely on second-order attitudes. Harry Frankfurt (2004) defends a theory of this kind. On Frankfurt’s theory, the only way I can be wrong about what is intrinsically valuable is if I am wrong about my own attitudes. In this case, wrong about what I value or care about. Frankfurt (2004, 26) writes: ‘the most basic and essential question for a person to raise concerning the conduct of his life cannot be the *normative* question of how he *should* live. That question can sensibly be asked only on the basis of a prior answer to the *factual* question of what he actually *does* care about.’ To deal with such disagreements, I attempt to signal points where alternative theories of value clearly come apart, and indicate where this would affect my arguments.

It is important to emphasise that many arguments apply no matter which theory of value is correct. Chapter 3, for example, asks whether we *value*, or *care about*, true belief for its own sake. This question can be asked without assuming anything about what makes something *valuable*. Even in the example just given, my assumption does not clash with all theories where psychological factors – such as desires – play a central role. According to ‘fitting attitude’ accounts of value, for instance, what is good is just what is

correctly or appropriately desired.⁷ Sidgwick (1907) claims, for example, that the good is what *ought* to be desired. My discussion, if read in this light, is just about whether true belief ought to be desired.

Valuing vs. Being Valuable

The previous remarks hint at a distinction that is worth making explicit. This is the distinction between a person *valuing* x and x actually being *valuable*. Though these often overlap, they are two separate concepts. As Brady (2009, 269) writes:

...it is a general truth in value theory that, although the fact that I do desire or care about something might incline us to think that that thing is worth desiring or caring about, it does not guarantee that it is. There is always the possibility that I desire or care about something that I ought not to desire or care about, that is, something that is not *worthy* of my concern. In other words, there is always the possibility that one of my ends or goals is not a *proper* end or goal. If so, we might think that the fact that I desire truth on a particular subject for its own sake does not guarantee that the truth on that subject is worth desiring, or is valuable as an end.

This essay largely assumes that there is indeed a gap between what a person *values* and what is *actually* valuable. The *mere* fact that a person values x does not entail that x is valuable. Though I do not defend this distinction in detail, it is highly intuitive and rarely denied.

⁷ For an enlightening discussion of ‘fitting attitude’ theories, and of Meta-Ethics generally, see Smith (2005).

The intuitive plausibility of the distinction is best illustrated with examples.

Consider, for instance, an extreme racist who places more value in the colour of a person's skin than anything else. A natural thing to say about this person is that they value something that is not valuable. That is, though they *regard* skin colour as valuable, they are *wrong* about this.

Conversely, it seems clear that a person can *fail* to value something that is *actually* valuable. Imagine, for instance, an exceedingly selfish teenager whose parents will do anything for him. They shower him with love and provide him with all the comforts one could ask for. Suppose that this child *only* values himself, and hence does *not* value his parents – or anything they do for him – at all. In this case, it seems natural to say that the teenager does *not* value something that is actually valuable.

Very few theories of value deny that this is a meaningful distinction. It is most pronounced on objectivist theories. According to these, what is valuable is determined by facts external to any individual. As such, we can be mistaken about what actually has value in a straightforward way: just as a person's belief that UFOs visit the earth is mistaken if UFOs don't actually visit earth, a person's values are mistaken if what they value is not actually valuable.

Most subjectivist theories also make room for this distinction. This includes theories that claim that what is actually valuable is entirely contingent on psychological facts. Consider David Lewis's (2000) theory, which argues that what is valuable is what we are disposed, under ideal conditions, to value. As he (68) writes:

Roughly, values are what we are disposed to value. Less roughly, we have this schematic definition: Something of the appropriate category is a value if and only if we would be disposed, under ideal conditions, to value it.

This means that, if I am not under ideal conditions, I may well value something that is not actually valuable. As Lewis (71) puts it, his theory ‘allows, as it should, that under less-than-ideal conditions we may value what is not really good.’

It is worth noting that this essay does not make any substantial assumptions about what it means to value something. I only assume that valuing is some kind of pro-attitude towards x . Perhaps valuing x is believing that x is valuable, or it could be desiring x , or desiring to desire x , or some kind of sentiment towards x . All of these theories are plausible, but an intuitive conception of what it means to value x is sufficient for our purposes. My claims about valuing should be compatible with any of the above theories.

The distinction between valuing and being valuable is important to make clear, as the views I discuss make different kinds of claims. LVT and CVT are claims about truth’s *actual* value. LVT claims that truth is instrumentally valuable, and CVT claims that truth is intrinsically valuable. PC, on the other hand, is a claim about what we *in fact* value. In itself, it takes no stand on what is actually valuable, nor on what makes something actually valuable. Its claim is just that – whether or not it is valuable – we in fact value true belief.

Intrinsic vs. Instrumental Value

For reasons defended below, this essay focuses primarily on the *intrinsic*, rather than the *instrumental*, value of true belief. It will be useful to get clearer on this distinction.

To say that x is *instrumentally* valuable is to say that x is valuable in virtue of leading to something else of value. To say that true belief is instrumentally valuable, then, is to say that we should believe true propositions because this is a means to valuable ends. It should be stressed that something is only instrumentally valuable only if it leads to something that is *legitimately* intrinsically valuable. It is a fact that drinking large quantities of antifreeze is an efficient means to the end of dying a horrible, painful death. But it seems wrong to claim, on the basis of this, that drinking large quantities of antifreeze is instrumentally valuable. The reason for this is simply that dying a horrible, painful death is not valuable.

Instrumental value can be distinguished from *intrinsic* value. Something is intrinsically valuable if it is valuable in itself or for its own sake. It is an end that instrumentally valuable things might lead to. To claim that true belief is intrinsically valuable, then, is to claim that we should believe, or that it is good to believe, true propositions *just because* they are true.

A simple example will help illustrate this distinction. The paradigmatic example of something that is instrumentally – but not intrinsically – valuable is money. Money is good because it leads to acquiring other good things. It allows you to eat nice food, live in a nice house, and provides access to opportunities and experiences that would not be otherwise available. These things, in turn, may only

be instrumentally good. The opportunities that money provides access to – education, for instance – may only be good because undertaking these opportunities is necessary in order to get something else good – a certain career, for example. This in turn may only be good for what it leads to, but eventually things of instrumental value must, for fear of infinite regress, lead to something that is *good tout court* – say, happiness. Such things are intrinsically good, or good for their own sake.

The above remarks illustrate, rather than defend, my understanding of intrinsic and instrumental value in this essay. There are some difficult issues in these areas. This is particularly true of intrinsic value, which is the subject of various disputes.⁸ I am employing this concept in what Grimm (2009, 246) calls the ‘standard way’, ‘to mean a value that is worth pursuing and realizing for its own sake’. Lynch (2004, 128) puts his understanding the same way. He writes that being an intrinsic good ‘makes something worth caring about for its own sake, as opposed to caring about it for what it leads to’. Similarly, Stich (1990, 93) writes: ‘I’ll begin by rehearsing some conventional wisdom. People value many things. In some cases the value is instrumental, in the sense that the thing valued is valued because it is believed to facilitate the achievement of other goals. Money is the standard example here. In other cases the thing valued is valued *intrinsically*, or “for its own sake”.’

At minimum, it seems fair to say that I employ this concept in the way it is standardly used in the literature on VT. But there is reason to think this

⁸ See Zimmerman (2015) for a good discussion of these disputes.

conception is standard in a wider sense. Zimmerman (2015, 1), in his overview of this topic, writes:

Intrinsic value has traditionally been thought to lie at the heart of ethics. Philosophers use a number of terms to refer to such value. The intrinsic value of something is said to be the value that that thing has “in itself,” or “for its own sake”, or “as such”, or “in its own right”.

As Zimmerman’s remark implies, as well as being the standard conception of intrinsic value, my usage is also the *traditional* conception of intrinsic value.

In short, when I ask whether truth is intrinsically valuable, I am asking whether truth has value in itself, or for its own sake. And when I ask whether truth is instrumentally valuable, I am asking whether believing true propositions leads to other things of value *in virtue of those propositions being true*.

On the Instrumental Value of Truth

I now turn to the substantive task of this chapter. This is to argue that the idea that true belief is instrumentally valuable – at least in any interesting sense – faces serious difficulties. By ‘interesting sense’, I mean any sense that explains or justifies why, if we want to achieve our goals, we should prefer our beliefs to be true *rather than false*. I argue that instrumental value is unlikely to explain why truth is something especially, or specifically, worth having.

My defence of this claim is relatively brief. Nonetheless, it is worth offering for two reasons. The first is that it provides a non-pragmatic justification for primarily focusing on intrinsic value. Though pragmatic considerations may be sufficient, a stronger justification is preferable. One basis for this claim is that, all else equal, it

is better to focus on the best form of a claim. These considerations suggest that LVT is not the best form VT can take.

More importantly, this conclusion gives my arguments against CVT additional *strength* or *urgency*. If instrumental value is left entirely untouched – and continues to be seen as a compelling option for grounding truth’s value – then conceding CVT may seem inconsequential. Put simply, bracketing instrumental value entirely provides *too much comfort* to defenders of VT. It leaves open the possibility that truth’s instrumental and intrinsic value are on a par, and that instrumental value alone is sufficient to get everything the defender of VT desires. My argument suggests this is not the case.

As they are discussed relatively briefly, the considerations in favour of my conclusion are not conclusive. Replies and counter-replies are no doubt available. Nonetheless, they strike me as *strong* considerations. They are, in any case, more than sufficient to motivate focusing primarily on CVT.

The Instrumental Value of Truth

According to LVT, true belief is valuable because it leads to other things of value. If I believe truly, I am more likely to achieve my goals. This claim is easy to motivate. No matter what has intrinsic value or why, it seems that truth will help us attain it. Suppose that pleasure is intrinsically valuable. It seems clear that it will greatly increase the amount of pleasure a person experiences if she has true beliefs about what will *bring* her pleasure.

Endorsements of truth’s instrumental value are common. Lynch (2004, 16), for example, writes:

Indeed, the most obvious reason to pursue true beliefs is that believing the truth can get us all sorts of other things we want. Believing the truth is practically advantageous.... Believing the truth is valuable because it is a means to other ends – sturdy bridges, cures for diseases, and safety. We can sum this up by saying that truth is instrumentally good.

Horwich (2006, 350) writes: ‘true belief is valuable because it *pays*: – it has evident practical benefits; we are more likely to get what we want if we base our deliberations and actions on true beliefs than if we base them on false ones’.

David (2001, 155) says that ‘we usually desire it [truth] because we think that having true beliefs will increase our chances of satisfying our other desires.’

Finally, Zagzebski (2003, 22) asserts that ‘True belief is surely a means to reaching our ends’.

Additionally, truth’s instrumental value is regularly taken for granted. This can be seen, for example, in the literature on the wrongness of lying and deception. It is often argued that one thing that makes these acts wrong is that they undermines people’s access to the truth, and that *this* is wrong because it undermines their pursuit of valuable ends. This claim clearly rests on the assumption that true belief is instrumentally valuable. Thomas Carson (2010, 89), to cite just one example, writes: ‘We are generally harmed when we are deceived because we cannot effectively pursue our ends and interests if we act on the basis of false beliefs.’

Before arguing against the instrumental value of truth, two preliminary points are in order. These are not controversial, but are important to keep in mind. For something to be instrumentally valuable, it must be *causally* relevant to the

accomplishment of a good end.⁹ The basic idea is this: x is instrumentally good if it is good (efficient, etc.) as a means to an end, and being a means to an end requires playing a causal role in bringing that end about. It is not sufficient for x to simply be *present* or, as Wrenn (2010) puts it, a ‘reliable symptom of a means’. Suppose that, in order to get your dream job, you need to read a series of books and then pass a test based on their contents. This job will bring with it everything you have ever desired. It seems clear that reading these books is instrumentally valuable. Now suppose that these books are all *blue*. Though it happens to be correct that reading these blue books will be good for you, ‘blueness’ is not part of the books instrumental value. This can be seen by the fact that, if you found copies in a different colour, they would be just as good. The causally relevant feature of the blue books is their content, not their colour. ‘Blueness’ is merely a ‘reliable symptom of means’, not causally relevant to you achieving your ends and hence not, itself, instrumentally valuable.

This raises a second point: To reject the claim that truth is instrumentally valuable, it is not necessary to argue – and I doubt anybody would argue – that no instrumentally valuable beliefs are true. This point alone – that some true beliefs are instrumentally valuable – is insufficient. For truth to be an instrumentally valuable property of beliefs, it must be the case that the success of true beliefs is *explained* by their truth. In other words, a belief has to be instrumentally valuable *because* it is true. Truth cannot be like blueness in the above example. It cannot merely be a feature that some instrumentally valuable beliefs happen to have. What is needed, then, is a general reason to think that there is some kind of causal

⁹ See Wrenn (2010) for an enlightening discussion of this requirement, and instrumental value generally. My example is inspired by a case that he discusses.

connection between true beliefs and instrumentally valuable beliefs. Only this would warrant the claim that, if we want to achieve valuable ends, we should generally want true beliefs. I argue that there are two reasons to doubt there is any such causal connection.

First Reason for Doubt

The first reason for scepticism is simply that there are innumerable counter-examples to the claim that true beliefs are instrumentally valuable. The truth is often bad for us, and believing truly often undermines our attempts to achieve valuable ends. This gives us reason to be suspicious of the idea that there is any strong, or tight, connection between true beliefs and instrumentally valuable beliefs. It shows, for example, that there are various instances where you should not try to believe truly if you want to achieve your goals.

Of course, one could always respond that, even if there are examples where truth is not instrumentally valuable, truth is still *generally* instrumentally valuable, and *generally* instrumentally preferable to falsehood. I argue below that this is not the case, but it is probably impossible to answer this charge with examples alone. It is also probably impossible to *defend* this claim with examples alone. At minimum, the fact that there are countless counter-examples to this claim gives us *good grounds* to doubt that there is any interesting causal connection between true beliefs and instrumentally valuable beliefs. It also demonstrates that the instrumental value of truth is not as 'evident' or 'obvious' as it is often claimed to be.

The most discussed examples of this kind are *overconfidence cases*. We are sometimes served better – we have a better chance of achieving our goals – if we

have false beliefs about our abilities rather than true beliefs. Pritchard (2007, 102) gives the following example: '[Suppose] one is unable to summon the courage to jump a ravine and thereby get to safety, because one knows that there is a serious possibility that one might fail to reach the other side. In such cases it seems that a false belief in one's abilities would be better than a true belief, if the goal in question (jumping the ravine) is to be achieved.'¹⁰

Another compelling case is suggested by Stich (1990). One end that many people agree is intrinsically valuable is *survival*. If nothing else, survival gains value from being a precondition to all other things of value. You can't do much if you're dead. It is clear, however, that true beliefs can get you killed. Stich (1990, 122-23) provides a potent example of such a case:

Is true belief always more conducive to survival than false belief?

Clearly the answer is no. To see the point, we need only reflect on the plight of poor Harry who believed that his flight left at 7.45 a.m. He wrote it down, ordered a cab the night before, and asked his wife to be sure he was out of bed by 6.30. Harry's belief was true, and he got to the airport just in time. Unfortunately, the flight crashed, and Harry died. Had Harry falsely believed that the flight left at 8.45, he would have missed the flight and survived. So true belief is sometimes less conducive to survival than false belief.

As Stich (1990, 123) notes, this result will generalise broadly:

True beliefs are not always optimal in the pursuit of happiness or pleasure or desire satisfaction, nor are they always the best beliefs to

¹⁰ Cf. James (1956).

have if what we want is peace or power or love, or some weighted mix of all these.

Though both these cases are convincing, the most persuasive class of counter-examples, to my mind, are *terrible truths*. One reason for this is that it is the nature of the truth itself that causes terrible truths to be instrumentally disvaluable. This suggests that the instrumental disvalue of certain true beliefs is *explained* by their truth. Things being as they are, there are also many such cases. Truth is often an efficient means to misery.

The horrible nature of many truths, and the suffering they are apt to cause, has not gone unnoticed. Perhaps the philosopher who has doubted the value of truth on this basis most forcefully is Nietzsche. This is reflected succinctly in his statement that ‘The truth is terrible.’¹¹ That he doubts the value of truth – particularly its instrumental value – is shown in various statements he makes.¹² For example:

The falseness of a judgement is for us not necessarily an objection to a judgement; in this respect our new language may sound strangest. The question is to what extent it is life preserving, species-preserving, perhaps even species cultivating. (1966, 4)

And a large part of why Nietzsche thinks that truth is instrumentally disvaluable is *because* he thinks that it is terrible. For Nietzsche, as Leiter (2014, 9) writes, ‘Since being able to live depends on illusion, any moral imperative to know only

¹¹ Nietzsche makes this claim more than once throughout his writings. One example comes from his autobiography *Ecce Homo* (1992, IV: 1). As will become clear, I owe much of this section to Brian Leiter’s (2014) brilliant paper of the same name.

¹² For in-depth discussion of Nietzsche’s views about the value of truth, see Leiter (2014), Ridley (2010), Allen (1992), and Gemes (1992).

the truth is necessarily an obstacle to life.’ Nietzsche (1974, 107) makes a claim of this kind when he writes:

Had we not approved of the arts and invented this type of cult of the untrue... sensate existence would be utterly unbearable. *Honesty* would lead to nausea and suicide.

A similar – though much weaker – claim about the instrumental disutility of truth is made by David Lewis (2000). As noted above, Lewis is an advocate of a dispositional theory of value. On his view, what is valuable is what we are disposed, under ideal conditions, to value. Unlike many who advocate similar theories, Lewis does *not* include full information in his ideal conditions. And part of the reason for this is that the truth may sap us of motivation. Lewis (2000, 81) writes:

‘Another unhelpful sort of knowledge is a vivid awareness that we are small and the cosmos is large; or a vivid awareness of the mortality of mankind, and of the cosmos itself. If such knowledge tends to extinguish all desire, and therefore all valuing, it will not help us to value just what is valuable.’

Though other philosophers have made similar claims,¹³ I believe the most powerful writer on this subject is the horror writer H.P. Lovecraft. Let me note just two examples. Lovecraft (1999, 139) opens one of his most famous stories – *The Call of Cthulhu* – with the following lines:

¹³ One particularly interesting, if depressing, example is the somewhat obscure essay ‘The Last Messiah’ by the Norwegian philosopher Peter Zapffe (2004).

The most merciful thing in the world, I think, is the inability of the human mind to correlate all its contents. We live on a placid island of ignorance in the midst of black seas of infinity, and it was not meant that we should voyage far. The sciences, each straining in its own direction, have hitherto harmed us little; but some day the piecing together of dissociated knowledge will open up such terrifying vistas of reality, and of our frightful position therein, that we shall either go mad from the revelation or flee from the deadly light into the peace and safety of a new dark age.

It might be thought that as this passage comes from his fiction we should not attribute these views to Lovecraft. This is reasonable, but in this case the art accurately reflects the author. In a 1918 letter, for instance, Lovecraft (*quoted in Houellebecq 2008, 23*) writes that ‘all rationalism tends to minimize the value and the importance of life, and to decrease the sum total of human happiness. In some cases the truth may cause suicidal or nearly suicidal depression.’

Of course, quoting people asserting that the truth is horrible and instrumentally disvaluable doesn’t prove much. A few specific examples should help motivate the plausibility of this claim.

Start with the obvious – but horrible – truth that each of us is going to die. Really think about the fact that one day you will be taking your last breath. You may well have to actually experience this moment; no matter what, your death will happen. All of your life will have just led to that last moment and once you are there *that is it*. All your ambitions, all the good moments, bad moments, your whole life, will be behind you and soon vanish.

Not only will you vanish from the world literally, but, as Leiter (2014) emphasises, also figuratively. Beyond perhaps your grandchildren, the likelihood that anybody will be remembered is slim to none. And, as Leiter (2014, 2) writes, ‘even regarding that fake bit of “immortality” – being recognized for having existed at some point in the intangible past, by grandchildren – one might ask: how could that constitute adequate satisfaction for the insatiable vanities of human life...?’.

