

# SPIRITUAL BLINDNESS, SELF-DECEPTION AND MORALLY CULPABLE NONBELIEF

KEVIN KINGHORN

*Wycliffe College, University of Oxford, UK*

It is not an uncommon practice to blame others for their beliefs. We say that the Nazi officers should not have believed Hitler's rhetoric as they did. We may argue that those who believe that a government need not provide free healthcare to all its citizens are morally callous. Sometimes we criticize others for whether or not they form *any* belief, claiming that they are irresponsibly naïve in believing too easily or that they are unduly cynical in not being willing to believe.

As to why it is appropriate to blame others for their beliefs, our language often suggests that belief is a voluntary matter within people's control. Consider the phrases, 'I refuse to believe that'; 'Why won't you believe me?'; 'I've decided that such a course of action would be imprudent'; and so forth. Yet despite such language, there seem to be compelling reasons to think that beliefs are *not* the kinds of mental states we can hold simply by choosing to hold them.

First, it is a common psychological experience that we are unable by an act of the will simply to acquire beliefs we may desire to hold. A woman who must testify in court and dreads choosing between honesty and loyalty may wish that she could believe that her friend is innocent. A religious person may be convinced that the way in which God judges the lives of people depends at least in part on whether they believe certain propositions contained in a creed. However, even though people may sometime have very strong reasons for wanting to believe certain propositions, is it often people's testimony that they are simply unable to acquire the beliefs they wish they could acquire.

Second, some philosophers have pointed out that, aside from any psychological impediments to voluntary belief acquisition, there seem to be *logical* problems with the idea of choosing to believe things. Beliefs are representational in nature; they are our 'maps' of the world in that they represent what we think is already true of the world. Thus, as Bernard Williams remarks, even if I could somehow find the psychological means to acquire a belief at will, 'it is unclear that . . . I could seriously think of it as a belief, i.e. as something purporting to represent reality'.<sup>1</sup> After all, if

I could somehow *choose* to believe something, I would realize that my belief stems from my own sovereign *choice* – and that the belief doesn't necessarily have any connection with facts about the actual world. Yet, if I realize that my belief doesn't necessarily have any connection with what's true about the actual world, then it's not a *belief* in the first place. Because our beliefs are our 'maps' of what the world is like, the act of 'choosing to believe something' would be analogous to a cartographer deciding where the lines of England, Wales, and Scotland should be on a map. If the cartographer is aware that his newly-drawn map reflects his own decisions and not necessarily the borders of these countries as they exist in the world, then he will be aware that his new map is not actually a map at all.

So let us take as a working assumption that people, rather than *choosing* which beliefs they will hold, instead *find themselves* believing things as the evidence before them seems to dictate. Does it follow from this that people cannot rightly be blamed for the beliefs they hold (or fail to hold)? Such a conclusion may be far too quick. In explaining how a nonbeliever can be culpable for a failure to see her evidence as pointing to the truth of certain Christian affirmations, many Christian writers have proposed that nonbelievers may be in a culpable state of 'spiritual blindness'. Indeed, we often speak of a person 'blinding' herself to some truth or being 'self-deceived' about some matter. I shall argue that these descriptions can indeed be appropriate ways of describing how a person can culpably come to hold (or fail to hold) some belief. In what follows I shall outline the conditions under which a person might rightly be described as having wilfully (and culpably) blinded herself to some spiritual truth.

There may be a variety of aspects of spirituality to which a person can be blind. To narrow our discussion, I shall focus on the question why someone might not believe that certain statements found in the Christian scriptures – such as statements regarding Jesus's divinity and recorded statements of Jesus as to how one should live one's life – are in fact instances of God's communication to all people. The recognition of such purportedly divine communication as in fact authentic is perhaps the most ready explanation as to what separates those who are followers of the Christian religion from those who are not followers of the Christian religion. And so it seems appropriate to focus our discussion of spiritual blindness on the question of what it would mean for a person culpably to fail to see that some statement does in fact come from God.

Also, it is important to emphasize the idea of a person *culpably* being blind to the fact that a particular statement comes from God. The general term 'spiritual blindness' might raise the question whether a person can be blind to a spiritual truth in much the same way that a person can be blind to colour distinctions. Our interest, however, is on a culpable failure to believe that some statement comes from God. And so we will need to

make some assumptions as to the general criteria for culpability. I take as a working assumption that people have moral beliefs and that some of our beliefs about what is morally good or bad involve questions of where our moral obligations lie. In cases where one believes that one is obligated to act a certain way, but also possesses a desire the realisation of which (one believes) is at odds with one's perceived obligation, one is faced with a certain kind of moral decision. When one then acts (or attempts to act) contrary to one's perceived obligation, one does something, so I shall take as a working assumption, for which one is morally culpable. Accordingly, if the Christian theist is rightly to describe spiritual blindness as a state for which one is morally culpable, she will need to show how spiritual blindness is the result of bad moral decisions of this kind.

In what follows I shall offer an account of how this scenario is possible. My purpose is not to provide an argument that the Christian God *does* exist or to examine the empirical question of how many instances of Christian nonbelief in our world really *do* stem from morally culpable decisions people make. Rather, my purpose is merely to outline the conditions that would need to be met if one is rightly to be described as morally culpable for a failure to see that some Christian affirmation is indeed true.

## I. HISTORIC RECOGNITION OF THE LINK BETWEEN ONE'S CHARACTER AND ONE'S BELIEFS

The idea that character flaws can prevent one from acquiring certain kinds of beliefs or knowledge is hardly novel. Aristotle linked vice with the destruction of one's practical wisdom.