Even in the highly unlikely event that – as our bodies rot away – one of us ‘survives’ for a bit longer through one of our books being read or something similar, this provides little solace because, just like us, all these readers end up dead. Not only do all the readers die, eventually (supposing nothing else does it first) the Sun will explode and the entire Earth will be destroyed. All of culture—Plato, Shakespeare, Hume, Kant, etc. – and all of history – the wars, the poverty, the charity, the improvements, the political hopes, dreams and struggles, everything – will be gone without having made even a speck of difference to anything in any important or intelligible sense.

Even if one is not bothered by such thoughts, these are hardly the extent of the misery awaiting us. As Leiter (2014, 2) writes:

...before all of us, and all our loved ones – meaning our dear children and our beloved siblings and spouses, and our parents and grandparents – before all of these persons dear to our heart enter oblivion, all will endure, to varying degrees, physical and psychological misery, mishap, and suffering. We will watch some loved ones die before their time; we will watch others suffer

excruciating physical ailments and psychological torments that will diminish their well-being and often their humanity.

Of course, it is not only that these various afflictions happen, although that is bad enough. What makes it all the more unbearable is that they *just* happen. They have no purpose, sense, or reason.

Truths such as these seem *bad*, and it is highly unlikely they are instrumentally valuable. More likely, believing such truths will make us miserable and drain any motivation we have to achieve our goals. From an instrumental point of view, things would be better if they weren't true. Consider the truth that we are going to die and be forgotten no matter what we do. It seems that, if it has any effect, believing this truth will most likely make all one's goals and desires seem entirely pointless. It is far more difficult, for example, to work really hard on – let alone justify working on – artistic or intellectual projects when considering the undeniable fact that all your work will be forgotten and completely disappear. This will probably happen during your life, but will definitely happen after it.

Note that this claim is more than just an intuition on my part. Empirical support for the instrumental disvalue of truth can be found in studies exploring the *depressive realism hypothesis*. According to this theory, as Dobson and Franche (1989, 419) put it, 'depressed persons may be more accurate in their perceptions of various situations than are non-depressed people'. This hypothesis is controversial, but it seems fairly robust in at least certain areas. On the basis of their research regarding this theory, Taylor and Brown (1988) conclude that:

[P]ositively biased illusions may represent a necessary part of mental health, and that realistic perceptions of the self and of the world may

more likely be associated with low self-esteem, depression, and mental illness than are more distorted perceptions.’

If Taylor and Brown are right, this strongly suggests that true beliefs can contribute to suffering. Given the depressing nature of many truths this hypothesis strikes me, at least, as eminently plausible.¹⁴

Terrible truths – in addition to numerous other examples – give us good reason to doubt that the connection between true beliefs and instrumentally valuable beliefs is as strong or tight as it may seem on its face. These examples also demonstrate that the instrumental value of truth is not as ‘evident’ or ‘obvious’ as it is often claimed to be.

Second Reason for Doubt

A second reason to doubt that truth is instrumentally valuable – at least in any interesting sense – is that true beliefs do not seem instrumentally *preferable* to other kinds of belief. For this reason, there does not seem to be any reason to *seek out* true beliefs for practical purposes, or to be particularly concerned about whether our beliefs are true. As I argue below, this claim extends to false beliefs. If you want to achieve your goals, there is no reason to prefer true to false beliefs. Any general preference for truth is arbitrary from the instrumental point of view.

Both Stephen Stich (1990) and Jonathan Kvanvig (2003) have argued that the value of truth cannot be explained instrumentally because other kinds of belief are equally useful. Kvanvig, for instance, argues that *empirically adequate* beliefs will do just as well as truth for instrumental purposes. He (2003, 39) writes:

¹⁴ For a review of this theory, see Dobson and Franche (1989).

‘[E]mpirical adequacy seems to have precisely the same value as truth. An empirically adequate theory is one that will never be revealed to be false in virtue of some false prediction it makes about the course of experience, and so no matter what use we make of our beliefs, we will be just as well off employing empirically adequate beliefs as employing true ones.’

As Kvanvig (2003, 39) also points out, ‘There are other concepts in the neighbourhood of empirical adequacy that have the same implication.’

Though I agree with Kvanvig, I will argue that a stronger claim can be justified. Both Kvanvig and Stich employ concepts that fall short of falsity. This is unnecessary. We can get the same result with flat-out false beliefs.

Consider what I shall call the *symmetry thesis*. According to the symmetry thesis, for any true belief, a false belief exists that is just as good from an instrumental point of view. If this is correct, then – from the point of view of instrumental value – there is no reason to prefer a belief *because it is true*. To see why, we need to start with an observation. If two beliefs lead to the same result with equal efficiency, then they are equally instrumentally valuable. Put another way, if belief y and belief z lead to result x with equal efficiency, there is no instrumental reason to prefer y to z . The choice between them is arbitrary. Once this is recognised, it can be seen that anytime a true belief will lead to x a false belief will do just as well in leading to x . For this reason, there is no general instrumental reason to prefer a true belief regarding x than a false belief regarding x . To prefer a belief regarding x merely because it is true is entirely arbitrary from an instrumental point of view.

Though it is hard to definitively prove the symmetry thesis¹⁵ – and there may be some exceptions – seeing it in action will establish its plausibility.

First, consider a toy example. Suppose that you want to grow a really beautiful garden, and in order to do so you need a particular kind of compost. In this case, it may seem like the best belief, from an instrumental point of view, is something like ‘If I use compost x , then my garden will grow to be beautiful’. It seems undeniable that this is an instrumentally valuable belief. It is not, however, the *best* belief one could have. Certain false beliefs are *just as good*. Since instrumental value comes *purely* in virtue of leading to x , these false beliefs are just as instrumentally valuable as this true belief. Consider the following belief: ‘It is necessary to put compost x on my garden *and* hum as I do so to help my garden grow. If I don’t put compost x on my garden, or if I don’t hum, it will not grow to be beautiful’. This claim is false – since humming is unnecessary – but it is *equally* instrumentally valuable. It leads just as well to the same end – a beautiful garden. This false belief can be put more succinctly as follows: ‘My garden will grow to be beautiful if and only if I use compost x and hum’.

Though it is difficult to prove, I cannot see why the symmetry thesis would not generalise widely. It certainly extends beyond toy examples. Consider Lynch’s claim above that true beliefs are valuable because they help us cure diseases.

¹⁵ One possible proof would be this: Let p be true, and my belief that p be instrumentally valuable. Now consider another proposition q that is false but plays no role in my practical reasoning, nor is it a proposition I care much about (e.g., that my neighbour three doors down watered her plants yesterday). Presumably it doesn’t matter what p is, a belief that p is arguably no more valuable than a belief that $p \& q$, where $p \& q$ is a false proposition (in virtue of the second conjunct being false).

Though true beliefs would be adequate for this purpose, they are certainly not necessary. The symmetry thesis shows that false beliefs are just as good. A surgeon, for example, could have nothing but false beliefs and still be the best surgeon in the world. She could save as many lives as she would have if she had nothing but true beliefs. This is because she could have only ‘if and only if’ false beliefs like those above. For example, ‘Surgery x will be successful if and only if I make an incision just above the patient’s ribs (etc.), *and* hum as I do so’. As above, this belief is false – humming is unnecessary. The surgery, however, will be just as successful. This argument could be extended to any medical field. This shows that a doctor, from an instrumental point of view, has no non-arbitrary reason to prefer true beliefs *because* they are true to false beliefs. *All* our medical knowledge could be false without a single additional person dying as a result.

To give one more illustration, return to the example we started with – pleasure-seeking. In this case too, one could have nothing but false beliefs and attain as much pleasure as one could with true beliefs. True beliefs will probably do the trick, but not any more so than false beliefs of the form ‘If I want pleasure, then I need to do both A and B’, where B is false and inconsequential. Above, I said that it seems clear that it will greatly increase the amount of pleasure a person experiences if she has true beliefs about what will *bring* her pleasure. We can now see that this is incorrect. You could attain all the pleasure in the world with nothing but false beliefs about what will bring you pleasure. A pleasure-seeker, from an instrumental point of view, has no non-arbitrary reason to prefer true beliefs about how to gain pleasure to false beliefs.

Assuming that the symmetry thesis does generalise, no true belief is *ever* the best instrumental belief. There will always be a false belief that is just as good. For this reason, there is no non-arbitrary instrumental reason to prefer a true belief, merely because it is true, over a false belief, merely because it is false. Seeking truth for practical purposes is no more rational than seeking falsity.

Two Objections

It might be objected that, even if it is correct, the symmetry thesis proves little. It is still the case that true beliefs *are* instrumentally valuable. The pleasure-seeker would still do well with all true beliefs, for example. The belief that x will bring me pleasure, if true, *will* serve the pleasure-seeker. It *is* instrumentally valuable.

This objection is very weak. This can be seen if we consider that, if the symmetry thesis is correct, everything the objector claims can be repeated about false beliefs. The pleasure-seeker, after all, would be just as well served with false beliefs. These false beliefs *are* instrumentally valuable.

A more promising objection might seem to be this: all my argument shows is that LVT needs to be *restricted*. It should not be claimed that true beliefs are instrumentally valuable no matter what. Rather, it should be claimed that only a certain class of true beliefs are instrumentally valuable. This would include, for example, means-end beliefs of the form ‘If I want x , then I should do A ’. It does seem to be the case that such beliefs will only be instrumentally valuable if they are true. Thus LVT, in an appropriately restricted form, is vindicated.

This objection seems correct as far as it goes.¹⁶ As above, however, this claim is incredibly weak. Again, if the symmetry thesis is correct, everything the objector asserts can be repeated about false beliefs. Consider a restricted version of the claim that false beliefs are instrumentally valuable (LVF). If we restrict LVF to include, for instance, all the false beliefs of the ‘if and only if’ variety discussed above then restricted LVF is as correct as restricted LVT.¹⁷ After all, such false beliefs are instrumentally valuable.

This shows that restricting LVT does nothing to vindicate any interesting version of LVT. The objection does not demonstrate that true belief is preferable to false belief. A restricted class of true beliefs may well be instrumentally valuable, but so is a restricted class of false beliefs. And there is no instrumental difference between what we can accomplish with the true ‘if, then’ beliefs and the false ‘if and only if’ beliefs. As such, there is no reason to prefer the restricted set of true beliefs to the restricted set of false beliefs.

The most natural thing to say, if the symmetry thesis is correct, is that truth and falsity are – like the blueness of a book – *equal or irrelevant* from the instrumental point of view. To keep *advocating* for the instrumental value of truth makes little sense, at least if this is supposed to tell us something about what we *should* believe. As we have seen, if the symmetry thesis is correct, we should no more try to acquire true beliefs than false beliefs. Nobody needs a single true belief to get everything they want. For this reason, whether my beliefs are true or false, in itself, should not concern me one bit. Neither of the objections

¹⁶ And nothing I say in this essay undermines this claim.

¹⁷ Restricted LVF could include other false beliefs as well, such as those that prevent you from getting on a plane that is going to crash.

undermines this claim. The kind of instrumental value the objector claims for true beliefs, and the kind that is compatible with the symmetry thesis, provides no action-guidance at all. This is because it does nothing to show that, if I want to achieve my goals, I should seek out or have a preference for true beliefs over false beliefs.

That the symmetry thesis, if correct, undermines LVT can also be seen by the fact that many claims made by instrumental value's defenders are incompatible with the symmetry thesis. Consider, for example, the following claims quoted above: 'Believing the truth is practically advantageous' (Lynch 2004, 16); 'we are more likely to get what we want if we base our deliberations and actions on true beliefs than if we base them on false ones' (Horwich 2006, 350); and 'having true beliefs will increase our chances of satisfying our other desires' (David 2001, 155). None of these can be sustained if the symmetry thesis is correct. This thesis says that, from an instrumental point of view, there is no reason to prefer true to false beliefs. Both can equally get you anything you desire. For this reason, it is not the case that truth is 'practically advantageous'. It has no advantages over falsity. Nor is it the case that 'we are more likely' to get what we want with true, rather than false, beliefs. We can get what we want equally well with both. For the same reason, it is incorrect that having true beliefs 'increase our chances of satisfying our other desires'.

The symmetry thesis also undercuts many arguments that assume the instrumental value of truth. Consider Carson's (2010, 89) claim, quoted above, that 'We are generally harmed when we are deceived because we cannot effectively pursue our ends and interests if we act on the basis of false beliefs.' This cannot be harm that deception causes if the symmetry thesis is correct. This is because this thesis

claims that we *can* just as effectively pursue our ends and interests on the basis of false beliefs as we can on the basis of true beliefs.

I have argued that there is no general instrumental consideration in favour of true beliefs. Nor is there any reason to think there is a general correlation between true beliefs and instrumental value as opposed to false beliefs and instrumental value. As there is no reason to prefer truth to falsity from an instrumental point of view, the claim that true belief is instrumentally valuable is – if meant in any interesting sense – incorrect. And, as we have seen, the price of its correctness is to strip LVT of all practical relevance.

Concluding Remarks

To conclude his discussion of instrumental value, Kvanvig (2003, 40) writes that ‘we cannot account for the value of true belief in terms of practical utility....

These remarks suggest that the value of truth is not in its capacity to further other interests we might have, but is rather intrinsic to truth itself.’ Kvanvig is half right. As I have argued, compelling considerations support the idea that the value of truth cannot be accounted for in terms of practical utility. This suggests, as he claims, that truth’s value is to be found, if it is to be found, in its intrinsic value.

Unlike Kvanvig, I believe – and will argue in what follows – that when we turn to the question of truth’s intrinsic value, we again come up empty.

As with instrumental value, I cannot pretend that my discussion of intrinsic value is the final word on the matter. Arguably, there will never be such a thing.

Nonetheless, in what follows, I focus less ambitiously on two issues that seem particularly pertinent in the present context. I hope to show that there are strong

reasons to reject the idea that truth is intrinsically valuable, and that a number of common arguments in CVT's favour are unconvincing.

Chapter 2: The Problem of Trivial Truths and Two Unsuccessful Solutions

Introduction

This chapter is concerned with the claim that true belief is *intrinsically* valuable (CVT). According to CVT, true belief is valuable *in itself* or *for its own sake*.

Thus, the value of true belief does not (or not just) arise from the fact that it leads to other things we value or desire. Rather, it is good to believe the truth simply *because* it is the truth. This would mean that, even if true belief is not instrumentally valuable, it is still valuable.¹⁸ The fact that a belief is true is *sufficient* for attributing value to that belief.

CVT is widely endorsed.¹⁹ This is evident within philosophy – particularly epistemology – and is also discernable in wider culture.

Epistemology is a hotbed for CVT. One reason for this is the vogue of explaining evaluative epistemic claims – such as whether a belief is justified or unjustified, rational or irrational – in teleological or consequentialist terms. On this view, a cognitive process is good – or earns positive marks – to the extent that it promotes or brings about things with intrinsic value. Conversely, a cognitive process is bad –

¹⁸ To be absolutely clear, I do not assume that true belief is *not* instrumentally valuable in this chapter. For all I say here, true belief could be instrumentally valuable. My concern is *exclusively* with intrinsic value.

¹⁹ Though not universally, of course. Complete ideological hegemony is rare in philosophy. An eclectic selection of philosophers who have disputed the idea that true belief is intrinsically valuable includes Nietzsche (e.g. 1966, sec 4), Rorty (e.g. 1998), and Stich (1990). A number of analytic Nietzsche scholars also seem (at minimum) tempted by this position. See, for example, Leiter (2002) and Gemes (1992).

or earns negative marks – to the extent that it fails to bring about or promote things with intrinsic value. And the thing most commonly claimed to possess intrinsic value is true belief. Thus, a cognitive process is good to the extent that it brings about or promotes true belief and bad to the extent that it fails to bring about or promote true belief.

Three examples will suffice for our purposes.²⁰ Alvin Goldman (2002, 52) expresses this view plainly when he writes:

I shall attempt to make a case for the unity of epistemic virtues in which the *cardinal virtue*, or the underlying motif, is something like true, or accurate, belief.... The principle relation that epistemic virtues bear to the core epistemic value will be a teleological or consequentialist one. A process, trait, or action is an epistemic virtue to the extent that it tends to produce, generate, or promote (roughly) true belief.²¹

William Alston makes similar claims. He argues that the evaluative aspect of epistemology involves identifying ways that our cognitive activities can be better or worse in relation to the goals of cognition. And the ‘primary function of cognition in human life’, according to Alston, ‘is to acquire true beliefs rather than false beliefs’ (2005, 28). He further suggests that, though having true beliefs is instrumentally valuable, it is ‘also of intrinsic value’ (31). Indeed, he argues that ‘it is as close to truistic as we can get in philosophy to take truth as a good-making

²⁰ All three are discussed further below. Some other defenders of this view, and hence of CVT, include: BonJour (1985), Lehrer (1990), and Foley (1987). See David (2001) for an extensive list.

²¹ As Grimm (2009, 20) makes clear, Goldman uses the term *cardinal virtue* to mean what is standardly meant by intrinsic value. That is, it is something that is pursued for its own sake.

characteristic, and falsity as a bad-making characteristic, of beliefs and other outputs of cognition' (31). Finally, Michael Lynch expresses the teleological view succinctly when he writes that '...the value of believing what is justified is parasitic on the value of believing what is true. Having justified beliefs is good because justified beliefs are likely to be true' (2004, 50).

This view of the importance of truth is not limited to defenders of teleological accounts. As David (2001, 151) writes, 'Epistemologists of all persuasions tend to invoke the goal of obtaining truth and avoiding error. This goal seems to be of special importance to epistemology. No other goal is invoked as frequently as this one. No other goal is given as much weight or is treated with as much respect as this one.' DePaul (2001, 172) makes a similar claim when he writes: 'I do not think that there is much question that the vast majority of epistemologists accept a theory of epistemic value very similar to the hedonistic theory just described. They take truth (or true belief) to be the only intrinsic epistemic good and falsity (or false belief) to be the only thing that is intrinsically bad.'

The idea that the truth is worth having for its own sake is also not confined to philosophy. It is deeply embedded in wider culture. It is not uncommon to hear of people quitting their jobs, or going on spiritual retreats, in search of the truth. However misguided the means, such examples demonstrate that people consider truth to be a valuable end. Stephen Grimm (2008, 725) provides the following anecdote:

The truth exerts a powerful attraction. Reading the newspaper over breakfast a few months ago, I came across the following quote from Ricky Williams, a running back for the Miami Dolphins who was in

the process of walking away from his million-dollar salary to pursue a career in holistic medicine. “I’m going to search for the truth”, Williams said. “Everything I’m doing in my life is about finding the truth”.

This almost romantic conception of the importance of truth – and a life committed to truth – is also a cornerstone of academic culture. This is reflected, for example, in numerous university mottos²² and biographies of great scientists and mathematicians.

Despite its prevalence, CVT faces a significant difficulty. This is the *problem of trivial truths*. Roughly, the problem is that many truths are so trivial that believing them seems to be entirely valueless. If this is correct, CVT should be rejected.

This is because CVT implies that such truths *are* valuable in virtue of being true. After all, a trivial truth is no *less true* than any other truth.

The above line of reasoning has been resisted in various ways. The central purpose of this chapter is to argue that the two most prominent ways of resisting the problem of trivial truths fail.

²² Here is an assortment of mottos from universities in the U.S (in English): Benedict College: *Truth and Virtue*; Brandeis University: *Truth, even unto its innermost parts*; CalTech: *The truth shall make you free*; Harvard University: *Truth*; Horace Man School: *Great is the truth and it prevails*; Indiana University: *Light and Truth*; John Hopkins University: *The truth will set you free*; Northwestern University: *Whatever things are true*; Northeastern University: *Light, truth, courage*; Rhodes College: *Truth, Loyalty, Service*; Stetson University: *For God and truth*; University of Miami: *Great is the truth*; University of Michigan: *Arts, Science, Truth*; University of Pittsburgh: *Truth and Virtue*; University of South Florida: *Truth and Wisdom*; Yale University: *Light and truth*.

Advantages of CVT and the Costs of Denial

Before criticising CVT, it is worth noting some of its attractive features and some costs incurred by denying it. This helps to explain both CVT's popularity and why many who endorse it are reluctant to renounce it despite recognising the force of the problem of trivial truths.

One point in CVT's favour – repeatedly emphasised by philosophers²³ – is that we sometimes want to believe the truth *just* to believe the truth. In other words, we consider certain truths worth having – and invest significant time and energy attaining them – even when they are not (at least obviously) instrumentally valuable. True belief sometimes *feels* intrinsically valuable to us.