Pericles and men like him have practical wisdom . . . because they can see what is good for themselves and what is good for men in general; . . . but the man who has been ruined by pleasure or pain forth with fails to see . . . that for the sake of this or because of this he ought to choose to do whatever he chooses and does.<sup>2</sup>

While Aristotle appealed in more general terms to the connection between epistemic shortcomings and corrupted character, writers within the Christian tradition have often used this purported connection in explaining why not all people form various Christian beliefs. St. Bonaventure maintained that the failure to hold an unwavering belief that God exists stems 'from a defect in the knower rather than from a deficiency in the object known'.<sup>3</sup> John Henry Newman remarked that there exist good arguments for the Christian faith, which are capable of convincing anyone who 'fairly studies' their premises.<sup>4</sup> Newman also noted that 'It is almost a proverb, that persons believe what they wish to be true, . . . [People] readily believe reports unfavourable to persons they dislike, or confirmations of theories of their own'.<sup>5</sup> William James went so

far as to state that ‘As a rule we disbelieve all facts and theories for which we have no use. Clifford’s cosmic emotions find no use for Christian feelings. Huxley belabors the bishops because there is no use for sacerdotalism in his scheme of life’.<sup>6</sup>

The comments of Newman and James go beyond merely pointing to some general character flaw that keeps one from forming Christian beliefs. Their point is roughly this: A person who fails to see his evidence as indicating that a certain proposition is true may fail to do so because of some sort of negative demeanour toward the possibility that that proposition is true.

Perhaps the two best-known articulators within the Christian tradition of this line of thought are John Calvin and Jonathan Edwards. Both authors take their cues from Romans 1:18–20, where St. Paul writes that

The wrath of God is being revealed from heaven against all the godlessness and wickedness of men who suppress the truth by their wickedness, since what may be known about God is plain to them, because God has made it plain to them. For since the creation of the world God’s invisible qualities – his eternal power and divine nature – have been clearly seen, being understood from what has been made, so that men are without excuse.

Both Calvin and Edwards insisted that the world provides ample evidence to support Christian beliefs. Edwards maintained that God gave humans an understanding of the divine, eternal things that concern them. As to central, Christian doctrines, Edwards concluded that, ‘if men have not respect to ‘em as real and certain things, it cannot be for want of sufficient evidence of their truth’.<sup>7</sup> Calvin likewise asserted that God

has been pleased . . . to manifest his perfections in the whole structure of the universe, and daily place himself in our view, that we cannot open our eyes without being compelled to behold him.<sup>8</sup>

At the same time, Calvin maintained that human corruption prevents many of us from seeing certain Christian truths that we ‘are desirous not to know’.<sup>9</sup> He asked,

how can the idea of God enter your mind without instantly giving rise to the thought, that since you are his workmanship, you are bound, by the very law of creation, to submit to his authority? – that your life is due to him? – that whatever you do ought to have reference to him? If so, it undoubtedly follows that your life is sadly corrupted.<sup>10</sup>

Following Calvin, Edwards asserted that

The mind of man is naturally full of enmity against the doctrines of the gospel; which is a disadvantage to those arguments that prove their truth, and causes them to lose their force upon the mind.<sup>11</sup>

As Nicholas Wolterstorff summarises, ‘characteristic of the Reformed tradition’ in which Calvin and Edwards feature prominently is the view that ‘cases of unbelief’ stem not so much from insufficient available evidence, but rather from ‘*resistance* to the available evidence’.<sup>12</sup>

It seems obvious enough from everyday observations that people’s desires can impact what they come to believe. As we observe, for instance, how sports fanatics tend to evaluate referees’ performances or how parents tend to evaluate their children’s athletic potential, it is not difficult to reach Francis Bacon’s general conclusion that ‘what a man would rather was true, that he the more readily believes’.<sup>13</sup> But how exactly can one’s desires affect one’s beliefs? And, if we are to think of people as morally culpable for being in a spiritually blind state where their desires affect their beliefs, what morally significant choices of theirs might we point to as culpably leading to their being in this state? To answer these questions we will need to examine the psychological literature in some detail. Our next three sections will be devoted to doing just that.

## II. HOW ONE’S DESIRES CAN AFFECT ONE’S BELIEFS

Francis Bacon anticipated the research findings of modern-day psychologists in commenting that ‘in innumerable ways, and those sometimes imperceptible, the affection tinges and infects the Intellect’.<sup>14</sup> Often, the uncomfortable nature of a belief provides the key to understanding how one’s desires affect the formation of that belief. A mother who very much wants it to be the case that her son is innocent of the crime for which he has been arrested might well come to believe in her son’s innocence despite contrary evidence that others consider fairly conclusive. Similarly, a husband might believe his wife’s attestations of faithfulness despite contrary evidence that more neutral observers find overwhelming. In each case, the uncomfortable nature of a particular belief ‘that *p*’ leads the person very much to desire that *p* not be true; and this desire in turn seems somehow to affect the formation of the person’s belief ‘that *p*’.