If anybody doubts this, various examples speak in its favour. Take, for instance, the long quest to prove Fermat's Last Theorem. This theorem – which states that no positive integers, a , b , and c can satisfy the equation $a^n + b^n = c^n$ for any integer value of n greater than two – was first conjectured by Pierre de Fermat in 1637 in the margin of a copy of *Arithmetica*. Fermat claimed that he had a proof for the conjecture that was too large to fit in the margin. This theorem is one of the most famous in the history of mathematics and, prior to its solution by Andrew Wiles in 1994 (358 years after it was conjectured), numerous brilliant mathematicians spent their lives in an obsessive pursuit to prove it.²⁴ It is perhaps true that these attempted solutions – and the successful solution – have some instrumental value. The effort to prove it led to various breakthroughs in other areas of mathematics, for example. Even so, it seems unlikely that mathematicians were interested in this problem

²³ See, for example: Grimm (2008), Goldman (1999), Lynch (2004) and Pillar (2009). Chapter 3 discusses this claim in detail.

²⁴ See Singh (1997) for an entertaining discussion of this.

exclusively for this reason. Granted, some might have been seeking career success or mathematical immortality. But, for most, the main motivation seems to have been a desire for truth for its own sake.

CVT has advantages over its denial in such cases. First, CVT can explain why people value truth for its own sake without needing to posit a pervasive and systematic error or mistake on their part. Further, it allows us to say that certain intuitively valuable truths are *actually* valuable. There is no doubt that the denial of CVT can be counter-intuitive in certain cases. For example, it commits one to the claim that the proof of Fermat's Last Theorem had no value unless it happened to have instrumental value. If it didn't – and surely some intuitively valuable truths don't – then we have to say it was entirely valueless.

A Potential Concern About CVT

It might be thought that we already have a reason to be sceptical of CVT. The claim that true belief is instrumentally valuable is weaker than the claim that true belief is intrinsically valuable. If a weak claim about x 's value fails – as I have argued it does in this case – then there is little hope of establishing a stronger claim about x 's value. Therefore, there is little hope of establishing that true belief is intrinsically valuable.²⁵

This worry is misplaced. The truth of one kind of value claim cannot be judged by the falsity of the other. This is because there is no strong connection between instrumental and intrinsic value – they are logically independent. Due to this, there is no reason to suspect that true belief lacks intrinsic value *just because* it

²⁵ This is a point put forward by Pritchard (2007).

lacks instrumental value. That these often come apart is demonstrated by some simple examples. Money is a paradigmatic example of something that possesses instrumental, but not intrinsic, value. Happiness seems to go the other way around. That is, it possesses intrinsic, but not instrumental, value. This is at least the case according to theories where happiness is the only end with intrinsic value.²⁶ Finally, if value pluralism is correct – broadly construed to include any theory with more than one basic value – then something could both be intrinsically and instrumentally valuable. That is, good in itself but also good because it leads to something else of intrinsic value. Some philosophers believe love is like this: good in itself, but also good because it leads to pleasure.²⁷ These examples help illustrate that something having or lacking one kind of value is not an *a priori* reason to suspect that it will have or lack another kind of value. Thus, true belief lacking instrumental value should not make us antecedently suspicious of the idea that true belief possesses intrinsic value.

The Problem of Trivial Truths

We have seen that CVT has a number of attractive features. Nonetheless, CVT faces a serious difficulty. This is the *problem of trivial truths*. This problem is widely discussed²⁸ and it is considered troubling enough by many defenders of truth's value to necessitate various revisions of CVT.

²⁶ Utilitarianism is an obvious example of this, but there are many more if we construe happiness to include, for example, eudaimonia.

²⁷ Two examples are Lynch (2004) and Frankfurt (2004).

²⁸ See, for example, Brady (2009), Goldman (1999), Grimm (2008), Heal (1987/88), Horwich (2006), Kvanvig (2003), and Sosa (2001).

Though certain compelling cases tempt us to conclude that true belief is intrinsically valuable, such cases are a small part of the story. CVT is committed to far more than the value of mathematics. The problem is this: The truth is often trivial and boring. When it is, it seems absurd to say that it is intrinsically valuable. However, though trivial, these truths are no *less true* than the most interesting truth ever discovered. For this reason, CVT is committed to the claim that these trivial truths are intrinsically valuable. In short, CVT is committed to ascribing value to the valueless.

To feel the force of this objection, consider three examples:²⁹

Mandy spends endless hours counting how many grains of sand are contained in various square centimetres of a beach.

Tom, who cannot operate a telephone, spends his days memorising the contents of foreign phonebooks from the 1970s.

Cassie commits her life to learning the number of times ‘the’ is said in each and every fast food commercial ever made.

It is clear that Mandy, Tom, and Cassie would acquire numerous true beliefs. Supposing they committed all of their cognitive energy to their respective tasks, and carried them out with sufficient diligence, they would additionally avoid false beliefs. The problem for CVT, of course, is that it is highly implausible that such true beliefs are intrinsically valuable. Not only do Mandy, Tom and Cassie seem to be gaining nothing of value, they seem to be wasting their lives. Grimm (2008, 726) puts the point well:

²⁹ These are adaptations of common examples from the literature. See, for instance, Sosa (2001), Grimm (2008), and Zagzebski (2003).

[I]f we think that pursuing the truth is intrinsically valuable, then why are we unapologetically indifferent to so many truths? If you propose an evening memorizing the phone book for Topeka, Kansas, and I decline, have I really missed an opportunity to enrich myself, from an epistemic point of view? If the truth is always intrinsically worth pursuing, then it seems that I have. And yet that conclusion seems ridiculous.

Nobody is willing to accept this conclusion as it stands. It is generally agreed that trivial truths undermine *absolutist* readings of CVT. This would include, for example, the claim that the value of truth necessarily, or even usually, *overrides* other values, or that truth is always, or ordinarily, *all-things-considered* valuable. While rejecting absolutist accounts, many philosophers want to maintain that true belief is intrinsically valuable. To accomplish this, a number of modifications to CVT have been proposed. In the sections that follow, I discuss the two most prominent responses to the problem of trivial truths. I argue that both are unconvincing.

Truth as a *Prima Facie* Value

The first response – a common move in value theory – is to weaken CVT to a claim about *prima facie* value. As this promises an easy way out, it is probably the most tempting response to the problem of trivial truths. Call this account of truth's value:

PFA: True belief is *prima facie*, but not necessarily *all-things-considered*, intrinsically valuable.³⁰

The main advantage of the PFA is that it allows reservations regarding trivial truths to be explained away while leaving CVT standing. Though all true beliefs are valuable in virtue of being true, trivial truths are not all-things-considered valuable. Their value is outweighed by other considerations. A number of philosophers endorse the PFA, including Michael Lynch, Paul Horwich and Jonathan Kvanvig.

Lynch writes (2004, 58):

Without a doubt, there are all sorts of true beliefs (say, beliefs about how many threads there are in my carpet) that are not worth having, all things considered. But the fact that I should not bother with those sorts of beliefs doesn't mean that it isn't still *prima facie* good to believe even the most trivial truth.

Elsewhere, Lynch notes that to say something is *prima facie* good is to say that it is 'good considered by itself but not necessarily good all things considered' (2009, 227). He then argues that limits on our time and capacity mean that many trivial truths aren't worth believing, but claims that if we didn't have such limits – if the truth was 'cost-free' – then 'it would be good to believe all and only what is true' (227).

Similarly, discussing the problem of trivial truths, Horwich notes that valuing one thing does not prevent us from valuing other things. Sometimes our values will conflict. When they do, we must 'decide that some are to be sacrificed for the sake

³⁰ This acronym is short for the '*prima facie* account of truth's value'.

of others' (2006, 351). 'In such a situation', Horwich writes, 'the sacrificed values continue to matter of course – but they are outweighed by more important considerations' (351). According to Horwich, trivial truths shows us that such a conflict can take place between the value of truth and other values, but that this does not require us to abandon CVT. The truth always has intrinsic value; we simply need to recognise that 'in many circumstances, the value of finding out the truth, or falsity, of a given proposition will be less than the costs of doing so' (351). Kvanvig (2003, 41), too, claims that:

The default position for any truth is that our general interest in the truth applies to it, though, of course, there can be special circumstances involved so that the general interest in the truth is overridden by other factors.

It will be useful to explain the PFA in slightly more detail. The PFA claims that all true beliefs – including trivial truths – are intrinsically valuable. The reason they do not seem to be in our daily lives – the reason we are indifferent to the vast majority of available truths – is that the value of believing most true propositions is overridden. For example, it will generally be more valuable to pursue truths that promise a practical payoff. For this reason, we are not in fact motivated to discover the truth about many things. But this fact, according to the PFA, does nothing to undercut the intrinsic value that every true belief possesses. If we were not "finite beings"³¹ – if we had enough time and opportunity – we would have some interest in, and recognise the value of, even the most trivial truth.

³¹ To use Kvanvig's (2003) term.

The PFA, then, makes the following claims: It is always good to believe the truth. This is genuinely good because true belief is intrinsically valuable. No matter what truth we are considering, it is good to believe it. Nevertheless, some truths are not good *all things considered*. Often the value of a true belief will be outweighed by other more important considerations and values. Trivial truths are just an example of something genuinely valuable being outweighed by other things with greater value. Even though the value of true belief can be outweighed, truth remains intrinsically valuable nonetheless. Thus, CVT is correct.

As noted, hedging in this way is a common strategy in value theory. The claim that something is always and everywhere good or bad is often too strong. A philosopher somewhere can almost always dream up a counter-example. To accommodate this, it is standard to claim that x is *prima facie* good or bad, or good or bad all else equal. As long as x has some positive or negative value, this response can be effective.

Even so, there is something unsatisfying about this response. Once weakened to *prima facie* value, a value claim becomes incredibly difficult – if not impossible – to refute. This is because its defender has an easy way out of any objection made: They can just assert that this is a case where all is not equal. Since *prima facie* claims are so hard to argue against, they are an effective way to preserve biases and force a theory to get the results one desires.

Against the PFA

I will now argue that the PFA fails. It does not save CVT from the problem of trivial truths. As mentioned, *prima facie* value claims about x are only effective when x clearly has *some* value or disvalue. Standard examples include that lying is *prima*

facie wrong and that promise-keeping is *prima facie* good. *Prima facie* claims are entirely ineffective in cases where x seems to have *no* value whatsoever.

Consider, for instance, the claim that rolling around on the ground for no reason is *prima facie* valuable. As above, to every counter-example given I can reply that this is a case where all else is not equal. But, unlike in the standard examples, this response does not seem *at all* plausible. This is because rolling around on the ground seems to lack not only absolute value, but *any* value. And by claiming that rolling around on the ground for no reason is valuable, albeit *prima facie*, I am committed to the proposition that it is a genuinely valuable activity.

I will argue that truth is more like rolling around on the ground than promise keeping. If it is wrong to call trivial truths valuable in a strong sense – a point that is widely accepted – then the idea that they are *prima facie* valuable should also be rejected. This is because certain truths seem to lack not only absolute value, but any value at all. If this is right, then the PFA fails for the same reason that the rolling around on the ground claim fails. Further, as will become clear, since the PFA is equally committed to the intrinsic value of trivial truths, the account preserves most of the counter-intuitive consequences that led to a reformulation of CVT in the first place.

The Problem

The PFA remains committed to the intrinsic value of trivial truths. It is committed, that is, to the claim that counting sand is a valuable activity. Downgrading the value of true belief to *prima facie* value does not avoid this consequence. This is because assigning *prima facie* value to something does not deny it value. It just means that

its value can be outweighed in certain situations. The problem is this: Trivial truths do not seem valuable at all.

How bad is this problem? We might not think there is much harm in attributing *some slight* intrinsic value to all truths if this secures the result that intuitively valuable true beliefs are actually valuable. This way of thinking is flawed.

The value assigned to trivial truths by the PFA is significant, as can be seen if we consider a number of conditional claims that should hold for any intrinsically valuable property. Start with the following – almost certainly correct – conditional: If x has intrinsic value, then there is something to be said in favour of x . One reason that this conditional seems correct is that, if there was nothing to be said in x 's favour, then the idea that it had value would seem meaningless.

Now consider some slightly stronger – or at least more strongly expressed – conditionals. If x has intrinsic value, one ought to be able to make the following claims: 'doing x wouldn't be a bad way to spend your life, or conduct your life, or a bad thing to orientate your life towards'. I believe these stronger conditionals must also be correct. If x is intrinsically valuable, it seems clear that a person who carried out x for her whole life will have done something valuable with her life – even if only slightly valuable. If this isn't the case, then it isn't clear why we would think that x has intrinsic value. At absolute minimum, it seems correct that, if x is intrinsically valuable, then a person would not have *wasted* – or done something *disvaluable with* – her life if she orientated it towards x .

The accuracy of these conditionals does not rest on a substantive theory of intrinsic value. It is not obvious how they could be rejected without rejecting the concept of intrinsic value itself. Even philosophers who are sceptical about the importance of

intrinsic value accept such statements. Harry Frankfurt (2004, 23; emphasis added), for example, writes:

Caring about something differs not only from wanting it, and from wanting it more than other things. It differs also from taking it to be intrinsically valuable. Even if a person believes that something has considerable intrinsic value, he may not regard it as important to himself. *In attributing intrinsic value to something, we do... imply that it would make sense for someone to desire it for its own sake – that is, as a final end, rather than merely as a means to something else.*

Now the key point: these conditionals are not satisfied by trivial truths. It seems clear that you would be wasting your life if you spent it counting sand for no practical purpose, and equally clear that you would not have done something valuable with your life if doing so was your only activity. Since trivial truths are obviously truths, we should therefore reject the claim that true belief *simpliciter* has intrinsic value. And since the PFA is committed to this claim – even if only to its *prima facie* value – we should reject the PFA. Trivial true beliefs lack any intrinsic value at all.

Note that such conditionals do not apply to instrumental value. It makes perfect sense to consider it pointless for somebody to spend her whole life doing *y* if she never attains *x*, where *x* is the thing that *y* is valuable in virtue of. Suppose – against conventional wisdom – that money is instrumentally valuable because it buys happiness. If a person spent her whole life accumulating money, but never purchased happiness, then she has wasted her life even though she spent it doing an

instrumentally valuable activity. After all, money was simply a contingent means to an end. Ends often justify means, but means rarely justify ends.

To make this criticism more concrete, consider the three cases (Mandy, Tom and Cassie) used to illustrate the problem of trivial truths. As noted, each person will acquire numerous true beliefs. According to the PFA, these truths are *prima facie* intrinsically valuable. That is, they possess genuine intrinsic value, but their value will be outweighed in certain circumstances by other values. As Horwich (2006, 351) puts it: ‘In such a situation, the sacrificed values continue to matter of course – but they are outweighed by more important considerations.’

The question is whether these truths possess any intrinsic value at all, however small. If they do – as the defender of the PFA claims – then conditionals like those above should hold. But I cannot see why anybody would think they do. First, take the – almost analytic – conditional: ‘If x is intrinsically valuable, then somebody who spent her whole life doing x would have spent her whole life doing something valuable’. This conditional seems absurd when applied to trivial truths. The idea that somebody who spent her whole life counting sand, learning names from a phonebook, or counting the number of times “the” is said in fast food commercials would have done something valuable with her life is ridiculous.

Even the weaker conditional – ‘if x is intrinsically valuable, then a person would not have *wasted* her life if she orientated it towards x ’ – is difficult to believe.

Consider an example:

Suppose that Jessie is anxious about her future. She has spent the last few months dwelling on various options and trying to decide what to do. She feels she is not making progress, so she decides to seek advice

from Alice, who she greatly admires. The two meet for coffee.

Eventually, Jessie says to Alice: ‘The main thing I’m scared about is wasting my life. I know that it’s hard to accomplish anything great, but I don’t want to get to the end of my life and feel like I’ve done nothing of value. What do you think I should do?’ Alice replies: ‘It is hard to say what a person should do with their life. But if you don’t want to waste it, spending all your time carefully counting the number of threads in different pieces of carpet is a safe bet.’

I take it that this is horrible advice. Spending your life carefully counting threads of carpet is a sure-fire way to waste it. If this is right, then it can’t be the case that trivial truths have intrinsic value. But the defender of the PFA – no less than the defender of an absolutist version of CVT – is committed to this claim. They would have to agree with Alice that this is not a bad way to spend your life. Since this is incredible – since acquiring such true beliefs seems to lack any value at all – the PFA cannot be correct.

A Different Way to the Same Conclusion

We can get to this conclusion another way, which I will briefly note. The PFA implies that if there were no overriding considerations – or if we had no overriding interests – then it would be good to acquire true beliefs about trivial matters. These truths are worth believing when all else is equal. This was expressed by Lynch when he wrote that, if his intellect and time were unlimited, it would be good for him to believe ‘all and only what is true.’ Kvanvig also makes this point when he says that if we were not ‘finite beings’ we would have an interest in every truth.

This consequence of the PFA is difficult to accept. Even if there were no overriding considerations – for whatever reason – it still does not seem valuable to memorise a phonebook or to count sand for no practical reason. Grimm (2008, 732) expresses this well when discussing Kvanvig’s claim:

Suppose we take away my finitude, at least in the sense of making me immortal. If at some point counting the motes of dust on my desk seemed worth doing from a purely intellectual point of view, then I can only conclude with Bernard Williams (1976) that immortality would be a tedious and dreary prospect indeed, and *itself* not worth having.

To put this slightly differently: If it is possible that memorising phonebooks could be a legitimately valuable way to spend your life, then being dead seems better than living certain valuable lives.

Further Issues for the PFA

It is worth highlighting a further point. This should also alleviate any lingering feeling that assigning *prima facie* intrinsic value to *all* truths is worth the cost. The examples of trivial truths most commonly discussed are similar to those above. But these examples do not do justice to just how trivial true beliefs can be. As Marion David (2005, 297) points out:

...“all truths” comprises not only each and every trivial truth of the “how many threads in my carpet”- variety, but also every conjunction of every two truths I already believe (trivial or non-trivial), including, of course, complex conjunctions whose conjuncts are themselves conjunctions of truths I already believe (trivial or non-trivial); it also

comprises every disjunction of (trivial or non-trivial) truths I already believe with any propositions you like; and lots of other redundant “garbage” of this sort.

This hints at a further issue for the PFA: Some trivial truths come without a cost. Recall that, according to the PFA, believing trivial truths is all-things-considered valuable provided that its value is not outweighed by other considerations. One hope was that this claim would allow those truths that are intuitively disvaluable – like counting sand – to be explained away. This overlooks the fact that we can acquire many trivial truths with minimal, or no, effort. Thus, not all trivial truths will be outweighed by other considerations. According to the PFA, such truths are all-things-considered valuable. The problem is that these easily gained trivial truths also seem valueless. David (2005, 299) writes:

We pick up trivial truths at no additional cost... maybe we are listening to the TV while doing something else that is actually worthwhile but doesn't require a lot of concentration – even commercials will be good for truths like “They are *saying* this is the best toothpaste of all times”, though probably not for “This *is* the best toothpaste of all times”. Do we really care about believing trivial truths that come for free? In the end, this is of course a personal question. But I suspect quite a few people will say: Not at all.

This argument can be extended. Not only are there truths that seem even more trivial than the common examples, there are true beliefs that seem *disvaluable*. Consider some examples of possible true beliefs: A person can have true beliefs about what it feels like to be hung, drawn and quartered; or have true beliefs about what it is like to torture somebody – and what it feels like to desire to torture somebody; or

have a true belief about what a particular child looks like naked; or have true beliefs about what it feels like to hate somebody solely because of her race or sex. There are many cases where – intuitively at least – it seems bad to have true beliefs. There is simply nothing to be said in their favour.³² Lynch, Horwich and Kvanvig are all committed to the claim that it is *prima facie* good to have true beliefs about, for example, what different children look like naked. That is, that – considered by itself – it is good to have such true beliefs. As discussed, by attributing intrinsic value to these beliefs they are also committed to various conditionals regarding them. For instance, that spending your life acquiring true beliefs about what different children look like naked is a valuable way to spend your life. To my mind, this is an absurd consequence of the PFA.