A belief may be uncomfortable or undesirable for any number of reasons. Sometimes a person may wish for something to be false simply because someone she dislikes is committed to it being true. For example, a person who has an unhealthy competition with a work colleague may wish that her colleague’s hypotheses be proved wrong; and a person who dislikes a politician may want the politician’s proposals to prove embarrassingly ineffective. In each case the person wants things to be a certain way; and certain beliefs (e.g., that her colleague is right and that the politician has good ideas) may therefore be uncomfortable for her. David Pears has pointed out that some beliefs may be uncomfortable because they stand in the way of a person performing certain actions – including akratic ones – that he wants to perform. Suppose, for example,

that a man at a party desires to have more drinks – but judges it best, because he must drive home, to stop at the two drinks he has already had. Pears states that the man ‘might tell himself, against the weight of his evidence, that it is not dangerous to drive home after six measures of whiskey, or he might forget, under the influence of his wish, how many drinks he had already taken’.<sup>15</sup> Here the man is resistant to holding a belief because the belief makes a desired action more difficult. In the end, we can imagine any number of scenarios in which a person might be motivated to hold (not hold) a particular belief.

In instances where certain desires do affect the formation or continued holding of one’s beliefs, it will sometimes be the case that one is unaware that one has these desires. The man from a previous example who believes his wife’s attestations of faithfulness may see his own belief as characteristic of the charitable way he judges other people – when in truth his belief stems from a desire to avoid the public embarrassment of a divorce. As other examples we might cite a wife who does not see that her affair is really an attempt to punish her husband; a son who does not see that his rebellion is really an attempt to get his parents to listen; and a daughter who does not see that her eating disorder is really an attempt to gain some control over her life. Psychotherapists commonly help people make these types of discoveries, which often are discoveries about why one performs the actions one performs. Put another way, they are often discoveries about the purposes one seeks to achieve. And so, among those beliefs affected by desires we should include beliefs as to what one’s purposes in acting truly are.

Even when one *is* aware that one has a certain desire, one may still not be aware that that desire plays a causal role in one’s own actions and in what one believes. A university student may be very aware of her own desire that her parents love her and may believe that her parents, who are both physicians, would undoubtedly love her if she were to follow in their steps as a physician. However, as Alfred Mele has pointed out, given that a person can be mistaken about her actual purposes in acting, we can suppose that this desire leads her to enter medical school without also supposing that she *believes* that this desire has anything at all to do with her decision.<sup>16</sup> It seems clear, then, that desiring *p* to be the case can affect whether one comes and/or continues to believe that *p* is the case – regardless of whether one is aware that one’s desire is having this affect or that one even has this desire.

Instances where one’s desires affect the formation of one’s beliefs sometimes fall under the description of irrationality in that one’s desires may keep one from meeting correct and/or one’s own standards of evidence assessment and proper reasoning. Because such irrationality involves one’s desires, we might refer to it generally as ‘motivated’ irrationality, or irrationality generated through ‘hot’ mechanisms. At the same time, it is important to note how one’s beliefs can be affected by *unmotivated* irrationality, or irrationality stemming from ‘cold’

mechanisms. So-called perversions of reason are cognitive failures of some kind, though they do not stem from wanting things to be a certain way.

Among the common kinds of unmotivated irrationality is what Daniel Kahneman and Amos Tversky term the 'availability heuristic'. According to the availability heuristic, 'subjective probability is evaluated by the difficulty of retrieval and construction of instances'.<sup>17</sup> Because it is easier for most people to think of words starting with 'k' than to think of words that have 'k' as the third letter, people may sometimes tend toward believing that there exist more words of the former sort than words of the latter sort.<sup>18</sup> Another source of unmotivated irrationality identified by Kahneman and Tversky is the 'representativeness heuristic'. A person may be asked whether the births of six children in a family are less likely to have the gender sequence of BBBGGG than the sequence of GBBGBG. According to Kahneman and Tversky's experiments, persons asked show on the whole a tendency to judge the first gender sequence as a less likely scenario than the second.<sup>19</sup> Because the first sequence can seem less 'representative' of the random process by which gender is determined, some people will accordingly see this first sequence as comparatively less likely to occur.

Yet another source of unmotivated irrationality involves the undue weight we can give to certain evidence if that evidence is in some way especially *vivid* to us. Richard Nisbett and Lee Ross identify three main ways in which information (which is potential evidence for a given conclusion) can strike us as especially vivid. First, some information can carry more emotional interest with us than other information. If one is told that 'Jack skidded off an icy road and demolished a parked car', and if the parked car happens to belong to a stranger, then we may be to a certain extent inclined to think that Jack was an unlucky victim of the icy road. But if the car is one's own or belongs to a close friend or relative, then we may be more inclined to conclude that Jack was careless or that he was driving too fast.<sup>20</sup> Second, information may be especially vivid to us if it is particularly concrete. Nisbett and Ross point out that the information that 'Jack was killed by a semitrailer that rolled over his car and crushed his skull' is more vivid and imagery-provoking than the information that 'Jack sustained fatal injuries in an auto accident'.<sup>21</sup> As such, this information is more likely to be recognised and recalled as evidence for conclusions such as that speed limits should be lowered on motorways. The third source of vividness that Nisbett and Ross identify involves the temporal, spatial and sensory proximity of the information. Nisbett and Ross explain,

The news that a bank in one's neighborhood has been robbed just an hour ago is more vivid than the news that a bank on the other side of town was robbed last week. The former bank robbery, accordingly, is likely to have a greater impact on one's views of the seriousness of the crime problem in one's city or the need for stiffer prison sentences for bank robbers.<sup>22</sup>

This list of ‘cold’ mechanisms is by no means exhaustive. Yet it does allow us to imagine any number of examples where one’s beliefs are affected by perversions of reason.

Sometimes the line between ‘cold’ explanations and ‘hot’ explanations will not be altogether clear. In testing a favoured hypothesis people will often tend to recognise confirming evidence more readily than disconfirming evidence. If this ‘confirmation bias’ stems simply from the fact that people tend to see more easily that which they expect to see, then it is cold mechanisms that will be affecting the formation of a person’s beliefs. If, however, this ‘confirmation bias’ is the result of a person’s desire that her initial hypothesis be proved right, then a hot explanation is in order. Where both hot and cold mechanisms are at work, it may be impossible to discern precisely what influence on belief formation any one mechanism is having.

Within hot explanations of the link between desires and belief formation, an understanding of cold mechanisms can help us explain how it is that desiring something to be the case can lead to one’s believing that something is the case. Taking as an example the mother who desires that her son be innocent of the crime of which he is accused, we might suppose that her desire leads to feelings of great emotional comfort when she considers evidence in favour of her son’s innocence. Though this evidence she considers might not be particularly good evidence – even by her own standards – the evidence might be especially vivid to her in virtue of her emotional reaction to it. Being vivid evidence, it might through cold mechanisms be given undue weight. Also, the emotional comfort the mother finds in contemplating this evidence might lead her frequently to do so, which would naturally tend to make this evidence more available to her as she considers various pieces of evidence for and against her son’s guilt. Further, if her desire leads her to form the belief that her son is in fact innocent, then her belief will constitute a sort of initial hypothesis for her. As such, she may tend to recognise evidence in favour of her son’s innocence more readily than contrary evidence – and this on the cold explanation that one sees more easily that which one expects to see.

Given that spiritual blindness – as outlined by Christian theists such as Calvin and Edwards – involves a desire for things to be a certain way, a person cannot rightly be described as blinding himself to some spiritual truth if his failure to believe that truth stems purely from cold mechanisms. At the same time, the presence of a hot motivation that something be a certain way does not by itself establish culpability – even when this motivation keeps one from holding some Christian belief that one otherwise would hold. As stipulated earlier, our working assumption is that culpability only arises when one decides to act in violation of what one believes to be a moral obligation. Thus, if the Christian theist is rightly to assign moral culpability for being in a state of spiritual blindness, she will need to point to some decision the spiritually blind person makes that contributes to his being in this state.



It is certainly in keeping with many Christian writers that people can and sometimes do play an active role in ‘blinding’ themselves to the truth of certain Christian beliefs. Aquinas, for example, maintained that sometimes ‘ignorance is directly and essentially voluntary, as when a man is purposely ignorant that he may sin more freely’.<sup>23</sup> And Joseph Butler, in his aptly titled sermon, *Upon Self-Deceit*, asserted that ‘It is as easy to close the eyes of the mind, as those of the body: and the former is more frequently done with wilfulness, and yet not attended to, than the latter’.<sup>24</sup> In exploring what it would mean for a person to perform the intentional action of ‘wilfully blinding himself’ to the truth of some matter, we shall need to examine in some detail the notion of self-deception. And as we seek to identify what sort of intentional actions might contribute to self-deception, the first thing we must recognise is that the notion of literal self-deception is deeply paradoxical, if not incoherent.

### III. THE PARADOX OF SELF-DECEPTION

There is nothing odd about the idea that a person may lie to himself. A novice archer may tell himself that he will certainly win an upcoming competition against archers of far greater skill; and an airline passenger may tell himself in the midst of a turbulent flight that he has already landed and is back safely at home with his loved ones. What *does* seem questionable is the further suggestion that one might be successful in getting oneself to believe one’s own lies. As Kant remarked, ‘a lie requires a second person whom one intends to deceive, and intentionally to deceive oneself seems to contain a contradiction’.<sup>25</sup> In order for self-deception to occur in a strict, literal sense, one would have to believe some proposition in one’s role as deceiver and at the same time believe the negation of that proposition in one’s role as deceived. But is it possible for a person to be in such a state?

Defenders of literal self-deception have commonly wanted to account for the possibility of such a state by positing some theory of ‘systems’ within persons, where the differing beliefs of the deceiver and deceived belong to different systems. Following Freud, such theories generally assign the role of deceiver to an unconscious system that censors uncomfortable beliefs from consciousness. But even if such rather speculative theories do show that literal self-deception is possible, there is still the question of just how often we need to appeal to the notion of literal self-deception to explain human behaviour. Alfred Mele argues that, in the vast majority of cases we typically associate with self-deceptive behaviour, people do not act with the intention of literally deceiving themselves.<sup>26</sup> Consider the mother who believes in her son’s innocence despite contrary evidence that the neutral observer finds

overwhelming. If she selectively attends to evidence with the result that she is ‘blinded’ to the truth about her son, is her behaviour best explained by supposing that she (at some level) believed that her son was guilty and then pursued the purpose of deceiving herself? Mele argues that it is more plausible to think that the mother’s purpose in selectively attending to evidence is simply to avoid unpleasant thoughts. Mele compares self-deception to akratic action on this point.

When an agent acts incontinently he acts intentionally; but that he act incontinently typically is not part of his intention. He does not aim at acting incontinently; this is not part of the action-plan that he wants to put into effect. Although, against his better judgment, he intentionally eats the sweet before him, he is not (at least characteristically) intending at the time to act incontinently.<sup>27</sup>

So, we might imagine that a given person has a desire that *p* be the case. We might further suppose that she is fully aware that she has this desire. We might even suppose that this desire motivates her to perform the intentional action of manipulating data in some manner – with the result that she comes to believe that *p* is the case. Mele’s point is that it is a mistake to think that this scenario warrants the conclusion that her *intention* in acting was to deceive herself into believing something she knew was not the case. Put another way, on Mele’s account, one’s motivation in manipulating data is simply a desire for things to *be* a certain way – rather than the desire that one come to believe that things are a certain way.