For all these reasons, I conclude that the PFA is implausible and thus fails as a response to the problem of trivial truths.

Restricted Versions of CVT

Based on our discussion so far, the problem of trivial truths may seem to have an obvious solution. This is to restrict CVT. According to this account, only a restricted class of true beliefs are intrinsically valuable. True beliefs that are not a member of this class fail to be intrinsically valuable. Call theories with this structure:

³² It might be the case that certain false beliefs about these things are not good either. If anything, it seems like having no beliefs – true or false – would be ideal. Even if this is the case, it does not help the PFA or CVT.

RVT: Only a restricted class of true beliefs are intrinsically valuable.

True beliefs that fall outside this class are not intrinsically valuable.³³

RVT is an appealing strategy, and is perhaps the most popular way to preserve CVT against the trivial truths objection. The reason for RVT's popularity is that it promises to capture two attractive ideas. First, it seems to allow us to say that certain intuitively valuable truths – such as mathematical truths – are valuable even if they lack instrumental value. Second, RVT can deny the implausible claim that trivial truths are intrinsically valuable. These can be denied membership to the class of intrinsically valuable truths. In short, RVT promises to hold onto some of the advantages of a pure version of CVT without incurring the costs.

The argument for RVT generally consists of three steps. These steps clearly demonstrate the above motivation. I will illustrate this argument with examples from Goldman's work.

The first step is to claim that true belief can be intrinsically valuable, or valuable for its own sake. A number of remarks show that Goldman endorses this claim.

For example, he writes (1986, 98):

³³ Chapter 1 discussed a restricted version of LVT. My concern in this section is *exclusively* with restricted versions of CVT. Though some of my criticisms here may apply to restricted versions of LVT, not all do. In any case, my only intention is to discuss restricted versions of CVT, and it is to restricted versions of CVT that 'RVT' refers. As such, this section leaves open the possibility that restricted versions of LVT could avoid all of the problems I discuss. It is worth noting, though, that if it turns out that a restricted version of LVT is the best response to the problem of trivial truths – or the best account of truth's value – this would just show that the problem succeeds. The objection is designed to show that truth is not *intrinsically* valuable, not that it is not instrumentally valuable. For this reason, to endorse LVT in the face of the problem is just to concede that the problem is correct and that truth is not intrinsically valuable. Chapter 2 is a success if it shows that truth is not *intrinsically* valuable. It is not supposed to show that truth is not instrumentally valuable.

Truth acquisition is often desired and enjoyed *for its own sake*, not for ulterior ends. It would hardly be surprising, then, that intellectual norms should incorporate true belief as an autonomous value, quite apart from its contribution to biological or practical ends.

Goldman's view is actually much stronger than this remark suggests. As noted, he is an advocate of a consequentialist approach to epistemology he calls *veritistic unitarianism*. Roughly, Goldman's view is that true belief is the only thing of epistemic intrinsic value, and that all other things of value – such as rationality, justified belief, etc. – derive their value solely from their relation to true belief – that is, these other things are merely instrumentally valuable. This is apparent from the remark quoted at the start of this chapter, and further support is easy to find. For instance (1999, 69):

From a veritistic standpoint, true belief is better than either ignorance or error, and higher degrees of belief in truths are better than lower degrees of belief.... These sorts of states have 'fundamental' veritistic value (in varying degrees), and practices that produce changes in veritistically valuable states have 'instrumental' veritistic value (V-value).

Goldman is not the only teleologist who endorses RVT. As the quotes at the beginning of this chapter make clear, Alston also endorses these claims. He argues that the 'evaluative aspect of epistemology involves an attempt to identify ways in which the conduct and the products of our cognitive activities can be better or worse vis-à-vis the goals of cognition.' And the 'primary function of cognition in human life', according to Alston, 'is to acquire true beliefs rather than false beliefs *about matters that are of interest to us*' (2005, 28; my italics). Alston also

endorses CVT. He claims that, though having true beliefs is instrumentally valuable, it is ‘also of intrinsic value’ (2005, 31).

The second step in the argument for RVT is to accept the problem of trivial truths at face value and concede that certain true beliefs have no value at all. This is where RVT diverges from the PFA. Goldman (1999, 88-89) writes:

Suppose S is ignorant about all of the following matters: What is the 323rd entry in the Wichita, Kansas, telephone directory? Who placed sixth in the women’s breast stroke at the 1976 Summer Olympics? What was the full name of Domenico Scarlatti’s maternal grandmother? Does S’s ignorance on all these matters constitute, or even contribute toward, the impoverished V-condition of his credal corpus? Does such ignorance imply that his credal state should receive a low V-value, or V-ranking?... I am inclined to say that his knowing no answers to them does not count against the V-value of his belief states.

The final step is to argue that this fact does not undermine CVT. Rather, what trivial truths show is that the class of true beliefs that are intrinsically valuable needs to be *restricted*. Goldman (1999, 89) writes:

In constructing a model of V-value, it is helpful to use a *question-answering* model.... In a question-answering model, agent S’s belief states... have value or disvalue when they are responses to a question that *interests S*... In short, V-value should always be assessed relative to *questions of interest*.

Contra the PFA, then, RVT claims that it is not correct that *any* true belief is intrinsically valuable. But it *is* correct that certain true beliefs are intrinsically valuable. Though I have illustrated RVT through Goldman, he is far from alone in endorsing it.³⁴

RVT promises to capture two compelling ideas. If it succeeds, RVT may be the ideal account of the intrinsic value of truth and a decisive response to the problem of trivial truths. I argue that RVT does not live up to this promise.

Against RVT

As a response to the problem of trivial truths, RVT appears to give us the best of both worlds. It recognises the force of the objection – and the implausibility of attributing value to trivial truths – *and* preserves the value of truth for its own sake. This appearance is an illusion. Despite what many of its defenders seem to believe, restricting CVT amounts to abandoning CVT.³⁵ This is because RVT cannot claim that any true belief is *ever* valuable in virtue of being true, nor can it claim that any true belief is *ever* worth having for its own sake. For this reason, RVT is not a defence of CVT at all. Hence, it is not a *response* to the problem of trivial truths. If anything, RVT simply concedes that the objection is correct and that truth is not valuable for its own sake.

Further, RVT violates the *spirit* of CVT. No defender of RVT can claim – in good faith – that truth *matters* or that truth is a particularly significant value. Many

³⁴ Other philosophers who endorse restricted versions of VT include: Coates (2009), Haack (1993), David (2005), Sosa (2001), Pillar (2009), and Grimm (2008).

³⁵ This line of argument seems to be suggested by Whiting (2013), but he does not develop it.

defenders of RVT do not seem to have realised this and continue to display a reverential attitude towards truth. Nobody who *genuinely* cared about truth would endorse RVT.

To be clear, even if RVT does abandon CVT, *this* does not show that it is incorrect. There are strong reasons to abandon CVT. My claims are different: First, RVT fails *as a defence of CVT* against the problem of trivial truths. This is how it is often employed and is the focus of this chapter. Second, many defenders of RVT are badly mistaken about the implications of their position for the value and importance of truth.

First Criticism of RVT

I begin with what may seem the less serious criticism. This is that RVT violates the *spirit* of CVT. To restrict the value of truth is to deny it of significance. Though it seems less serious, this criticism paves the way for the second criticism. This is because it supports my claim that defenders of RVT are badly mistaken about what their position involves. This is hard to believe – both criticisms seem fairly obvious – but it is easier to accept once we realise that multiple examples support this charge.

A reverential attitude towards truth pervades much of the writing on RVT. This can be seen in a number of the above quotes. To demonstrate it more fully, I shall discuss Christian Pillar's fascinating article *Desiring the Truth and Nothing But the Truth* (2009). Pillar provides a lucid example of this outlook.

Note first that Pillar is an advocate of RVT. In response to the problem of trivial truths, he writes (195):

I do not think the existence of uninteresting truths points to a deep problem. Thus, it does not need anything ‘deep’ to solve it. I do not know anyone in Irkutsk. The entries in its local phone book are certainly of no interest to me. Nevertheless, I think of myself as being interested in truth. Not in any truth, obviously, but in the correct answers to questions that are of some concern to me. I will restrict a person’s interest in truth to those propositions the person finds, in some way or other, interesting.

This remark demonstrates that Pillar endorses RVT.³⁶ It may not seem to show that he regards truth as particularly important or weighty. All he says, after all, is that ‘I think of myself as being interested in truth.’ I think of myself as being interested in poker. By this, I just mean that I am more than indifferent towards it and enjoy playing it. But if I never played poker again, this would be no great loss to me. To use Frankfurt’s phrase, my interest is not *wholehearted*.

Pillar’s remark is deceptive out of context. As he makes clear throughout the article, the sense of interest he is discussing is *wholehearted* interest. This is shown, for example, by his frequent use of ‘lover of truth’ to describe the person who is interested in truth.³⁷

Further support for this claim is easy to find. For instance, Pillar discusses a case in which a person is seriously ill and knows she would be happier not believing the truth about her condition. He considers what a *genuine* lover of truth would do

³⁶ This is also expressed in other pronouncements, such as: ‘Our interest in truth is given by the first part of the bi-conditional – to believe p if p (for all interesting propositions p)’ (2009, 208).

³⁷ E.g. pages 193, 204, and 210.

in this case. He considers two possibilities. First, she could display a *Part 1* interest in truth.³⁸ On this conception, we want ‘to believe p if p is the case’ (2009, 200). As Pillar (2009, 205) writes:

Being concerned for truth is being open to the world. Let me believe whatever it is that the world might throw at me. It is a concern for my beliefs that is specified by how the world is. Such a concern is of the form ‘If the world is such and such, let me be such and such’.

In this case, she would want to believe whatever her actual condition is, even if this makes her miserable.

Alternatively, she could display a *Part 2* interest in truth. This involves being concerned with whether the *world* matches what we happen to believe. It takes the form: ‘if I believe such and such, then let the world be such and such’. In the ill woman’s case, she would believe that she was well, and then want the world to match this belief.

Pillar rejects a Part 2 interest as a *genuine* interest in the truth – which he considers himself to possess. In doing so, Pillar makes it clear that he believes that truth *matters*. A number of examples bear this out:

Part 2 allows for trade-offs which, in my view, are incompatible with a person who is interested in truth.... Even if it would be beneficial to me to live in the illusion of health when I am ill, such an interest is incompatible with, what in my view, is an interest in truth. (204)

³⁸ This classification is based on William James’ (1956) distinction between wanting to believe the truth and not wanting to believe falsehoods.

In contrast to Part 2, Part 1 does... capture the strength of character that we might ask from a lover of truth, who shuns the pleasant illusions and prefers to live in a way that keeps him in touch with reality, even if it contains things one would rather avoid. (204)

When it comes to the goods that believing can bring, we want the truth and nothing but the truth. (205)

Special circumstances have to be put in place in order to make a Part 2 concern legitimate. In any case, it is not part of a concern for truth... Being concerned for truth is being open to the world. Let me believe whatever it is that the world might throw at me. (205)

As these remarks indicate, Pillar considers truth to be significant and weighty. For instance, he sees it as more important than the kinds of happiness or comfort we can get from false beliefs.

Defenders of RVT are not entitled to claims – such as Pillar’s – about the weight or importance of truth. Those advocates of RVT who continue to make such claims are mistaken – or self-deceived – about the implications of their own position. Really *caring* about truth is undermined by RVT.

Begin with a general point: If you care about x – or take x to be a weighty value – you should value x even when it is difficult. It is *too easy* to care about – or think you care about – something when it is entangled with other attractive things.

Consider an example. Suppose that philosophers were usually paid millions of dollars a year. And suppose that a university was hiring a new philosopher. This university is determined to hire somebody who *really* cares about philosophy, and

not somebody who cares about the money. Finally, imagine that openings at universities are so rare that there is no reason to suppose that anybody will ever get more than one chance (though there are plenty of other high-paying jobs around).

One obvious, though questionable, test for whether the candidates genuinely cared about philosophy would be to advertise the job as lower paid than the usual millions – say, enough for a comfortable middle-class life, but no more. This would divide those who care about money from those who care about philosophy. Indeed, it would be natural to say that those who pulled out – even if they enjoyed philosophy – did not *really* care about philosophy at all, or at least did not consider it valuable in a weighty way. Without the money, they did not consider it worth doing. It would not ring true, for instance, if they claimed that they were a *lover* of philosophy. After all, they were not expected to do philosophy for free – or to starve to death for it – only to do it for less than millions. Even if such philosophers considered *themselves* to be true lovers of philosophy, this test would seem to show that they were deluded or self-deceived.³⁹

³⁹ There is a possible objection to this example. It might be claimed that the test is faulty because it would only succeed in dividing those who value philosophy intrinsically from those who value it instrumentally. And that this would not divide those who *really* care about, or love, philosophy from those who don't, because an advocate of the instrumental value of philosophy can still love it. There may be something to this objection, but I don't think it is very strong. This is because I doubt that valuing philosophy instrumentally in the example *is* compatible with really loving philosophy. Thus, even if it does divide on this basis, it does its job. To my mind, giving up philosophy *just* for more money – especially when you would have had a materially comfortable life if you had stuck with it – is a good indicator that you don't consider philosophy a very weighty or significant value. Further, caring about something *just* as a means to more money does not seem to be the kind of thing we have in mind when we say that somebody *loves* something. In any case, if this example is rejected, further considerations in favour of my claim are provided below.

The standard for *genuinely* caring about truth is no different. Pillar makes this point when he says that being interested in truth is incompatible with comforting or pleasurable illusions. It is too easy to claim you care about truth – and perhaps, like Pillar, even think you do – when it is only considered in cases where you get, or at least don't lose, other good things along with it.

For this reason, the only useful test for whether a person *really* cares about truth is a case where truth is isolated from other good things. If you do not consider truth valuable in these kinds of cases, then – like the philosopher who turns down the job because they will not become a millionaire – the natural conclusion is that they don't *really* care about truth, or at least that they do not give it much weight. It would not ring true, for example, if you claimed to *love* truth. Even if you consider yourself to be genuine lovers of truth, such cases would seem to show that you are deluded or self-deceived.

Now the key point: Defenders of RVT do not care about truth in the exact kinds of cases that would pick out a *genuine* lover of truth. This is because trivial, uninteresting, or boring truths are a *perfect example* of the kind of case where a genuine lover of truth would endorse its value. These are cases where truth is the only possible thing that could have value, and hence truth is isolated from other things that could be considered valuable. In these cases, however, defenders of RVT claim that truth *has no value at all*. This could be seen in two examples above. Pillar discussed an uninteresting truth and said that he did not care about it and Goldman noted a number of trivial truths and claimed that they had no value. As such, defenders of RVT cannot legitimately claim to love truth.

A number of further points demonstrate the tension between RVT and a reverential attitude towards truth:

(1) With some possible exceptions, anybody can endorse the value of anything if it is appropriately restricted regardless of her other values. For this reason, to say that you endorse a restricted version of the value of x is to say very little about how important you consider x . As a rule, it is a sign that you do not actually think x is all that important.

For example, I believe that pleasure is intrinsically good. It is good, I think, in every single case in which it occurs. For this reason, I am willing to claim that I have the kind of interest in pleasure that Pillar claims to have in truth. I further believe that suffering is intrinsically bad. It is bad, that is, in every case that it occurs.

Despite these beliefs, I am happy to endorse a restricted version of the value of suffering. According to this principle – call it RVS – suffering is always valuable when it causes more pleasure in the future for the victim than her not suffering would have. Despite endorsing RVS – that is, endorsing the value of suffering in the same way that RVT endorses the value of truth – it seems obviously misleading to call me a *lover* of suffering. There would be something absurd about me writing an article called *Desiring Suffering and Nothing But Suffering* in which I made statements like ‘When it comes to the goods that sensations can bring, we want suffering and nothing but suffering.’ A *genuine* lover of suffering would rightfully accuse me of being a fraud who doesn’t really care about suffering. Defenders of RVT are open to similar charges when they claim to love truth.

(2) There is a very basic reason to think that RVT does not cohere with a reverential attitude towards truth. This is the fact that defenders of RVT do not care about a large class of truths. They claim that countless truths possess no value at all. Nobody would take somebody seriously who claimed to be a genuine lover of pleasure but denied that numerous instances of pleasure had any value. Pleasure is pleasure after all. Similarly, there is no reason to take seriously somebody who claims to love truth – or claims that truth is significant – but denies that numerous truths have any value.

(3) Suppose – plausibly – that Pillar’s description of what it is to have a genuine interest in truth is correct. RVT is clearly incompatible with this conception. For one, RVT denies that I should ‘believe whatever it is that the world might throw at me.’ On the contrary, RVT instructs me *not* to believe countless things the world throws at me. This is because, according to RVT, I have no reason to believe any truth that falls outside the restricted class. If this class includes trivial truths or truths that I do not find interesting, then it tells me not to believe most truths the world throws at me on any given day.

Further, and against a Part 1 interest, RVT tells us *not* to ‘keep in touch with reality, even if it contains things one would rather avoid.’ Instead, it tells us that *anytime* we are not interested in a particular truth – and hence would rather avoid it – we should not worry about believing it.

In short, Pillar claims that a genuine interest in truth is ‘of the form “If the world is such and such, let me [believe] such and such”’. If this is correct, then Pillar’s own interest in truth is not genuine. The world is such and such in countless ways that I am not interested in, and RVT tells me not to believe any of it.

(4) RVT often lines up with a Part 2 interest in truth. This is the kind of interest that Pillar – again plausibly – rejects as a genuine interest in truth. On this conception, I want the world to be whatever way I believe it to be – ‘If I believe such and such, let the world be such and such.’

To see the problem, return to the health example. Regarding this case, Pillar (2009, 204) writes: ‘Even if it would be beneficial to me to live in the illusion of health when I am ill, such an interest is incompatible with, what in my view, is an interest in truth’. This statement is rejected by RVT. Whether a person should live in this illusion or not depends *entirely* on whether she is interested in the truth about her condition. That is, whether she *wants* to know the truth. If she is not interested, then RVT says she has no reason to care about her condition.

According to Pillar, such a claim is incompatible with a *genuine* interest in truth. If this is correct, then Pillar does not have a genuine interest in truth.

This point can be extended. Imagine that somebody, Fred, considers himself to be a genuine lover of truth. Suppose that Fred has become so depressed that he has entirely lost interest in life, including the truth of all propositions. Fred endorses RVT and believes that a genuine interest in truth involves wanting to ‘believe p if p (for all interesting propositions p).’⁴⁰ Pillar would have to agree with Fred that – though he doesn’t care about the truth of any proposition – he displays a *genuine* interest in truth. But that seems absurd. Surely Fred is a prime example of somebody who does *not* love or care about truth. It is certainly odd to say that he has an *interest* in truth, as Pillar would have to say. After all, Fred is not interested in anything.

⁴⁰ This is one of Pillar’s (2009, 208) formulations of RVT.

RVT appears to be an easy way to preserve the significance of truth without having to endorse the value of trivial truths. Things are not so easy. As the above example illustrates, RVT does not sit well with the idea that truth *matters*. Many defenders of RVT do not realise this implication and continue to exalt truth. This cannot be sustained. The price of RVT is the rejection of truth's significance. If one is not willing to pay this price – and wants to keep making claims like Pillar's – RVT must be rejected.

Second Criticism of RVT

RVT has a deeper problem. It is incoherent *as a defence* of CVT. This is because RVT is incompatible with the idea that truth is valuable in its own right, or for its own sake. For this reason, RVT cannot save CVT from the problem of trivial truths – as it is commonly employed to do. Indeed, it is not a *response* to the problem at all. It amounts to nothing less than abandoning CVT, which is exactly what the objection claims we should do.

As we have seen, various philosophers endorse both RVT and CVT. Goldman, it will be recalled, claims that truth is the only thing with intrinsic epistemic value. The value of everything else is merely instrumental. Alston endorses a similar teleological view – he writes that truth is the goal of cognition and of 'intrinsic value' – along with RVT.