In the end, we should acknowledge that it is *possible* to interpret instances of self-deceptive manipulation of data along Freudian lines. However, it is just as possible to explain self-deceptive behaviour – where a person’s desires cause her to manipulate data with the result that she believes what she desires to believe – without appealing to Freud’s more speculative thesis involving the intentional, deceptive actions of an unconscious mental system.

If we construe self-deceptive behaviour along these more conservative lines, we might ask just how appropriate the term ‘self-deception’ really is. Purposeful, interpersonal deception clearly involves the intention of one party to lead another party to believe something that the first party knows not to be the case. This description, of course, accords well with a Freudian account of literal self-deception, where one’s unconscious system acts with the intention that one’s conscious system believe something that one’s unconscious system knows not to be true. However, on the more conservative – and, it seems to me, more plausible – account of self-deceptive behaviour, one does not simultaneously hold contradictory beliefs. If, in the vast majority of cases we typically associate with self-deceptive behaviour, one does not actually act with the intention of deceiving oneself, then is the term ‘self-deception’ misleading?

Even if the term is a bit misleading in a strict sense, there are conceptual reasons why we might want to retain the term 'self-deception' in describing certain kinds of human behaviour. Despite the dissimilarities between interpersonal deception and our more conservative account of self-deceptive behaviour, there remain some central points of correlation. In both cases a person's desire motivates her intentionally to manipulate data. Also, in both cases the manipulation of data may lead to some person<sup>28</sup> believing something that, *ceteris paribus*, she would not otherwise believe. And this resultant belief, in both cases, helps to satisfy the initial desire that motivated the manipulation of data.

So let us keep the term 'self-deception' and specify that, in our discussions, it will refer to a process the result of which is that one is 'blind' to some truth one would otherwise recognise. This process begins with some desire, or 'hot' motivation, that things in the world be a certain way. Acting on this desire, one performs the intentional action of manipulating data – with the result that one fails to see something one would otherwise have seen: *Viz*, that one's evidence on balance points to the truth of some proposition. Even though one might not act with the intention literally to deceive oneself, it nonetheless seems appropriate to describe this process as wilful self-deception. For in this process one is motivated by the uncomfortable nature of some proposition to perform some act that contributes to one's own failure to see that proposition as probably true.

Having sketched an account of what 'self-deception' amounts to, we now need to examine in more detail what kinds of decisions might contribute to the self-deceptive process. It is to this question that we now turn.

#### IV. SELF-DECEPTIVE STRATEGIES

We can begin our examination by categorising self-deceptive strategies as being one of two types. First, there are ways in which a person can bias her own interpretation of information she is considering. Second, there are ways in which a person can influence which pieces of information are available to her for further consideration. As an example of the first type, we might imagine a scenario where Tom is hopelessly in love with Tina. Tina has consistently declined Tom's invitations to dinner and has indicated to Tom that she doesn't much care for his general demeanour. Tom's friends may tell him what should be obvious for all to see: Tina simply is not interested in Tom. Yet, because of Tom's desire that Tina love him, he may continue to interpret her behaviour as coming from one who very much *is* interested but who is 'playing hard to get' as a way of encouraging his attention.<sup>29</sup> Tom has the same information that his friends have; and we can suppose that, were it not for his strong desire, Tom would reach the same conclusion that his friends have reached.

However, his strong desire has clearly affected his interpretation of the data he is considering.

Similarly, we might suppose that a mother is told that her teenage son has been arrested for drug distribution. The mother acknowledges that the arrest *is* evidence that her son is guilty, but there is also the evidence that he was a kind, innocent, naïve boy just a few short years ago – certainly not the kind of boy who would turn to a life of crime in just a few years.<sup>30</sup> When the mother considers the evidence for and against her son's guilt, she may be repeatedly drawn to mental images of her son doing things from previous years such as blowing out birthday candles and playing in the garden. If the mother did not have such a strong desire that her son be innocent, she would readily dismiss this evidence as not particularly good evidence. However, because the evidence of the mental pictures is so comforting and thus vivid to her, it is not difficult to imagine that she might come to give this evidence undue weight and subsequently conclude that her son must be innocent.

We might expand on the example of the mother to turn our discussion to the second type of self-deceptive strategy: Namely, the influencing of which pieces of evidence are available for one's further consideration. We can imagine that the mother's mental images of her son during happier times provide a great deal of emotional comfort. Subsequently, she may find these mental pictures dominating her thoughts whenever she considers the question of her son's guilt. Given that her focus on these mental pictures comes at the expense of possible reflection on contrary evidence, her consideration of evidence becomes (unbeknownst to her) selective. And the selective consideration of evidence is an obvious enough means by which the process of self-deception might occur.

To cite another example, we might imagine a literary historian who publishes a book claiming that Shakespeare did not write many of the works attributed to him. We imagine that his book sparks a host of responses in the form of articles from other historians. Because the historian very much wants his original thesis to be correct, he may specifically look for all the articles he can find by historians he suspects will be sympathetic to his own position. Conversely, he may read only a sampling of articles by those of his fellow historians whom he suspects will not receive his position with much charity. Thus, as the historian investigates further his original hypothesis, his consideration of evidence will be selective and therefore biased. Moreover, we might imagine that, in reading through the articles he does survey, the historian adopts the following practice. When he comes to a passage supporting his original thesis, he says to himself, 'Yes! That's right', and he then re-reads the passage. However, when he comes across data contradicting his own hypothesis, he typically reads quickly further along, hoping that the author will eventually offer a rebuttal to this data. As a result of this pattern of reading, the historian might well remember small and detailed

pieces of supporting evidence – despite failing to remember more obvious and compelling contrary evidence.

In discussing how a person might influence which pieces of evidence are available for her further consideration, we have thus far focused on how one might gather and attend to certain evidence at the expense of gathering and attending to contrary evidence. In addition to restricting in this way the flow of evidence against a given conclusion, humans also seem quite capable of *generating* evidence supportive of a given conclusion. Suppose that Jane has a disagreement and subsequent unpleasant verbal exchange with a store clerk while shopping one day. When she reflects on the event, she may be resistant to the idea that she is in any way to blame for the angry escalation of words. Inasmuch as her desire to be the innocent party colours her interpretation of the words that were actually spoken, she might be described as considering evidence in a biased manner. However, after repeated reflection upon the event, she may come to consider evidence that was *not* part of the actual exchange that took place. As she rehearses the scene again and again in her mind – playing, of course, both the role of herself and the store clerk – her own lines become more and more innocuous as the clerk’s lines become more and more acerbic. Her original statement, ‘I need to return this’, eventually becomes ‘I wonder if I could return this’. The clerk’s original statement, ‘Do you have a receipt?’, eventually becomes ‘Well, where’s your receipt?’. Also, the tone of what is spoken changes, so that the clerk’s tone of voice becomes more and more rude. Given that Jane is now considering evidence that she herself generated, it is not surprising that she would reach the conclusion that she is in no way to blame for the argument. And if the store clerk reflects on events in a corresponding manner, then it would not be surprising if he were to reach a completely opposite conclusion. In everyday parlance we might say of Jane that she has ‘convinced herself’ that she was in the right.

Another prominent method by which one can generate evidence involves the notion of pretending. A miserly man might yet wish he were a charitable person – or, at least, wish that others would view him as a charitable person. Such a man might then decide to act like a charitable person would act. He might say ‘Hear, hear!’ when speeches are made about helping the poor, and he might adopt phrases that he has heard fêted philanthropists use. By pretending in this way to be a charitable person, the miserly man might actually come to believe that he is more charitable than he actually is. As Daryl Bem has argued,

Individuals come to ‘know’ their own attitudes, emotions, and other internal states partially by inferring them from observations of their own overt behavior and/or the circumstances in which this behavior occurs.<sup>31</sup>

And in addition to the ways in which our beliefs about ourselves can be influenced by our own observations of our own pretending, Daniel

Gilbert and Joel Cooper point out that pretending can have far-reaching effects in virtue of the fact that it is a strategy that involves other people. Given that our conceptions of ourselves hinge largely on our perception of how others perceive us, Gilbert and Cooper note that ‘the targets of our self-presentations may reciprocally convince us of the validity of their impressions’.<sup>32</sup> The two cite a number of studies in support of the links in this ‘self-presentational feedback loop’, and they maintain that the complex nature of social interaction can blur the distinction in our own minds of where our pretending ends and our true attitudes begin.

If our behavior is shaped by subtle social forces, we may fail to recognize these as the true causes of our actions, concluding that we do indeed possess the dispositions we were once merely pretending to possess.<sup>33</sup>

In the end, the social aspect of pretending does seem to allow for a kind of ‘feedback loop’ through which a good deal of far-ranging self-deception can occur. And this point is only magnified when we consider that the initial act of pretending may be subtle and sometimes difficult even for the pretender to recognise – as when one does not recognise that one’s own noble attempt to ‘put on a brave face’ is in fact partly an effort at pretending that one’s past enterprises were not foolhardy. We end this section, then, having identified a number of different ways in which one’s decisions can contribute to one’s failure to form beliefs one would otherwise have formed.

#### V. MORAL RESPONSIBILITY FOR SELF-DECEPTION

Having outlined some of the kinds of decisions that can contribute to a self-deceptive process, let us now consider the extent to which one is morally responsible for making such decisions. Some of the examples of self-deception from the previous section provide obvious ways in which one might make decisions for which one is morally responsible. Pretending that something is so and rehearsing scenes in one’s mind are two obvious possibilities. If one believes that it is wrong to pretend to be charitable or to rehearse repeatedly in one’s mind the blameworthy behaviour of a store clerk, then to do so would be to perform an act for which one is morally culpable.

But what of Tom, who misinterprets Tina’s refusal of his advances as an attempt to ‘play hard to get’? Or what of the mother who believes that her son is innocent? In these examples, the person in question somehow manipulates data in such a way that s/he comes to believe something that perhaps we are inclined to think s/he really should not believe. But before we can talk about moral responsibility for the self-deceptive process in either case, we must first identify a decision the person makes for which s/he might be morally responsible. Yet, in each case it is not readily

apparent that any *decision* led to the failure to believe some truth that should have been obvious.

Now, we might suppose that in each case the subject had moments when the truth began to dawn on him or her. It is common enough to have doubts creep to mind as to whether that which one believes might really be false after all. If we imagine this to be the case in the examples under consideration, we might imagine that the subjects wilfully chose to ignore such doubts when they arose, turning their thoughts instead to other matters.