Neither Goldman nor Alston seem to consider endorsing both RVT and CVT to be particularly problematic. As far as I can tell, Alston says nothing about any tension between the two. Goldman does note some tension, but I presume he does not think it is too serious given that he spends the rest of his book, and much of his later work, demonstrating Unitarianism in action. He writes (2002, 61):

Admittedly, the dimension of interest does complicate our story. We can no longer suggest that higher degrees of truth-possession are all that count in matters of inquiry. But can't we incorporate the element of interest by a slight revision in our theory? Let us just say that the core epistemic value is a high degree of truth-possession *on topics of interest*. Admittedly, this makes the core underlying value a somewhat 'compound' or 'complex' state of affairs. But, arguably, this is enough to preserve the idea of thematic unity, and thereby preserve Unitarianism.

This is not so. It is incoherent to advocate both RVT and CVT. There are a number of ways to get to this conclusion. I demonstrate this incompatibility with two arguments. The first I shall call the *no difference* argument and the second the *further question* argument.

Before laying out these arguments, it is worth noting that my claim is intuitive. The versions of RVT that are of interest to us – those that are supposed to defend CVT against the objection from trivial truths – argue that many true beliefs are intrinsically valuable *and* that countless beliefs are not intrinsically valuable despite being true. There is clearly something paradoxical about this. It is a bit like somebody claiming to love everything about her husband despite hating his face and personality.

I start with the no difference argument. Recall that, according to CVT, true belief is valuable in itself or for its own sake. According to RVT, only a restricted class of true beliefs are valuable in themselves. When a true belief falls within this class, it is intrinsically valuable and when it doesn't it lacks intrinsic value.

One simple way to see that RVT abandons CVT is to ask *why* the value of truth is restricted. It seems clear that RVT cannot answer this by appealing to truth *itself*. Again, a trivial truth is no less true than the most interesting truth ever discovered. Any answer that could be given would have to appeal to features of the restricted class that have nothing to do with truth. That is, factors that are *external* to truth. But once this move is made – and those truths that don't live up to the criteria are denied value – any claim that the value being assigned to truth is intrinsic is undermined. It simply makes no sense to claim that one truth possesses value solely in virtue of being true, but that another truth possesses *no* value.

We can put this slightly differently: If certain true beliefs are valuable in virtue of being true, then it is entirely arbitrary to deny any other true belief intrinsic value. Both beliefs are true regardless of external, non-intrinsic, differences.

The reason this is a 'no difference' argument should now be clear. If truth is sufficient for attributing value to a belief, then *any* true belief is intrinsically valuable. Trivial truths and interesting truths are both true. Thus, if one kind of true belief is intrinsically valuable – as RVT claims – then all truths should be intrinsically valuable. There is simply no *intrinsic* difference between the two types of truth. This makes RVT's claim that certain truths but not others are intrinsically valuable arbitrary. No relevant difference can justify such a claim. It is certainly possible, of course, that one truth is *more* valuable than another for non-intrinsic reasons. For example, one truth may be more instrumentally valuable, or more interesting. But this has nothing to do with CVT.

Next consider the *further question* argument. I will begin with some strong statements, and then defend them. A consequence of RVT is that no true belief is

ever valuable in itself, for its own sake, worth having in its own right, or worth pursuing or attaining in virtue of being true. Not only can true belief not play the role Goldman and Alston claim it can, truth itself is *entirely* lacking in value according to RVT.

The reason is this: RVT cannot claim that belief *x* is valuable *because* it is true. There is always the further question: ‘Does belief *x* fit into the restricted class?’ This fact denies truth any value for its own sake. To say that true belief is intrinsically valuable is to say that true belief is *sufficient* for value. All you need to know is that a belief is true and you know it is valuable (at least a little bit). This is denied by RVT. According to RVT, there are two necessary conditions for a belief to be valuable. Belief *x* is valuable because (1) it is true *and* (2) it falls within the relevant restricted class. For this reason, true belief is not *sufficient* for value according to RVT, and, hence, true belief is not intrinsically valuable according to RVT.

Consider an example. For simplicity, suppose that what determines whether a particular true belief falls into the relevant class is the believer’s interest in that truth. And suppose that ‘interest’ here just means whether I have a pro-attitude towards believing this truth.⁴¹ Now imagine that a person, call him Ryan, stumbles upon a true belief. Ryan comes to believe, truly, some fact about evolution that would change the face of biology. Is this true belief valuable for its own sake? CVT and the PFA have no problem answering this question.

According to both accounts, the answer is simply ‘yes’. The belief is valuable in

⁴¹ This is probably not the best version of RVT. Similar examples are easily constructed for any other restriction, and the basic point applies independently of how CVT is restricted.

virtue of being true. What does the RVT say? As the example stands, despite Ryan having a true belief, RVT says nothing. It attributes *no* value to this truth as it stands. This is enough to show that RVT abandons CVT. For us to determine the answer, we need to know a further fact: does the belief fall into the restricted class? If Ryan finds it interesting, it is valuable and, if not, it is not. Regardless of whether this is a good account of truth's value, it is clearly not an account of truth's *intrinsic* value. RVT *never* attributes value on this basis.

In sum: *No* true belief ever has value in itself according to RVT. This is because a belief does not gain, or possess, *any* value in virtue of being true. It possesses, or gains, value in virtue of being true *and* fitting within the relevant restricted class. The fact that a belief is true does not itself tell us *anything* about whether it is valuable. There is always a further question before its value can be correctly determined. This makes it clear that RVT abandons CVT. According to RVT, then, true belief is *not* intrinsically valuable.

Though this point seems obvious, we have seen that a number of defenders of RVT do not realise it. As RVT abandons CVT, it cannot be used as a *response* to the problem of trivial truths, nor can a defender of a teleological account of epistemic value endorse RVT. Indeed, *nobody* who believes that truth is valuable for its own sake can consistently endorse RVT.

To conclude this section, I will consider a possible objection to my argument.

This should also serve to clarify my claims.

Consider the following reply: RVT *is* compatible with the value of truth for its own sake. RVT values true belief intrinsically just like other responses to the problem of trivial truths do, such as the PFA. It just values true beliefs that fall

within a restricted class. So long they do, however, *that* true belief is valuable for its own sake.

This response is misleading. It is easy to interpret this as claiming that RVT is committed to the intrinsic value of *truth*. But that is not correct. It is correct that *one* of the necessary conditions for the value of a belief, according to RVT, is that it be true. But this is insufficient. It is also necessary that the belief fit into the relevant restricted class. For this reason, RVT cannot claim – as the objector asserts – that *truth* is valuable for its own sake.

Note also that it is not the case that RVT values *both* of these conditions independently but claims that it is better when they are combined. Neither condition is valued independently. This is clear from the fact that a true trivial belief is regarded as valueless and a false but class-relative belief is regarded as valueless. It is *only* when these two necessary conditions are fulfilled that the belief is valuable. Neither condition – including truth – stands alone or is valued alone. Nor, for that matter, would the account be better if they *were* valued independently. If RVT valued truth on its own, it would have to value trivial truths and then it would have all the same problems as the PFA.

In any case, RVT does *not* value truth independently. Consider a close analogy. According to the classical definition of lying, a person lies if and only if they (1) say something false, and (2) do so with the intention to deceive. From this, it follows that if someone *only* says something false (1) without the intention to

deceive (2), they do not lie. Similarly, if they say something with the intention to deceive (2), but it is not false (1), they do not lie.⁴²

According to RVT, a belief is valuable because (1) it is true, and (2) it falls within the relevant restricted class. From this it follows that a belief is not valuable if it is true (1) but doesn't fall within the restricted class (2). Similarly, it is not valuable if it falls within the restricted class (2), but is not true (1). (1) alone possess no value – put another way, a true belief alone is *valueless* – and hence a belief is not more valuable because it possesses (1), just as a statement is not more of a lie if it possess (1) but not (2).

Now return to the objection. According to the objection, RVT values true belief intrinsically in exactly the same sense as other theories. We can now see that this is incorrect. The PFA considers (1) to be sufficient for the value of a belief. This makes sense, since this is what it means to say that truth is intrinsically valuable. RVT rejects this claim. RVT *always* asks the further question: 'But does it fit into the relevant restricted class?' Thus, RVT is incompatible with CVT.

To reiterate, the fact that RVT abandons CVT does not show that it is incorrect. There are good reasons to abandon CVT. Chief among these is that – as was seen in our discussion of the PFA – (1) is implausible. As the trivial truths objection suggests, trivial truths are not at all valuable despite being true. RVT agrees with this claim, which in my view is a point in its favour. My aims have been different. RVT has been employed as a defence of the value of truth for its own sake. As we saw with Goldman and Pillar, it has also been used as an explicit response to the

⁴² For a good overview of the debate regarding the nature of lying, see Mahon (2008).

problem of trivial truths that is supposed to preserve the intrinsic value of truth.

This response fails. RVT simply abandons the idea that truth is valuable – or important – for its own sake.

Concluding Remarks on the Problem of Trivial Truths and CVT

The problem of trivial truths is a serious strike against CVT. The idea that trivial truths are either worth pursuing, or worth having, is implausible. CVT – by ascribing value to such truths – wrongly ascribes value to valueless beliefs. As it stands, this is a strong reason to reject CVT.

Various philosophers agree that trivial truths are a problem for CVT. Most think this problem can be avoided. I have discussed two prominent responses to the problem. Both fail. In principle, a response may be available that solves the problem. For this reason, my argument is not a knock-down argument. But there is reason to be sceptical of this prospect. Aside from weakening the value it assigns, it is difficult to see what else could be done without abandoning CVT.

We have seen that one promising possibility accidentally abandons CVT. This is perhaps symptomatic of the difficulty, and undesirability, of holding onto it. If my argument is right, then, in future, rejecting CVT should be no accident.

Chapter 3: Do We Value Truth For Its Own Sake?

‘All men by nature desire to know’

– Aristotle

Introduction

Various defenders of the idea that truth is intrinsically valuable (CVT) appeal to the alleged fact that we *value* truth for its own sake. As this is a psychological claim, I shall refer to it as:

PC: We in fact, or actually do, value truth for its own sake.⁴³

My purpose in this chapter is to argue that we should reject PC. We do *not* value truth for its own sake, nor do we value truth in any way that would support CVT.

PC and Its Uses

PC claims that – whatever its actual value – we in fact value true belief for its own sake. I shall understand ‘we’ as meaning *most people*. This seems to be how others who argue for PC understand this term. There is also reason to think that this is the *best* way to understand PC. Consider two alternatives.

First, PC might mean that the person who proposes it *herself* values truth for its own sake. This version of the claim is not worth discussing. One reason for this is

⁴³ It is important to emphasise that this chapter is concerned exclusively with whether we value truth *intrinsically*. For all I say here, it may be the case that we value truth *instrumentally*. Indeed, I suspect that this is the case. One reason for my focus is simply that arguments using PC to defend CVT seem more common than arguments using our attitudes towards truth to defend LVT.

simply that no defender of PC seems to be making such claims. Rather, words like ‘we’, ‘our’, and ‘us’ proliferate. This at least implies that they are referring not only to themselves. Further, if they *do* only mean to refer to their own values, the claim is not worth considering for the additional reason that it is uninteresting. PC is only interesting if it is meant to carry some weight, and, as we shall see, PC is often employed to support CVT. If it is to do so, it will need to be more general than just a self-report.

‘We’ could also be interpreted as meaning that *everybody* values truth. This seems to be what Aristotle is suggesting when he says that ‘All men by nature desire to know’. Though this claim *would* be interesting, it is highly implausible. Many people are similar in many ways, but there is also significant diversity. Some people, for instance, are fetishists about shoes, and others are sexually attracted to cars. For this reason, it would be surprising if anything turned out to be valued without exception. This claim is, if nothing else, too ambitious as an armchair claim. As Kornblith (2002, 150) writes:

It is important to note as well that any attempt to gain universal applicability by appeal to goals that all humans in fact have will almost certainly run afoul of the facts. Human beings are a very diverse lot; some of us are quite strange. It is hard to imagine making a plausible case for any particular goal or activity that is genuinely universally valued.

The claim that *most people* value truth avoids both of these problems. It is ambitious enough to be an interesting claim. If most people really do value truth for its own sake, then that may well support CVT. It is also the strongest claim

that can be plausibly – though defeasibly – made from the armchair. This is because it seems to be a fair assumption that, all else equal, *most* people are similar to us. As Frank Jackson (1998, 32) puts it, ‘my intuitions reveal the folk conception in as much as I am reasonably entitled, as I usually am, to regard myself as typical.’ Thus, if an example convinces me that I don’t value truth, then I have some reason to believe that it will convince most people.⁴⁴ Further, as noted, conceiving of ‘we’ as ‘most people’ also coheres well with what defenders of PC actually say.

To value truth intrinsically is to regard a true belief as valuable *in virtue of* being true.⁴⁵ To help illustrate this, imagine two people who value truth in different ways. Suppose John believes that truth is instrumentally valuable – he thinks that true beliefs will help him achieve his goals. If he found out this was not the case, he would no longer care whether his beliefs were true. Kate, who values truth intrinsically, would react very differently if she found out that true belief was not instrumentally valuable. Though she would now think that true belief has *less* overall value, she would still care about believing truly. This is because she values true beliefs *just because* they are true. As is probably clear, John does not value truth for its own sake. He does not care whether his beliefs are true *on the basis of* truth. He values useful beliefs, and happens to think that true beliefs are useful. Kate, on the other hand, values truth *for its own sake*. According to PC, people are

⁴⁴ Though, again, this will always be defeasible. It could turn out that, in some cases, you are atypical. Without reason to think this however, it seems reasonable to assume that you are in the norm. Most of us are average in most ways.

⁴⁵ As noted in Chapter 1, my understanding of ‘valuing’ is intuitive. Everything that I say in this chapter should be compatible with any plausible theory of valuing.

generally like Kate and not like John. I shall argue that we are *not* like Kate. Like John, we do not value true beliefs *just because* they are true.

The Plausibility of PC

PC is easy to motivate. Indeed, certain examples may make PC seem obviously correct. Two frequently employed examples are scientific and mathematical truths. Though many scientific truths are instrumentally valuable, we – or at least many of us – seem to value them for more than their *mere* usefulness. This idea can be put as follows: Certain instrumentally valuable truths would still seem valuable to us *even if* they were not practically beneficial. DNA sequencing may be one example. There is little doubt that progress can be made on various medical goals through this process. Nonetheless, even if this failed, there seems to be some value simply in believing truths about our fundamental nature. Grimm (2011, 723) makes a claim of this kind:

It is a commonplace among scientists... that certain questions *are* pursued simply for their own sakes, or from a pure desire to know. Indeed, whole fields of science, such as cosmology, seem to be driven by precisely by such a ‘pure’ desire. Even if no practical benefit were to emerge from such research, it is often said, the research would be worth pursuing.

Kitcher (2004, 216) makes a similar point:

Thoughtful and perceptive people throughout history have sometimes entertained a question not because the answer would enable them to do something practical, something they couldn’t have managed

without it, but simply because the question itself fascinated them.

When we view a completely pragmatic account of the sciences as inadequate, I think we're responding to this (almost?) universal sense of curiosity. Our aim ... [is] simply to answer the questions.

Many of us also care about certain truths that have no practical value at all.

Goldman makes this claim when he writes: 'The dinosaur extinction fascinates us, although knowing its cause would have no material impact on our lives.' Truths in pure mathematics are perhaps the paradigmatic example of this.⁴⁶ As Lynch (2004, 15-16) writes:

We care about the truth for more than just the benefits it brings us....
There are times in our lives when we simply want to know for no other reason than the knowing itself. Curiosity is not always motivated by practical concerns. Consider extremely abstract mathematical conjectures. With regard to at least some such conjectures, knowing their truth would get us no closer to anything else we want.

Mathematical truths are not the only useless truths we seem to value. Philosophy itself provides a further example. To quote Grimm (2011, 723) again:

[T]here is a long tradition in philosophy of claiming that philosophical questions are worth pursuing simply for their own sake – to the point where Aristotle could claim that it was precisely the practical *uselessness* of philosophy that made it so fine and impressive.

⁴⁶ I give an example of such a case on page 39.

Given the force of such examples, it is hardly surprising that PC is widely endorsed.⁴⁷ A few further examples – including more explicit statements by those quoted above – will help illustrate its popularity. Goldman (1986, 98) writes: ‘Truth acquisition is often desired and enjoyed *for its own sake*, not for ulterior ends’. And Lynch (2004, 12; emphasis added) writes: ‘Nobody likes to be wrong. If anything is a truism, that is. And it reveals something else *we believe* about truth: that it is good. More precisely, it is good to believe what is true.’ Alston (2005, 31) makes the following claim:

[T]he attainment of knowledge and understanding are ... *of intrinsic value*. “All men by nature desire to know”, said Aristotle, and this dictum has been reaffirmed by many of his successors. Members of our species seem to have a built-in drive to get the truth about things... and to understand how and why things are as they are and happen as they do.

Carl Hempel’s views resemble Alston’s. According to Hempel (1965, 333), we have an intrinsic desire for truth that is rooted in our ‘sheer intellectual curiosity, in [our] deep and persistent desire to know and to understand [ourselves] and [our] world. So strong, indeed, is this urge that in the absence of more reliable knowledge, myths are often invoked to fill the gap’. Hartry Field (2001, 120)

⁴⁷ I do not want to give the impression that PC is universally accepted. Two philosophers who deny that we value truth for its own sake are Stephen Stich (1990) and Hilary Kornblith (2002). Stich is probably the most interesting example. He claims that, though we value many things for their own sake – his main examples are health, happiness, and the welfare of one’s children – true belief is not among them. While his arguments are very different from mine, I agree with much of what Stich says and my project in this chapter is very similar to his. Like Stich (1990, 101), I believe that ‘once we have a clear view of the matter, most of us will not find any [intrinsic] value... in having true beliefs.’

claims that ‘What we desire is the infinite conjunction of all claims of the form ... ‘I believe “p” only if p’. According to Engel (2009, 188), ‘Truth is indeed a property of our beliefs that we value.’ Finally, defending both CVT and PC, Jones (1997, 423) writes: ‘It should be uncontroversial that we do and should value true believing.’

Arguments That Make Use of PC

Though the central aim of this chapter is to argue that PC should be rejected, there is more at stake than a mere psychological claim. PC is often used to support CVT. For this reason, PC’s failure – as well as being of independent interest – would undermine various arguments in favour of CVT.

PC has been used to support CVT in (at least) two ways. First, PC has been used as a *premise* in arguments for CVT. One example is the *best explanation argument*. This argument claims that the best explanation of the fact that we *value* truth intrinsically is that truth is intrinsically *valuable*. Kvanvig (2003) provides one instance of this argument. He (41) first defends PC:

[W]e do have a [purely intellectual] interest in the truth.... It is the nature of the interest to lack specificity: We do not have an individuated interest in the truth of the claim that our mothers love us, that the president is not a crook, that Wyoming is north of Mexico, and so on. What we have is a general interest in the truth, and that interest attaches to particular truths.... The default position for any truth is that our general interest in truth applies to it...

Kvanvig (42) then claims that CVT is the best explanation of this fact: ‘I propose that the conclusion that truth is intrinsically valuable is the best explanation of the data before us’.

Others claim that PC provides direct *evidence* for CVT. In these cases PC is not used as a premise, but is instead taken to directly support CVT. The form these arguments take is heavily influenced by the theory of value being employed.

According to some versions of subjectivism, for instance, PC is *sufficient* to establish CVT. This is because, on these theories, what possesses value is determined by an aspect of the agent’s psychology. For example, what she cares about, what she values, what she desires, or what her second order desires are. For this reason, an accurate psychological claim can directly reveal what *in fact* has value. Thus, assuming PC is accurate, the fact that we *value* – or care about or desire – true belief for its own sake demonstrates that true belief is *valuable* for its own sake.

Horwich (2006, 351) seems to argue along these lines. He implies that valuing true belief intrinsically is just what it *is* for true belief to be intrinsically valuable:

[T]here is widespread sentiment to the effect that certain items of knowledge are desirable regardless of any practical use to which someone might decide to put them. Knowledge is valuable, as we often say, “for its own sake”. In the second place, without some such assumption, it would be hard to justify our pursuit of truth in fields of inquiry such as ancient history, metaphysics, and esoteric areas of mathematics—fields that may not be expected to have any pragmatic

payoff.... For these reasons I think we should acknowledge that true belief has a *non-instrumental* value—a value for its own sake.

Objectivists about value also invoke PC to support of CVT. According to many versions of objectivism, PC *cannot* tell us directly what has value. This is because, at least in principle, there is always the possibility that our beliefs regarding what has value fail to correspond to the external facts – that is, to what is *actually* valuable. Nonetheless, even for advocates of these versions of objectivism, PC still often plays the role of evidence for CVT.

Lynch provides one example of this. Lynch does not believe we can move directly from the fact that we *value* true belief, or that we *believe that* true belief is valuable, to the claim that true belief *is* valuable. He writes, after defending PC: ‘Of course, none of this proves that everyone accepts this, or for those of us that do, that we are *right* to accept it’ (2004, 19). Nonetheless, he believes that: ‘[T]he fact that we care about something is very good evidence that ... it [is] worthy of caring about’ (15).

Lynch’s final point is hard to object to, even if – for an objectivist – it is somewhat depressing. Though there may be an unbridgeable gap between what we value and what is valuable, it seems obvious that we can’t even begin to talk about what is valuable without invoking our beliefs about what has value. Nobody seems to be arguing about the intrinsic value of plastic. This is not even a candidate for being valuable. And surely a big part of the explanation for this is that we don’t happen to value plastic. If we were a very different kind of animal, perhaps plastic would play a prominent role in our philosophical disputes. This is somewhat dispiriting if objectivism is true. This is because it could be the case

that we are an extremely limited species who will never grasp or appreciate what is actually valuable. Perhaps we are doomed to always miss out, just as ants will never experience happiness. Though it would be depressing, there is not much we can do about this if it is true.

This links with a final point. There is good reason to think that the relevance of PC's accuracy for assessing CVT is more fundamental than just the fact that PC has been used in arguments for CVT. It seems highly plausible – as a number of the quotes suggest – that part of the temptation to endorse CVT in the first place, part of its intuitive appeal, stems from the fact that PC seems correct. If a person values truth for its own sake, they seem far more likely to endorse CVT. Nobody wants to believe that what they regard as valuable is actually worthless. If, on the other hand, it turns out that we do not value truth intrinsically, this may make CVT seem less attractive and rejecting it less counter-intuitive.

My Argument and a Concern

Despite different assumptions, the above arguments all require PC to be *correct*. My purpose in this chapter is to argue that PC should be rejected. We do *not* value true belief for its own sake. The first part of my argument is negative. I argue that two defences of PC do not actually provide it with any support. This task is worth undertaking because both arguments seem compelling at first glance. Indeed, they have often been treated as decisive. With this task out of the way, I then argue directly that we do not value truth for its own sake.

My positive and negative arguments are not entirely independent. Unless the obvious arguments for PC are dispensed with, there will seem to be easy

objections to – or strong considerations that outweigh – my direct arguments against PC.

To conclude, I return to the arguments that make us of PC to defend CVT. I show that, once PC is rejected, these arguments are entirely unconvincing. Further, I argue that rejecting PC erodes much of the motivation for endorsing CVT.

A Concern

Before getting underway, there is a possible concern that is worth briefly addressing. The question of whether we *value* true belief intrinsically is ultimately a descriptive – or empirical – question. This means that many of my claims are speculative. PC cannot finally be settled until the appropriate empirical work has been done. Though this is undeniable, and perhaps worrying, a few points bear mentioning. First, it can be shown from the armchair that arguments and examples that are supposed to prove a psychological thesis fail to do so. Interpretation *is* an armchair activity. Much of this chapter is concerned with this task. Thus, even if armchair speculation is worthless, many of my arguments would be unaffected. Nor would this affect my arguments regarding which philosophical claims are supported or undermined by which empirical claims. The last section of this chapter argues that, *if* PC is false, this would undercut various arguments for CVT, as well as much of its plausibility.

Further, this is not a unique problem for me. As we have seen, various philosophers make the empirical claim that PC is *correct*. As far as I can tell, these claims are being made without a shred of systematic data; at most, other philosophers are cited. As such, I am at least on equal footing with defenders of PC. And, of course, this concern is not a unique problem for arguments about PC.

My point is not that this kind of speculation is a good thing, nor that it can be justified. My claim is just this: If armchair speculation cannot be justified, we are all going down with the ship together.

First Negative Argument: Tempting Examples That Prove Nothing

Certain truths seem to be valuable for their own sake – we seem to *value* them for their own sake. I noted two examples: We seem to value certain scientific truths over and above their instrumental value and we seem to value certain truths despite them entirely lacking instrumental value – mathematical truths being the paradigmatic example. These examples are frequently utilised by defenders of PC. Indeed, these examples are sometimes taken to be adequate – or sufficient – to establish PC.

Examples like these are difficult to dispute. As such, my aim in this section is not to discredit them directly. Instead, I argue that these examples alone, regardless of our reactions to them, tell us almost nothing about the accuracy of PC. It is a mistake to suppose that substantial conclusion can be drawn from them. Further, the reason this is a mistake highlights a requirement that any argument for PC must satisfy.

Why the Examples are Inadequate

The central reason that scientific and mathematical examples offer little support to PC is that they lack the required level of generality. According to PC, we value truth intrinsically – we value beliefs *because* they are true. This means that, if a belief is true, we value it (at least a little). The examples discussed show, at best, that we *sometimes* – when it concerns particular areas of interest – want true

beliefs just to believe the truth. This is the wrong kind of conclusion. This is because, even if this psychological claim is true, it does little on its own to show that we want truth *for its own sake*.

This can be seen if we consider the numerous ways that we could hold our positive reactions to such cases fixed, but nonetheless undermine PC by changing the surrounding facts. For example:

The Inconsistent Truth Hater: Imagine that a person, call him Peter, values every scientific and mathematical truth ever cited by a philosopher. And suppose that he wants to believe these truths just because he wants to believe the truth about science and mathematics. But now imagine that Peter *despises* every other truth it is possible to believe. For instance, he actively avoids believing the truth about every other area when this is possible – he reads newspapers that he knows are unreliable, for example – and he hates the fact that he can't avoid forming innumerable true beliefs no matter what he does.

It seems absurd to claim that Peter values truth *for its own sake*. He values truth about areas that happen to interest him. And, in these areas, he wants the truth just to have the truth. But he clearly doesn't value truth *intrinsically*. He wants the truth *because* it meets a particular criterion – or appears in a particular context – not for its own sake. Despite not valuing truth for its own sake, Peter would agree that the truth is worth having in all the common examples cited to support PC. This alone is enough to show that citing particular examples is entirely inadequate to establish PC, no matter how compelling the examples are.

This last point also reveals an adequacy requirement that any argument must meet if it is to support or establish PC. Our interest in truth must be shown to be sufficiently *general*. This is a necessary condition. Unless we value a sufficiently general class of truths, PC fails. There is simply no reason to accept it if we have not been convinced that it is really *truth itself* that we value. It can't be the case that, like Peter, we only care about some truths, but don't want most others. If this was the case, there would be no reason to say that we value *truth* for its *own* sake.

Defenders of PC seem to be aware of this requirement, and often claim that we do value truth in this general way. Recall Hartry Field's (2001, 120) claim that, 'What we desire is the infinite conjunction of all claims of the form ... 'I believe "p" only if p'. Kvanvig (2003, 41) makes a similar claim when he writes: 'What we have is a general interest in the truth, and that interest attaches to particular truths in the manner of instantiation in predicate logic. The default position for any truth is that our general interest in truth applies to it'.

Sufficient generality may not be the only burden that arguments for PC need to meet. But it is one obvious burden. It is also all that I need because, as I argue below, it *cannot* be met. The main point to note for now, however, is that citing the usual examples – no matter how compelling they seem – does almost nothing to establish PC.

Second Negative Argument: A Failed General Argument

This section discusses a more promising argument for PC. This argument, if plausible, has no problem with the generality requirement. According to Michael Lynch, we have a 'basic preference' for true belief. A basic preference is 'a

preference for something that can't be explained by our preference for other things' (2004, 15). Applied to true belief, this amounts to agreeing with the following claim: 'where the belief that p and the belief that not- p have identical instrumental value, it is better, just on grounds of truth alone, to have the true belief rather than the false one' (2004, 21). Lynch believes that most people would agree to this claim. If Lynch is right, then his argument would give us the right kind of conclusion. It would show that the generality requirement can be met.

Lynch also believes that this would support CVT. This is because, according to Lynch, having a basic preference for true belief is good evidence for the claim that true belief is intrinsically valuable. He writes: '... the fact that we care about something is very good evidence that ... it [is] worthy of caring about. Accordingly, if you care about truth for its own sake, then you presumably believe our last truism, namely that truth is worthy of caring about in just that way' (2004, 15).

Lynch defends this psychological claim through two thought experiments: the *Experience Machine* and *Russell World*.⁴⁸ These are both prominent thought experiments in their own right.⁴⁹ Both are also of interest for truth's value independently of their use by Lynch. This is because – whether it is cashed out as

⁴⁸ These are both explained in detail below. I delay this exposition until after I have explained some general problems that thought experiments can suffer. Pairing my explanations of the thought experiments with my criticism of them allows the material to flow more smoothly.

⁴⁹ This is particularly true of the Experience Machine. See Feldman (2011) for a discussion of the various uses that this example has been put to. As far as I can tell, Lynch is the only philosopher in this debate who has used the Experience Machine in a *sustained* way to support PC and CVT. Others, such as Kvanvig (2003, 41-2), have used similar examples for the same purposes, however. My arguments should also apply, *mutatis mutandis*, to all similar examples.

a basic preference or not – they seem to demonstrate *something* important about truth's value. Both make it very tempting to conclude that we either value true belief for its own sake, or that truth is intrinsically *valuable*. Neither conclusion is warranted. The Experience Machine and Russell World tell us *nothing* about the value of truth. Though I focus on Lynch, my arguments concern the thought experiments themselves. As a result, if my arguments are correct, any uses of these thought experiments to support PC or CVT will fail. As it is not obvious that these thought experiments fail at first glance – and the temptation to employ them for these purposes is powerful – they are worth discussing in detail.

Lynch (2004, 21) expresses his reasons for thinking that the Experience Machine and Russell World support PC as follows:

This... is what I mean by saying we can learn about what we believe from these science-fiction stories: for many of us, our intuitive reactions to these cases suggest that we have a basic preference for the truth; that this preference matters to us; and thus that *we believe* that truth is worth caring about for its own sake. If you prefer not to live in the vat, or in a Russell world, then you implicitly accept that where the belief that *p* and the belief that not-*p* have identical instrumental value, it is better, just on grounds of truth alone, to have the true belief rather than the false one.⁵⁰

⁵⁰ Lynch sometimes refers to the device in this example as a vat. This ambiguity is not important for our purposes. As Lynch sets up the example, both entering a machine and being a brain-in-a-vat would have the same implications.

Success and Failure

In order to assess whether the Experience Machine and Russell World support PC, it is necessary to first examine what would make them successful or unsuccessful.

There are (at least) two requirements a thought experiment needs to satisfy to support PC. Call these the *minimal standards of success*. First, we need to judge that true belief is indeed valuable in the example. If an example is proposed and true belief seems worthless to us in the situation described, it is clearly unsuccessful. Second, it needs to be the case that our judgement cannot be explained by other features of the case. It has to be clear that our intuitions are actually reacting to the perceived *intrinsic* value – rather than, say, the instrumental value – of true belief. Suppose a thought experiment is proposed and true belief seems valuable in the situation described. This will not support PC if the content of the true belief would obviously make us happy. This is because it will be unclear whether we are valuing true belief *because* it is true or because it would make us happy. It needs to be relatively unambiguous that our reaction gives us some reason to favour PC over alternative claims.

A thought experiment can fail for various reasons. These include reasons that are simply the converse of the above minimal success conditions. Since I will be arguing that the Experience Machine and Russell World fail, it will be useful to spend a bit of time on a few of these reasons – and one in particular. Illustrating these problems with simple examples will make my criticisms of these examples much clearer. It will also make it plain that these problems warrant rejecting Lynch's argument.

The most obvious reason a thought experiment fails to support a proposition is if it invokes an intuition that directly contradicts that proposition. Suppose I am defending the view that pain is not bad for the person who experiences it. To motivate this I ask you to imagine the following case:

Mark, a mathematician, has been working on a problem all day. He decides to take a walk to clear his head. As he is walking he has a breakthrough. Distracted by this, he forgets to look when he crosses a road and gets hit by a car. This breaks many bones and collapses his lungs. Lying on the ground, Mark is trying to scream out in pain but a lack of oxygen makes his cries almost inaudible.

I then say: ‘as this case demonstrates, being in pain is not necessarily bad for the person who experiences it’. It is clear that this example fails to support my claim. The situation described seems horrible for Mark.

Thought experiments rarely fail in this obvious way. More commonly, an example will simply fail to generate the desired intuition. Suppose I am arguing that all current accounts of knowledge are insufficient. I claim that a necessary condition of knowledge is being able to say any known proposition backwards. This, I claim, can be seen by the fact that we want to deny ascribing knowledge in cases where a person cannot do this. I ask you to imagine Jane, who has a justified, true belief about p – plus whatever further condition the reader prefers. Suppose p is that the sun rose today. During an interview with Jane, we ask her various questions about her claim to know p . She answers in a way that satisfies all current accounts of knowledge. But when we ask her to say ‘the sun rose today’

backwards she fails. She tries a few times but keeps messing it up. I then ask, rhetorically: ‘Does Jane *really* know that p ?’

This example fails to support my claim about knowledge. Unlike the Mark case, it does not elicit the opposite intuition. The fact that Jane cannot say p backwards does not seem to *support* that she knows p . If she could say it backwards, I would still think she knew p . The backward saying of the proposition just seems irrelevant.

A Less Obvious Reason

Thought experiments can fail in more subtle ways. A thought experiment fails when the intuition it elicits – even if it agrees with the authors overall assessment of the case – is not the *right* intuition to support the authors claim. Put more carefully: A thought experiment fails if it does not track the intuition required to support p , where the thought experiment is used to support p . Call this a *tracking failure*. I will argue that both the Experience Machine and Russell World – when employed to support PC or CVT – exhibit this problem.

When a tracking failure occurs, the thought experiment does not support p . The reason for this is simple. Our assessment of x does not support p if x and p are unrelated. Where this is the case, our evaluation of x simply tells us nothing about p .

There are two broad classes of tracking failure. The first is *direct* tracking failure. This occurs when it can be shown that our assessment of the case is reacting to features that are irrelevant to assessing p . The second class of tracking failure arises from uncertainty. For convenience, we can call this *indirect* tracking failure. This occurs when it is *unclear* whether we are reacting to features of the case

relevant to assessing p , or whether we are reacting to irrelevant features of the case. Indirect tracking failure, then, is the converse of the second minimal success condition. Indirect tracking failure has the same consequence as direct tracking failure. Since it is unclear which features of the case are being reacted to, the case does not support p because it gives us no reason to endorse p . As with direct tracking failure, the example does not allow us to conclude anything about p .

A Tracking Failure Test

There are obvious instances of tracking failure. But – unlike the other reasons discussed – it is often not obvious. The Experience Machine and Russell World are instances of non-obvious tracking failure. Neither tells us anything about whether we value true belief for its own sake.

As this is not obvious, we require a test. Suppose a thought experiment is presented that is intended to track our intuitions about the desirability of x . It is designed to demonstrate that we desire x for its own sake. The example presents a situation where we lack x . We are supposed to think that this situation is bad. And we do think this situation is bad. Does this result support the badness of x ?

I propose the following tracking failure test: If we re-imagine the thought experiment so that we now possess x – and all else remains equal – then we should think the re-imagined example is better than the original version. That is, we should think it is better if we have x than if we don't. If, on the other hand, we think the re-imagined scenario is no better than the original – even though we now have x and all else is equal – we should conclude that the desirability of x was not what we were reacting to in the original example. For this reason, the original example does not support the desirability of x .

Consider an example. Suppose I am arguing that x is intrinsically bad. I design a thought experiment to motivate this claim. You are asked to imagine the following:

A torturer is causing you extreme psychological and physical pain. She does various things to you. She hits you with a crowbar; plays loud music at random times; deprives you of sleep; calls you names; and tells you the date of her anniversary.

The situation described is clearly bad. From this, I conclude that the example supports my claim. But suppose that the example was designed to demonstrate that somebody telling you the date of her anniversary is bad. Do our negative intuitions support this conclusion?

My conclusion would be a mistake. The example does not support the claim because it fails to track the appropriate intuitions. My test demonstrates this. To implement it, we need to re-imagine the case. Imagine that – instead of telling you the date of her anniversary – the torturer tells you the date of her birthday. Everything else is the same. If my claim was correct, then this change should make the situation (at least slightly) better – or less bad⁵¹. After all, a supposedly intrinsically bad feature of the case has been removed. But the re-imagined situation seems just as bad as the original. This shows that our intuitions about the badness of the original example were not tracking the badness of the torturer telling us her anniversary. Our negative evaluation needs to be explained by something other than my claim that this is bad. For this reason, the example does not support my claim.

⁵¹ Assuming that telling you the date of her birthday is not also bad.

It might be thought that this manoeuvre is suspicious. The following idea might seem plausible: if you change any single feature of a thought experiment where multiple features are doing work, it will always be the case that we get the same intuition after the single feature has been altered.

This idea would undermine the test if correct. But it is not. Consider the same example with a different feature adjusted. Imagine that the torturer does not cause you physical suffering – she does not hit you with a crowbar. According to my test – if our intuitions are tracking the badness of physical suffering – the re-imagined example will seem better – or less bad – than the original example. And it seems obvious that this situation – although still bad – *is* less bad. Being physically and psychologically tortured seems worse than being only psychologically tortured. Thus, if the original example was designed to support the badness of physical suffering, it would pass the test. Our intuitions are – at least partly – tracking the badness of physical suffering.

My Argument

I shall argue that the Experience Machine and Russell World are both analogous to the anniversary case. For this reason – among others⁵² – neither supports PC or CVT. Lynch presents two bad situations. Both involve a lack of true belief. From these two points, Lynch concludes that we believe the situations are bad *because* they involve a lack of true belief. This is a mistake. Unlike the physical suffering example, our negative reactions to the Experience Machine and Russell World tell us nothing about the value of true belief. This is because – like the anniversary case – our intuitions about these cases are not tracking the value of true belief at

⁵² I discuss two problems for each example.

all. Thus, our intuitive reactions to these examples do not support the claim that true belief is, or that we regard it as, intrinsically valuable.

The Experience Machine

The first thought experiment Lynch (2004, 15) presents is the *Experience Machine*:

Suppose I had a machine that allowed you to experience whatever you want. Once inside, floating in the tank, you live in a virtual reality of your own design – one filled with experiences of adoring friends, marvellous adventures, spectacular food, good sex, and deep conversations. None of it would be real, of course, but it would seem to be. It could even be arranged so that once inside the machine, you completely forget that you *are* inside the machine. There is only one catch. Once inside, you can *never come out*. Would you do it?

Lynch expects the answer to this question to be ‘no’. This is a reasonable expectation. Intuitions against entering the machine seem to be fairly robust.⁵³ I will grant that we should not enter the machine. This is open to doubt, but the ambitious project of rejecting it is beyond what I require. My claim is that the Experience Machine does not tell us anything about the value of true belief. It could be the case that it tells us about other things – such as whether hedonism is true.

⁵³ This at least seems to be the case among philosophers. It would be hard to explain the prominence of the example otherwise. There is also support for the idea that most people would not enter the machine – though there is much interesting debate about why – in the experimental philosophy literature on this subject (e.g. De Brigard 2010 and Weijers 2013).

Lynch (16) believes that our reaction against entering the machine supports PC.

He draws the following lesson from this example:

Most of us will probably say no [to entering the machine]. While we certainly wouldn't mind being in the machine for a few hours or even weeks, we wouldn't want to spend the rest of our life in a virtual world. Others of us, whose actual lives are filled with tragedy or poverty, might be more inclined to opt in for the long term. But even so, most would prefer having their problems truly disappear to living a life where they only *seem* to disappear. The machine produces beautiful illusions, but we want more than illusions. We want the truth, warts and all.

As this remark illustrates, Lynch believes that our desire not to live in a virtual world is explained by our desiring true beliefs. I will argue that no such conclusion can be drawn from our negative reaction to the Experience Machine.