Even if we assume that the subjects in our examples did not make choices of this kind, it still would be too quick a move to conclude that no choices by the subjects whatsoever were involved in the self-deceptive process. For there might have been any number of other, more subtle decisions made by the subjects that contributed to their own deception. Take the case of Tom, who misinterprets Tina's refusal of his romantic advances. If Tom does not experience any nagging doubts about the accuracy of his own interpretation, it might nonetheless occur to him to ask certain questions. For example, he might think to ask his friends whether they agree with his interpretation. He might also think to reflect on whether his romantic desires have previously led to mistaken conclusions regarding the attitudes of women he has pursued in the past. The choice not to pursue such questions might well amount to a choice that contributes to his own deception. And if Tom believes that he might have an obligation to ask such questions, then he would be morally culpable for a refusal to pursue the matter.

If we suppose that no such questions occur to Tom, we might ask why this is so. Perhaps Tom is simply not in the habit of asking such questions, and it therefore does not occur to him to ask them. We might then ask how Tom came to have this particular habit (of not asking). If the habit owes its development in part to past refusals to pursue such questions when they occurred to him, then the contributing factors in his current deception *do* include choices he has made. Indeed, it seems possible to imagine any number of decisions Tom could have made in the past that might be part of the process of self-deception that led to his current beliefs about Tina's hidden motives in refusing his advances. And if Tom believed at these earlier points that it might be wrong for him not to ask such questions, then the conditions for moral culpability are again met. In the end, there may be any number of morally significant decisions Tom makes that contribute to his failure to see the truth about Tina.

Let us turn now to consider the specific question of how someone might be morally responsible for a failure to recognise a purportedly divine communication as in fact authentic. We should first note that there is any number of reasons why a person might be motivated to engage in self-deceptive behaviour in such a matter. That is, the belief 'that a purportedly divine statement is (or might be) true' might be

uncomfortable for any number of reasons. One might have invested time and energy in trying to disprove the existence of God and therefore desire that one's initial hypothesis be proved right. One might dislike certain religious proponents and desire that they be proved wrong in their beliefs. Perhaps the most potentially powerful motivation is that the purportedly divine communication has implications as to what one's moral obligations might be.

Take, for instance, the affirmations in Christian scripture as to Jesus's divinity and to what directives he issued. Jesus's recorded directives may well be uncomfortable to an individual because they preclude various actions she desires to perform. Indeed, the recorded directives of Jesus involve such sensitive areas as one's finances, vocation and sexual practices. Thus, it is not difficult to see how one might be motivated to reach the conclusion that Jesus was not in fact divine. Given that the motivation to avoid uncomfortable moral obligations seems potentially the most powerful motivation to engage in self-deceptive behaviour on religious issues, let us focus our remaining discussion on this motivation.

An active resistance to what one believes to be one's moral obligations may take any number of forms. An obvious example is a decision to perform some action that one believes God has identified as wrong. Another slightly less obvious example is the decision wilfully to turn one's attention away from the question of whether it is possible that God may in fact have identified that action as wrong. There are also much less obvious examples of resisting what one believes are, or might be, one's moral obligations. Take, for instance, the various ways in which one might resist the truth about oneself. C. R. Snyder, in discussing strategies of excuse-making,<sup>34</sup> observes that a person whose behaviour does not meet her own personal standards may engage in a 'reframing performance strategy', where she works to diminish the extent of her failure. Thus, if a person finds that her behaviour conflicts with what she believes or suspects to be the truth about God's commands, she may remind herself and others that her behaviour does not really hurt anyone or that her 'sin' is not one of the more serious ones. Alternatively, Snyder points out, a person may engage in a 'transformed responsibility strategy', where she acknowledges her failure but downplays her responsibility for it. She may, for instance, take a 'consensus-raising approach' by reminding herself and others that everyone fails in this manner; or she may perhaps take a 'consistency-lowering approach' by seeking to explain her failure as a temporary and uncharacteristic lapse. There may be no end to the ways in which one can seek to avoid the conclusion that one's own behaviour falls short of what God has, or may have, commanded or encouraged. We may sabotage the efforts of friends so that our own behaviour appears good by comparison.<sup>35</sup> We may seek to avoid conclusions about our own behaviour and motives by taking performance-inhibiting drugs such as alcohol so as to handicap our own



performances and thereby create a ready-made, external explanation for our moral failings.<sup>36</sup> We may even in certain cases embrace the diagnosis of a mental or physical illness because the conclusion 'I'm unwell' is more comfortable than the conclusion 'I'm wilfully engaged in unacceptable behaviour'.<sup>37</sup>

Instances of resisting the truth about oneself in manners such as these constitute subtle ways one might resist such moral conclusions as: 'God may want me to act in a certain way, and I may be in the wrong if I do not act in this way'. We have already explored some of the many ways in which one's active resistance to some proposition *p* can lead to one's failure to see that *p* is true. Thus, it is not difficult to imagine how a person's active resistance to the kind of moral conclusion outlined above can contribute to a self-deceptive process whereby one fails to believe that this proposition (and others like it) are true.

By contrast, if a person does *not* resist, but instead attempts to act in accordance with, her perceived moral obligations, then she can in no way be described as engaging in morally culpable self-deception. If such a person fails to believe, e.g., that the Christian scriptures contain authentic instances of divine communication, this 'spiritual blindness' cannot rightly be described as a wilful or morally culpable condition. From the Christian perspective, of course, her beliefs about the lack of authentic divine communication in the scriptures will be *false* beliefs. But if she has consistently acted in accordance with what she *believes* to be the demands of morality, she has done all she could do. Any remaining disparity between the objective facts of God's revealed commands and her own subjective beliefs about these commands cannot be attributed to any sort of 'culpable blindness' or 'self-deceit'.