First Argument Against the Experience Machine – Direct Tracking Failure

The Experience Machine does not support PC or CVT because – in this context – it is guilty of both classes of tracking failure. That is, direct and indirect tracking failure. I begin with direct tracking failure.

As a reminder, this occurs when our assessment of the case is reacting to features that are irrelevant to assessing the claim the example is employed to support. It is not obvious that this example has this problem. My test shows that it does.

Begin by imagining a different version of the Experience Machine. Call the adjusted version TEM*. TEM* includes all the features of the Experience

Machine, but TEM* has an additional feature that the original example lacks. In TEM*, the machine is set up so that, once inside, any belief you already have or form once inside the machine contains a tacit prefix – ‘According to my experiences, ...’. Prefixing your beliefs in this way requires no energy on your part and is of no inconvenience. The prefix is just an automatic part of any relevant belief.⁵⁴ Indeed, you don’t even really notice this fact about your beliefs – or, at least, you notice it no more than you now notice that many of your beliefs refer to the actual world. Currently, I believe that ‘I am typing on a computer in an office’, but I don’t generally think ‘In the actual world, I am typing on a computer in an office’. In this way, your beliefs in TEM* feel the same as they currently do. Everything else in TEM* is the same. There are still ‘experiences of adoring friends, marvellous adventures, spectacular food, good sex, and deep conversations’ (2004, 15). It is still the case that:

None of it would be real, of course, but it would seem to be. It could even be arranged so that once inside the machine, you completely forget that you *are* inside the machine. There is only one catch. Once inside, you can *never come out*. Would you do it?

It seems to me that TEM* is no more appealing than the Experience Machine. The features of the Experience Machine that make it unappealing – whatever these are – still seem to be present in TEM*.

Notice, however, that in TEM* most of your beliefs will be true. It will be true that ‘according to my experiences, I am a bestselling author’, or that ‘according to my experiences, I am a great philosopher’, or that ‘according to my experiences, I

⁵⁴ By ‘relevant belief’ I mean beliefs about things external to you.

have adoring friends and a loving wife'. It will be the case that all your beliefs prior to entering the machine that were true will still be true – although in adjusted form.

If Lynch is correct that our reactions to the Experience Machine are tracking the badness of false beliefs – and our desire not to get in is explained by our desire for true beliefs – then it should be the case that the original example seems worse than TEM*. Knowing that your beliefs will be true once you enter the machine, are you now tempted?

I suspect those who find the prospect of living in the Experience Machine unappealing would also find the prospect of living in TEM* unappealing.

Assuming I am right, this result shows that our intuitions about the badness of the Experience Machine are not tracking the value of true belief. This means that the lesson Lynch draws from this example – ‘The machine produces beautiful illusions, but we want more than illusions. We want the truth, warts and all’ – is mistaken. Wanting the truth does not explain why we don’t want to get into the machine. We still don’t want to get into the machine even if we have the truth. Since the Experience Machine does not track the value of true belief, it cannot support PC or CVT.

Second Argument Against the Experience Machine – Indirect Tracking Failure

If not because we want true beliefs, why do people find the Experience Machine unappealing? This question points to a second problem with this example: it suffers from indirect tracking failure. This occurs when it is *unclear* whether we are reacting to features of the case relevant to assessing a claim, or whether we are reacting to irrelevant features of the case.

The Experience Machine displays this vice because a negative reaction to the situation described could be triggered by numerous factors and explained in various ways. For this reason, we can never conclude from a negative reaction to this example that it is reaction to any *particular* feature of the case. Put another way, it is always open to reasonable doubt that any one particular factor is the important part of the explanation for our negative intuitive reaction. It is – just because of the nature of the case – unclear which of our values, desires, or beliefs are doing the work.

A sample of the possible preferences that could explain a negative reaction to the Experience Machine includes: a desire to be respected, admired, idolised or feared; a desire to help make other people's lives better; a desire for power over people; a desire to change the world; a desire to spend time with family and friends; a desire for more than mere experiences – i.e. to be a bestselling author rather than just to have the experience of being a bestselling author. The list could go on and on. But the central point should be clear: The move from a negative reaction to the Experience Machine to a claim about any particular preference, let alone a 'basic preference', is always going to be uncertain and open to reasonable doubt.

Since it is unclear which features of the Experience Machine are causing the negative reaction, the example does not support PC or CVT. It gives us no reason, by itself, to endorse CVT rather than any other claim compatible with the many possible explanations of our reaction. As with direct tracking failure, then, the example does not allow us to conclude anything about the correctness or plausibility of CVT.

In fairness to Lynch, he is aware of this second problem. He writes (2004, 17):

Some may protest that we want more than mere experiences out of our life, and it is *this* fact – not any preference for truth – that makes us prefer the real world over the vat.

His motivation for proposing Russell World is – at least partly – to undercut this objection. In the next section, I argue that, although Russell World does not suffer indirect tracking failure, it does suffer direct tracking failure. It also has another problem: it fails to meet the minimum success conditions.

Russell World

The second thought experiment Lynch proposes to support PC is based on a situation conceived by Bertrand Russell – that the world actually began yesterday.⁵⁵ Lynch (2004, 17) writes:

Suppose that, unbeknownst to us, the world began yesterday – it *seems* older, but it isn't. If I really lived in a Russell world... almost all my beliefs about the past would be false. Yet my desires would be equally satisfied in both worlds. This is because the future of both worlds unfolds in exactly the same way. If I believe truly in the actual world that if I open the refrigerator I'll get a beer, then I'll get a beer if

⁵⁵ In Russell's original example, he considers whether the world began five minutes ago. His point was that nothing in the future or present could ever disprove this hypothesis. He writes: 'There is no logical impossibility in the hypothesis that the world sprang into existence five minutes ago, exactly as it then was, with a population that "remembered" a wholly unreal past. There is no logically necessary connection between events at different times; therefore nothing that is happening now or will happen in the future can disprove the hypothesis that the world began five minutes ago' (Russell 1921, 159-60)

I open the refrigerator. Since events in the Russell world are just the same as in the actual world once it begins ticking along, I will also get that beer in the Russell world if I open the refrigerator, even if (in the Russell world) I believe falsely that I put it there yesterday. In other words, whatever plans I accomplish now, I would also accomplish if the world had begun yesterday, despite the fact that in that case, my plans would be based on false beliefs about the past.

Russell World seems to avoid the second objection to the Experience Machine. For example, our reactions cannot be explained by the fact that we desire more than *mere* experiences. In Russell World, my friends would be real people, anybody I helped would actually be helped, and I could change the world.

Another point in favour of Russell World is that it seems to succeed in separating out intrinsic from instrumental value. This ensures that, if the example suggests that we value true belief, it suggests that we value true belief for its own sake. Lynch explains this when he writes: ‘whatever plans I accomplish now, I would also accomplish if the world had begun yesterday, despite the fact that in that case, my plans would be based on false beliefs about the past.’

Despite Russell World allowing more than mere experiences – and despite our false beliefs about the past having no instrumental disutility – most of us, according to Lynch, would not want to live in this world. He writes (2004, 18):

Yet, given the choice between living in the actual world and living in the Russell world, I strongly prefer the actual world. Of course, once “inside” that world, I wouldn’t see any difference between it and the real world; in both worlds, after all, events crank along in the same

way. For the fact remains that thinking about the worlds only insofar as they are identical in instrumental value, *there is a difference right now* between the two worlds that matters to me. Even when it has no effect on my other preferences, I – and presumably you as well – prefer true beliefs to false ones.

First Argument Against Russell World – Direct Tracking Failure

Russell World may well avoid indirect tracking failure. But it does not avoid direct tracking failure. As with the Experience Machine, it is not obvious that Russell World has this problem. My test shows that it does.

Lynch elicits the intuition that Russell World is bad – or suboptimal – by asking whether we would rather live in the actual world or in Russell World. Since we would rather live in the actual world, and because we have true beliefs about the past in the actual world, Lynch concludes that our preference for the actual world – and against Russell World – tracks true belief.

This is not so. To see this, imagine that the choice is between the actual world, Russell World and an alternative called RW*. Just like in Russell World, the world in RW* began yesterday. The difference is that in RW* you also come to *believe* that the world began yesterday. For this reason, you no longer have false beliefs about the past. You believe, truly, that none of these beliefs refer to an actual past.

If the value of true belief explains our preference for the actual world over Russell World, then we should regard the choice between the actual world and RW* as equivalent. After all, neither involves any false beliefs about the past. But it seems to me that the actual world is still preferable to RW*, and I suspect that most who

prefer the actual world to Russell World would agree. Whatever causes our negative reaction to Russell World still seems present in RW*. If this is correct, it can't be the value of true belief that is causing our intuitions against Russell World. And if Russell World does not track the value of true belief, it cannot support PC or CVT.

Second Argument Against Russell World – Standard of Success Failure

The second problem for Russell World is that the example – even if it gets the reaction Lynch wants for the reasons Lynch wants – fails to do what is required to support PC or CVT.⁵⁶

Marion David discusses this problem. It arises from the structure of the example.

This is, as David (2005, 299) writes:

... my actual beliefs are held fixed, i.e., it is stipulated that I have the beliefs that I *actually have*; and it is asked whether I would prefer to live in the normal world or in the alternative world of the scenario.

(my emphasis)

This form is problematic for the following reason: even if we don't want to live in Russell World because we want true beliefs, this does not demonstrate that we believe CVT. A different view is equally compatible with our rejection of this world. This is that we want *our own* beliefs to be true. There is nothing inconsistent about valuing the truth of your own beliefs but not valuing the truth *for its own sake*. An example – which is similar to *the inconsistent truth hater* – illustrates this distinction:

⁵⁶ My arguments in this this section could be extended to the Experience Machine.

Imagine a person – Chris – who is lazy, anti-intellectual, irrational, and prejudiced. Chris is, nevertheless, attached to the particular beliefs he holds. He desires that his beliefs be true. He has no interest in verifying his beliefs, inquiring further into the areas that he holds beliefs about, or in trying to attain further beliefs. He wants his irrational and prejudiced world view to be true, but he doesn't care at all what else is true. He does not care about the truth *as such*.

Notice that – because he cares about the truth of his own beliefs – Chris would no more want to live in Russell World than somebody who *does* believe CVT.⁵⁷ But it is clear that Chris does not value true belief in the way CVT specifies, nor does he capture the spirit of CVT. I doubt that Lynch had any intention of commemorating Chris when he wrote a book championing the significance of truth.

David (299) nicely expresses the second problem with Russell World when he writes:

[Our reaction to Russell World], it seems to me, ought to be irrelevant to the issue at hand. The expected response, namely that I prefer to live in a normal world, indicates at best that I want my beliefs, the beliefs that I actually have, to be true, that I want to live in a world in which the beliefs that I actually have are true. This does not indicate that I want to believe whatever is true. To show the latter, especially to show that I care about believing what is true for its own sake rather

⁵⁷ Granting, for the sake of argument, that we don't want to live in Russell World because we value true beliefs.

than for its instrumental value, Lynch has to present a scenario where I am in *one* world but with *two* radically different belief-sets, one made up of true beliefs and the other made up of false beliefs, which nevertheless are equally valuable on purely instrumental grounds; and he has to convince me that I prefer to have the set with the true beliefs, even though its cash-value is no higher than the one with the false beliefs.

Our reactions to Russell World are not adequate to support PC or CVT – even if they are the ones Lynch is after. This example simply fails to test for the intuitions that would require or compel us to accept either claim.

In sum: the Experience Machine and Russell World fail to support PC, and fail to demonstrate that we have a basic preference for true belief. The examples that Lynch employs to support his basic preference argument are unconvincing. Thus, we have been given no more reason to accept PC or, by extension, CVT. In the next section, I argue that we may be able to make a stronger claim. I then turn to some direct reasons to reject PC.

A Neutral Result?

It might be thought that the failure of the Experience Machine and Russell World to support PC is a neutral result: Though they give us no reason to accept PC, their failure also gives us no reason to reject PC. We are exactly where we started. I will end my discussion of these examples by noting a reason that this may not be the case. I do not rely on this claim, as our reactions to the Experience Machine, Russell World and their alternatives can be explained in various ways. Nonetheless, the point seems interesting enough to note.

In the original versions of both examples, we did not have true beliefs. In the adjusted versions, we do have true beliefs. All else seems to be equal between the examples. According to PC, we value true belief for its own sake. Thus, the adjusted examples contain something that we value which is lacking in the original examples. For this reason, if PC is correct, it seems that we would prefer the adjusted examples to the originals. But we do not. This is analogous to the anniversary case. When the example did not seem worse after changing the anniversary to a birthday, the natural conclusion was that we did not regard telling somebody your anniversary as bad. It did not seem that we had been given no reason one way or the other to think that telling somebody your anniversary was bad. Even though the test only demonstrates that our intuitions are not tracking the relevant feature in the original case, it is plausible that if the feature was something we valued our intuitions would have tracked it in the re-imagined case. But true belief didn't seem to make a difference. This might suggest that we do not value true belief intrinsically. If we did, then situations would seem better, all else equal, when true belief is present.

Against PC

Thus far, I have made two claims: (1) Though we value certain truths, these cases do little to establish PC; and (2) though there are compelling thought experiments that seem to establish PC – or at least one promising version of PC – these are, at best, indeterminate.

I believe stronger claims can be justified. This section argues directly that PC should be rejected. I present three considerations in favour of this conclusion. These all draw, in different ways, on the problem of trivial truths. As discussed in

the previous chapter, certain truths are so trivial that they seem to be entirely valueless. Though this problem is generally put forward as a direct challenge to CVT – and is convincing in this guise – I argue here that trivial truths also undermine the claim that we *in fact* value true belief for its own sake.

Briefly put, the trivial truths objection as a challenge to PC is this: At any point in the day, you could acquire numerous true beliefs. Most of these true beliefs would be trivial. You could, for example, count the hairs on your arm, or count the pieces of thread on a piece of carpet. As well as seeming valueless, we do not seem to *value* these truths. This is suggested by the fact that we never, at least for no practical reason, try to acquire them. Yet trivial truths are unquestionably truths. Thus, if we valued truth intrinsically, we would value these truths. But there is little indication that people *do* value them. As such, there is little reason to think that most people value true beliefs for their own sake.⁵⁸

Two Generality Problems

My first argument – which has two components – is that no plausible psychological claim about our attitude to truth will be sufficiently general to vindicate PC. Though it is difficult to prove this definitively, considering the kind of generality discussed above gives us good reason to think it is correct. Further, I will also consider the kind of generality built in to PC – that is, that *most people* value true belief. Both kinds of generality are believed to hold by defenders of PC.

As a reminder, the first kind of generality claimed on behalf of PC is the following. People value a sufficiently general class of *truths*. That is, our interest

⁵⁸ For more detail on the problem of trivial truths, see Chapter 2.

in truth for its own sake has a sufficiently *wide scope*. Hartry Field's (2001, 120) statement nicely captures this claim: 'What we desire is the infinite conjunction of all claims of the form ... 'I believe "p" only if p'. This is a *necessary* condition. PC should be rejected if it cannot be met. It can't be the case that we only care about some truths but don't care at all – or actively don't want – countless others.

A second kind of generality – that *most people* value truth – is also often claimed to hold. This claim would be vindicated if, for example, it turned out to be the case that this value was 'built in' to us, or a 'universal' fact about us. That this is the case is claimed by a number of philosophers. Recall Alston's (2005, 31) claim that 'Members of our species seem to have a built-in drive to get the truth about things', and Kitcher's (2004, 216) declaration that humans have an '(almost?) universal sense of curiosity'.

Once trivial truths are taken into account, it becomes clear that neither of these possibilities is plausible, or so I will argue.

Sufficiently General Class of Truths

As noted, considering certain examples makes it difficult to deny that some truths are valued for reasons other than practical utility. These examples establish, at most, the following: we *sometimes* want to know the truth about a particular area just to know the truth about that area. As argued above, this claim does little to establish PC. It is insufficiently general as it stands. With trivial truths in mind, a stronger claim than can be made: our interest in truth is *not* general enough for PC to be plausible. This is because the problem of trivial truths suggests that, more often than not, we *don't* care about truth. After all, we are constantly turning down opportunities to acquire true beliefs. And this psychological claim – that we

sometimes, but usually don't, care about believing the truth just to believe the truth – gives no support whatsoever to the idea that we regard true beliefs as worth having for their own sake – that is, *simply because* they are true.

The problem for defenders of PC, then, is this: Our interest in truth doesn't extend to a sufficiently general class of truths. Though it is difficult to be precise about such things, it is hugely implausible that a tiny proportion of all available truths is sufficient. We do not care in the slightest about the vast majority of available truths. But they are true, so we would value them, at least a little bit, if we valued true belief for its own sake. Since we do not value them, we should reject PC.

There is a possible response to this line of argument. This has some similarities to the *prima facie account* (PFA) discussed in Chapter 2. It also fails for similar reasons. The response is the following: We *do* value, or care about, every truth – including the most trivial – but we just care about other things more, at least much of the time. Kvanvig makes this claim in response to a case discussed by Ernest Sosa. Sosa (2003, 156) considers a person scooping up a handful of sand and carefully counting the grains '[a]t the beach on a lazy summer afternoon'. As he notes, this 'would give us an otherwise unremarked truth'. According to PC, we value this truth, we see it 'as at least a positive good, other things equal.' This view Sosa finds 'hard to take seriously. The number of grains would not interest most of us in the slightest.' In response to this, Kvanvig (2003, 41) writes:

We are finite beings, with limited time and resources for enhancing our well-being; and without some special situation in which counting the grains brings pleasure to a person, perhaps only by passing the time in a way not completely boring, our general interest in enhancing

our well-being comes into conflict with our general interest in the truth. Perhaps it is even true that most of the time when our interest in enhancing our well-being conflicts with our interest in the truth, the former overrides the latter. In any case, there is no obstacle to interpreting Sosa's example in this way.

This is incorrect. There are numerous obstacles to explaining Sosa's example, and related examples, in this way. As these are the similar to the problems faced by the PFA, let me note just one here. We can often get trivial truths for free. They do not always conflict with our general interest in our well-being, or other such interests. On Kvanvig's view, when this is the case, we would want these true beliefs. But we don't. Consider an example inspired by David (2005). Suppose that the television is on as you are brushing your teeth. Since you are brushing your teeth, there is not much that you can do to increase your well-being. On the television, there are advertisements running. Further, after finishing brushing, you intend to turn the television off. Would you care if somebody turned it off before you finished brushing? Would you feel that you had lost something of value? If you value trivial truths, then you would answer 'yes'. As David points out, 'even commercials will be good for truths like "They are *saying* this is the best toothpaste of all times"'. My strong intuition, however, is that most people would say 'no'. Truths such as this are not valued by most of us *at all* – it is not just that their value is outweighed in certain situations. As Sosa says of his own example, such truths 'would not interest most of us in the slightest.' And *this* fact about us – that we don't value countless easily available truths, even when they come without a cost – is incompatible with PC. According to PC, we value true beliefs *simply because* they are true. Trivial truths demonstrate that this is not the case.

Our interest in truth, then, is insufficiently general to establish PC. We do not value, in the slightest, most truths that are available to us most of the time. As such, it is entirely implausible that we value truth for its own sake. Thus, PC should be rejected.

Do Most People Value Truth Intrinsically?

Trivial truths do not only undercut the first kind of generality claimed by defenders of PC. They also undercut the second. That is, that desiring truth for truth's sake is a near universal part of human psychology.

This conclusion follows from my previous claim. If most people do not value the truth that 'they are *saying* this is the best toothpaste of all time' – even when acquiring this truth would have no cost – then most people do not value truth for its own sake.

We can also get this conclusion another way. To see this, first consider a different formulation of the same question: Is an intrinsic desire for truth a *standard* part of human psychology? If Alston, Kitcher, and others are right, the answer to this is 'yes'. Most people have this desire, and it is not unusual.

When the question is framed this way, we can see that PC is not a standard feature of our psychology. This can be illustrated with an example. Suppose that one of your family members suddenly began spending all his time meticulously counting the number hairs on his body. And imagine that, in explanation of this action, he sincerely claimed that he was doing this because he values truth for its own sake. Would his behaviour strike you as *normal*? My feeling is that at least most of us would consider this activity *strange*. We may even think that there was something *wrong* with him, and be tempted to send him to a psychiatrist.

This would not be our reaction if PC was a standard feature of our psychology. After all, this person is simply carrying out an activity PC claims that most of us value. If PC is correct, we would think both that this person was doing something *valuable*, and we would *value* what he is doing. *I don't think these things, and I doubt most people do.* This suggests that valuing truth for its own sake is far from standard.

There is a possible response to this. It might be claimed that this argument is indecisive because, even when uncontroversial values are considered, we always think that it is strange if a person does *nothing but* pursue one thing. Consider pleasure, for example. Though most of us value pleasure, it is not uncommon for people to claim that a completely hedonistic lifestyle is unhealthy or unworthy in some way. By the line of reasoning above, we should conclude that we *don't* value pleasure. But that seems absurd.

There is something to this objection. It may well be correct that *any* kind of single-minded excess strikes us as odd. This is not, however, a convincing response to my claim. There is a key difference between the two cases. It would be strange if somebody spent *any* of their time counting the strands of hair on their body for no practical reason, whereas somebody occasionally seeking pleasure for its own sake is not *at all* unusual. Even adjusting for the implicit strangeness of single minded obsession, then, a desire for truth for its own sake remains distinctly odd.

The following may seem to be a concern: As a number of quotes at the beginning of this chapter demonstrate, people *do* respect, and value, people who spend their life searching for truth in certain areas. There is something almost heroic about a

scientist on a pure quest for truth, for instance. We do not think that they are mentally ill. How can we explain this? I think the most plausible explanation is that it is not truth *for its own sake* that we value in these cases. We value something else that truth is a means to, and then mistake this for valuing truth intrinsically. That this is going on is suggested by the hair counting example. In both cases, the scientist and the hair counter are after truth. Both end up with true beliefs. As the cases are symmetrical in this sense, it is hard to see how truth *itself* could account for our different attitudes. The difference, it seems to me, is simply that we *care* about our fundamental nature, for instance, but don't care *at all* about the number of hairs on our body. Truth, considered by itself, has nothing to do with it.

The claim that most of us value truth for its own sake is implausible. If we did, then most of us would think it both normal and valuable to count the hair on our bodies. Now, some people *may* think this, but it does not seem to be standard. On the contrary, this value would strike many of us as decidedly *abnormal*.

Do We Have a Basic Preference for True Belief?

Even granting the previous arguments, it might be thought that certain versions of PC can still succeed despite the problem of trivial truths. Return to Lynch's claim. Recall that, according to Lynch, we have a *basic preference* for true belief. A basic preference is 'a preference for something that can't be explained by our preference for other things' (2004, 15). Applied to true belief, this amounts to agreeing with the following claim: 'where the belief that *p* and the belief that not-*p* have identical instrumental value, it is better, just on grounds of truth alone, to

have the true belief rather than the false one' (2004, 21). Lynch believes that most people would agree to this claim (2004, 15).

I have argued that the Experience Machine and Russell World tell us nothing about the value of true belief. Though they seem like they would, this appearance is deceptive. I did not argue that Lynch's claim that we have a basic preference for truth was *incorrect*. I will now argue that there is good reason to doubt that we have any such preference.

At first sight, trivial truths may not seem to undercut Lynch's claim. The problem of trivial truths tells us that we do not care about having true beliefs about countless matters, but we may still prefer to have true beliefs *all else equal* in such cases. If this is correct, then it might be thought that the objection from trivial truths does not undermine PC. But this is not correct. Trivial truths also give us reason to be suspicious of this psychological claim.

To motivate this, I will present an example that meets Lynch's description of a basic preference for true belief. That is, there are two beliefs – one true and one false – with equal instrumental value. If Lynch's claim is correct, we would prefer the true belief to the false belief on the grounds of truth alone. I suggest that, when trivial propositions with equal instrumental value are considered, it is not at all obvious that we care whether we believe truly.⁵⁹

Consider the following example:

⁵⁹ Note also that this example fulfils Marion David's (2005, 299) requirements, discussed above, for a legitimate test of whether we value truth intrinsically.

Suppose there is a contest. The winner receives ten million dollars.

These are the terms: whoever reports the number of grains of sand on a particular beach wins. But there is a catch: the number of grains the contestant reports does not have to be true for her to win. The only requirement is that she *believes* it to be true when she reports it. To test for this, the contestant is given an infallible lie detector test.

Those running the contest – extremely rich trust fund kids who stand to inherit billions when their parents die – are nihilistic sadists who gain pleasure from the humiliation of others. They figure it will be funny to watch otherwise respectable people – academics, lawyers, teachers, etc. – degrade themselves by desperately counting grains of sand. They see this as an apt analogy for life: people act like it matters – putting on suits, fighting for ‘justice’, working hard, trying to understand the world, helping kids get a good start in life, etc. – when really they are just desperate self-deceived animals spending their time doing nothing more valuable than counting sand. The sadists do not themselves know how many grains of sand are on the beach, nor do they desire to find out. They figure testing for sincere belief will be sufficient to ensure that the contestants spend large amounts of time counting sand. To avoid people thinking the contest is a hoax the sadists and the contestants all sign a legally binding contract that ensures that a contestant who has an honest belief about the number of grains will receive the money.

The sadists overlook a possibility. Imagine there is a hypnotist who can induce sincere beliefs in people at will. Jordan goes to this

hypnotist. Like the sadists, the hypnotist has no idea how many grains of sand are on the beach. Jordan gets the hypnotist to induce a belief that there is some random number of grains. It doesn't matter what the number is – either way she wins the money. She then returns to the beach and submits the number to the sadists. The lie detector test is administered and she passes. Begrudgingly, the sadists transfer ten million dollars to her account.

Now consider two alternatives. One: the hypnotist's random guess about the number of grains on the beach is false. Two: the hypnotist's random guess is true. Nobody – including Jordan – will ever confirm which of these alternatives has taken place. Both alternatives clearly have equal instrumental value – she gets the money either way. Would Jordan care whether her belief is true?

My intuition is that Jordan would not care one bit about whether her belief is true. *I* simply wouldn't care either way. Supposing that my reaction is shared, then trivial truths give us reason to doubt not only that we often care about true belief, but that we have even a basic preference for true belief. And if we do not value believing truth even when nothing is on the line, it is hard to see what hope there is for PC.

Final Thoughts on PC and Its Relationship to CVT

I have argued that PC should be rejected. The arguments and examples commonly used to support PC fail. Further, there are a number of direct arguments against PC.

To conclude, I will consider some consequences for CVT if PC is rejected. The first point to note is that, once PC is rejected, arguments that make use of PC to support CVT don't even get off the ground. To get a sense of this, consider some of the arguments noted at the start of this chapter.

Begin with the *best explanation argument*. According to this argument, CVT is the best explanation for the fact that we regard true belief as intrinsically valuable – for the fact that we *value* true belief for its own sake. This argument obviously fails without PC. If people do not value truth for its own sake, then there is no fact that needs to be explained by CVT.

Further, CVT does *not* seem to be the best explanation for the more plausible psychological claims we have considered. It is strange, for example, to claim that CVT is the best explanation for the fact that we sometimes, but usually don't, care about believing the truth just to believe the truth. This explanation would amount to little more than the claim that we are mistaken most of the time. After all, we are constantly turning down opportunities to acquire true beliefs.

Next consider the *evidence* argument. On this view, PC gives us direct evidence of the truth of CVT. What kind of evidence this is claimed to be depends on the theory of value that is endorsed. But rejecting PC completely undermines this claim on any theory of value. If PC is not accurate, it is not evidence for anything.

Rejecting PC may do more than merely remove positive reasons to endorse CVT. It may itself give us reason to reject CVT. As we have seen, it is commonly claimed that what we value – including what we desire and care about – is good evidence for what is *valuable*. Though this claim is strongest on subjectivist

accounts of value, it is not exclusive to such accounts. Objectivists also claim that what we value is good evidence for what is valuable.

If this is the case, then regarding something as *lacking* value may well be good evidence for it *not* being valuable. The plausibility of this claim may be easier to see if we first consider an uncontroversial example. I do not believe that plastic is intrinsically valuable – I do not value plastic for its own sake. If I try to think of the possible reasons to regard plastic as intrinsically valuable, none of them strike me as compelling. My reaction to such thoughts is to claim that plastic is not *valuable*. It is not to remain neutral on the topic. I assume that it is the same for you. Believing that something has no value – not *valuing* it – generally leads us to conclude that it is *not valuable*.

Now return to truth. I have argued that we do not value truth intrinsically.

Assuming this is correct, this fact seems to give us some evidence that truth is not intrinsically *valuable*. If it is correct that *valuing* something is good evidence for it being valuable, then it is hard to see why the reverse would not hold. I cannot see any basis for such an asymmetry. It strikes me as plausible that our values could be either evidence for *and against* what is valuable, or evidence for neither. But it is hard to see why our values would *just* be evidence for what is valuable.

I have argued that we do not value truth for its own sake. If this is right, and if it is right that psychological claims can support or even establish value claims, then – as well as rejecting PC – we have found further reason to reject CVT.

Conclusion

In his paper *Why Do We Value Knowledge*, Ward Jones (1997, 423) claims that ‘It should be uncontroversial that we do and should value true believing’. Many writers have made similar claims. This essay has argued that both parts of this statement are open to serious doubt.

In Chapter 1, I argued that there are reasons to be sceptical of the idea that truth is *instrumentally* valuable. As well as there being powerful counter-examples to this idea, true beliefs do not seem to be practically superior to false beliefs. As such, instrumental considerations seem to give us no reason to prefer true to false beliefs. Chapter 2 argued that there is a strong reason to reject the *intrinsic* value of truth (CVT). This is the problem of trivial truths. Though this problem is widely discussed, its power is frequently missed. I looked at two prominent responses to this problem, and argued that both fail. For this reason, CVT still has a significant strike against it. Finally, Chapter 3 argued that, when trivial truths are kept in mind, it becomes highly implausible that we in fact *value* truth for its own sake. Nor is this claim bolstered by compelling arguments in its favour. Rejecting PC also has implications for CVT. This is because a number of arguments for CVT assume this psychological claim. As a result, the failure of this claim has the consequence of undermining these arguments.

The arguments in this essay are not decisive. More work needs to be done before confident conclusions can be drawn. It may also be the case that various theses in the vicinity of VT are consistent with my claims. Nonetheless, my arguments

suggest that the most plausible conclusion is this: We do not value true believing for its own sake, and there is little reason to think that we should value it at all.

References

- Allen, B. 1992. 'Nietzsche's Question, "What Good Is Truth?"' *History of Philosophy Quarterly* 9: 225-240.
- Alston, W. 2005. *Beyond Justification: Dimensions of Epistemic Evaluation*. Ithaca, NY: Cornell University Press.
- Aristotle. 1958. *Metaphysics*. Oxford: Clarendon Press.
- Augustine. 2009. *Confessions*. H. Chadwick (trans.). New York: Oxford University Press.
- 1887. *De Mendacio*. H. Browne (trans.) in P. Schaff (ed.) *Nicene and Post-Nicene Fathers* (Vol. 3). Buffalo, NY: Christian Literature Publishing Co.
- Bacon, F. 1986. *The Essays*. New York: Penguin Classics.
- BonJour, L. 1985. *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.
- Brady, M.S. 2009. 'Curiosity and the Value of Truth', in A. Haddock, A. Miller and D. Pritchard (eds.), *Epistemic Value*. 265-283. Oxford: Oxford University Press.
- Coates, A. 2009. 'Explaining the Value of Truth'. *American Philosophical Quarterly*. 46: 105-115.
- David, M. 2005. "On 'Truth is Good.'" *Philosophical Books* 46: 292-301
- 2001. 'Truth as the Epistemic Goal', in M. Steup (ed.), *Knowledge, Truth, and Duty*. 151-169. New York: Oxford University Press.
- De Brigard, F. 2010. 'If You Like It, Does It Matter If It's Real?' *Philosophical Psychology* 23: 43-57.
- DePaul, M.R. 2001. 'Value Monism in Epistemology', in M. Steup (ed.), *Knowledge, Truth, and Duty*. 170-185. New York: Oxford University Press.

- Descartes. 1996. *Meditations on First Philosophy*. Cambridge: Cambridge University Press.
- Dobson, K. and Franche, R.L. 1989. 'A Conceptual and Empirical Review of the Depressive Realism Hypothesis.' *Canadian Journal of Behavioral Science* 21: 419-433.
- Engel, P. 2009. 'Pragmatic Encroachment and Epistemic Value', in A. Haddock, A. Miller and D. Pritchard (eds.), *Epistemic Value*. Oxford: Oxford University Press.
- Feldman, F. 2011. 'What Do We Learn from the Experience Machine?', in J. Meadowcroft and R. Bader (eds.), *The Cambridge Companion to Robert Nozick's "Anarchy, State, and Utopia."* Cambridge: Cambridge University Press.
- Field, H. 2001. *Truth and the Absence of Facts*. Oxford: Oxford University Press.
- Foley, R. 1987. *The Theory of Epistemic Rationality*. Cambridge, MA: Harvard University Press.
- Frankfurt, H. 2004. *The Reasons of Love*. Princeton: Princeton University Press.
- Frege. 1967. 'The Thought', in P.F. Strawson (ed.), *Philosophical Logic*. 17-38. Oxford: Oxford University Press.
- Gemes, K. 1992. 'Nietzsche's Critique of Truth'. *Philosophy and Phenomenological Research*. 52: 47-65.
- Glanzberg, M. 2014. 'Truth'. *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2014/entries/truth/>>.
- Goldman, A. 2002. 'The Unity of Epistemic Virtues', in *Pathways to Knowledge*. New York: Oxford University Press.
- 1999. *Knowledge in a Social World*. New York: Oxford University Press.

- 1986. *Epistemology and Cognition*. Cambridge, MA: Harvard University Press
- Grimm, S. 2011. ‘On Intellectualism in Epistemology.’ *Mind*, 120: 705-733.
- 2009. ‘Epistemic Normativity’, in A. Haddock, A. Miller and D. Pritchard (eds.), *Epistemic Value*. 243-264. Oxford: Oxford University Press.
- 2008. ‘Epistemic Goals and Epistemic Values.’ *Philosophy and Phenomenological Research*, 77: 725-44.
- Haack, S. 1993. *Evidence and Inquiry: Towards Reconstruction in Epistemology*. Cambridge, USA: Blackwell.
- Heal, J. 1987/88. ‘The Disinterested Search for Truth.’ *Proceedings of the Aristotelian Society* 88: 97-108.
- Hempel, C. 1965. ‘Aspects of Scientific explanation’, in *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: Free Press.
- Heraclitus. 1979. *The Art and Thought of Heraclitus*. C.H. Kahn (trans.). Cambridge: Cambridge University Press.
- Horwich, P. 2006. “The Value of Truth.” *Nous* 40 (2): 347–360.
- Houellebecq, M. 2008. *H.P. Lovecraft: Against the World, Against Life*. London: Gollancz.
- Jackson, F. 1998. *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. New York: Oxford University Press.
- James, W. 1956. ‘The Will to Believe’ in *The Will to Believe and Other Essays on Popular Philosophy*. New York: Dover Publications.
- Jones, W. 1997. ‘Why Do We Value Knowledge?’ *American Philosophical Quarterly* 34: 423-439.
- Kitcher, P. 2004. ‘The Ends of the Sciences’, in B. Leiter (ed.), *The Future for Philosophy*. Oxford: Oxford University Press.

- Kornblith, H. 2002. *Knowledge and its Place in Nature*. New York: Oxford University Press.
- Kvanvig, J. 2003. *The Value of Knowledge and the Pursuit of Understanding*. Cambridge: Cambridge University Press.
- Lehrer, K. 1990. *Theory of Knowledge*. Boulder, CO: Westview Press.
- Leiter, B. 2014. 'The Truth is Terrible', forthcoming in D. Came (ed.), *Nietzsche on Morality and the Affirmation of Life*. Oxford: Oxford University Press.
- 2002. *The Routledge Philosophy Guidebook to Nietzsche on Morality*. London: Routledge.
- Lewis, D. 2000. 'Dispositional Theories of Value', in *Papers in Ethics and Social Philosophy*. Cambridge: Cambridge University Press.
- Lovecraft, H.P. 1999. *The Call of Cthulhu and Other Weird Stories*. New York: Penguin Classics.
- Lynch, M.P. 2009. 'Values of Truth and Truth of Values', in A. Haddock, A. Miller and D. Pritchard (eds.), *Epistemic Value*. 225-242. Oxford: Oxford University Press.
- 2004. *True to Life: Why Truth Matters*. Cambridge, Mass: MIT.
- Mahon, J.E. 2008. 'Two Definitions of Lying'. *International Journal of Applied Philosophy*. 22:2 211-230.
- Mill, J.S. 1982. *On Liberty*. Harmondsworth: Penguin.
- Nietzsche, F. 1992. *Ecce Homo: How One Becomes What One Is*. R.J. Hollingdale (trans.). New York: Penguin Books.
- 1974. *The Gay Science*. W. Kaufman (trans.). New York: Vintage.
- 1966. *Beyond Good and Evil*. W. Kaufman (trans.). New York: Vintage.
- Pillar, C. 2009. 'Desiring the Truth and Nothing But the Truth'. *Nous*. 43: 193-213.

- Plato. 1973. *Plato: The Collected Dialogues, Including the Letters*. E. Hamilton and H. Cairns (eds.). Princeton: Princeton University Press.
- Pritchard, D. 2007. "Recent Work on Epistemic Value." *American Philosophical Quarterly* 44 (2): 85-110.
- Ridley, A. 2010. 'Perishing of the Truth: Nietzsche's Aesthetic Prophylactics.' *British Journal of Aesthetics* 50: 427-437.
- Rorty, R. 1998. *Truth and Progress: Philosophical Papers Volume 3*. Cambridge: Cambridge University Press.
- Russell, B. 1921. *The Analysis of Mind*. London: Allen & Unwin.
- Schwitzgebel, E. 2014. 'Belief'. *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2014/entries/belief/>>.
- Sidgwick, H. 1907. *The Methods of Ethics*. 7th ed. Indianapolis: Hackett.
- Smith, M. 2005. 'Meta-Ethics', in F. Jackson and M. Smith (eds.), *The Oxford Handbook of Contemporary Philosophy*. 3-30. Oxford: Oxford University Press.
- Sosa, E. 2003. 'The Place of Truth in Epistemology', in M. DePaul and L. Zagzebski (eds.), *Intellectual Virtue: Perspectives from Ethics and Epistemology*. 155-180. Oxford: Oxford University Press.
- 2001. 'For the Love of Truth?', in A. Fairweather and L. Zagzebski (eds.), *Virtue Epistemology: Essays on Epistemic Virtue and Responsibility*. Oxford: Oxford University Press.
- Stich, S. 1990. *The Fragmentation of Reason*. Cambridge, MA: The MIT Press.
- Strickland, L. 2013. 'Philosophy and the Search for Truth.' *Philosophia* 41: 1079-1094.

- Taylor, S.E. and Brown, J.D. 1988. 'Illusion and Well-Being: A Social Psychological Perspective on Mental Health. *Psychological Bulletin*, 103: 193-210.
- Weijers, D. 2014. 'Nozick's Experience Machine is Dead, Long Live the Experience Machine!' *Philosophical Psychology* 27: 513-535.
- Whiting, D. 2013. 'The Good and the True (or the Bad and the False).' *Philosophy*. 88: 219-242.
- Wrenn, C.B. 2010. "True Belief is Not Instrumentally Valuable", in C. Wright and N. Pedersen (eds.), *New Waves in Truth*. New York: Palgrave Macmillan.
- Zagzebski, L. 2003. "The Search for the Source of Epistemic Good." *Metaphilosophy* 34 (1/2): 12-28.
- Zapffe, P. 2004. 'The Last Messiah'. G. Tangenes (trans.). *Philosophy Now*, 45.
- Zimmerman, M.J. 2015. 'Intrinsic vs. Extrinsic Value'. *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.), forthcoming URL = <<http://plato.stanford.edu/archives/spr2015/entries/value-intrinsic-extrinsic/>>.