In support of this characterisation of wilful spiritual blindness as a refusal to respond positively to moral truths one recognises, we might look again to the writings of St. Paul in the first part of Romans. We earlier noted Paul's declaration here that 'what may be known about God' is 'clearly seen' by those who do not 'suppress the truth by their wickedness'.<sup>38</sup> Paul goes on in this passage to cite one's stance toward truth in identifying that which humans do to incur God's judgment. He writes that, 'for those who are self-seeking and who *reject the truth* and follow evil, there will be wrath and anger'.<sup>39</sup> I take the account of culpable spiritual blindness outlined in this section as a plausible spelling out of Paul's general line of argument here.

## VI. CONCLUSION

In this essay I have explored the notion of wilful spiritual blindness, and I have sought to shed light on this notion through an analysis of morally culpable self-deception. I have not put forward any particular view as to

the extent to which spiritual blindness does account for the failure of actual nonbelievers to form Christian beliefs. Some non-theists, of course, have contended that *theistic* beliefs are themselves prime examples of wish fulfilment and perhaps also self-deception. Feuerbach and Freud are two of the better known defenders of this line of argument. I have in no way attempted to adjudicate this dispute between theists and non-theists over whose beliefs are in fact the result of wish fulfilment or self-deception. In the end, perhaps both parties are right with respect to some people. That is, perhaps it is the case that the beliefs both of some theists and of some atheists are formed as a result of their wishing things to be a certain way.

My project has been merely to make sense of the idea – put forward by writers such as John Calvin and Jonathan Edwards – that someone might fail to hold Christian beliefs due to one's own 'wilful spiritual blindness'. Toward that end, I have outlined the conditions that would need to be met in order for a person rightly to be described as culpable for her failure to hold certain Christian beliefs – e.g., the belief that a particular message does in fact come from God and accurately identifies where her moral obligations lie. When a person considers the possibility of such a truth, she may find the possibility uncomfortable. Perhaps, e.g., the message speaks of a moral obligation that conflicts with certain behaviours in which she desires to engage. Motivated by a desire that the content of the message be false, she may engage in intentional actions that she believes to be wrong and that contribute to a self-deceptive process the result of which is that she fails to believe what she otherwise would have believed: Namely, that the message under consideration does in fact come from God and does accurately identify where her moral obligations lie. In such a scenario, a person can rightly be described as culpably 'blinding herself' to some spiritual truth.

## Notes

1 Bernard Williams, 'Deciding to Believe', reprinted in *Problems of the Self* (Cambridge: Cambridge University Press, 1973), p. 148.

2 Aristotle, *Nicomachean Ethics* (Oxford: Clarendon Press, 1925), IV, 1140<sup>b</sup>.

3 St. Bonaventure, *Disputed Questions on the Mystery of the Trinity* (St. Bonaventure NY: The Franciscan Institute, 1979), q. I, i, 1.2.3.

4 John Henry Newman, *The Theological Papers of John Henry Newman on Faith and Certainty* (Oxford: Clarendon Press, 1976), p. 27.

5 Newman, *Theological Papers*, p. 27.

6 William James, *The Will to Believe and Other Essays in Popular Philosophy* (New York: Dover Publications Inc., 1956), p. 10.

7 Jonathan Edwards *The Works of Jonathan Edwards* (New Haven, CT: Yale University Press, 1970), III, 157.

8 John Calvin, *Institutes of the Christian Religion* (Albany, OR: Sage Software, 1996), I, v, 1.

9 Calvin, *Institutes*, I, iii, 2.

10 Calvin, *Institutes*, I, ii, 2.

11 Edwards, *Works*, II, 307.

12 Nicholas Wolterstorff, 'Is reason enough?' in R. D. Geivett and B. Sweetman (eds.), *Contemporary Perspectives on Religious Epistemology* (Oxford: Oxford University Press, 1992), p. 145.

13 Francis Bacon, *Novum Organon* (Oxford: Oxford University Press, 1855), I, §xlx. Strictly speaking, Bacon's conclusion does not seem to be universally true. Fear, for example, may lead a child to believe that there are monsters lurking under his bed – and this despite his desire that all monsters be kept well away from him. Still, our concern will be with cases of the kind Bacon describes here.

14 Bacon, *Novum*, I, §xlx.

15 David Pears, *Motivated Irrationality* (Oxford: Clarendon Press, 1984), p. 12.

16 This example and others like it are provided by Alfred Mele, *Irrationality* (Oxford: Oxford University Press, 1987), pp. 154–5.

17 Daniel Kahneman and Amos Tversky, 'Subjective probability: a judgment of representativeness', *Cognitive Psychology* 3 (1972), p. 452.

18 See Daniel Kahneman and Amos Tversky, 'Availability: a heuristic for judging frequency and probability', *Cognitive Psychology* 5 (1973).

19 Cf. Kahneman and Tversky, 'Availability', p. 432.

20 Richard Nisbett and Lee Ross, *Human Inferences: Strategies and Shortcomings of Social Judgment* (Englewood Cliffs, NJ: Prentice-Hall Inc., 1980), p. 46.

21 Nisbett and Ross, *Human Inferences*, p. 47.

22 Nisbett and Ross, *Human Inferences*, pp. 49–50.

23 Thomas Aquinas, *Summa Theologica* (Westminster, MD: Christian Classics, 1981), I–II, lxxvi, 4.

24 Joseph Butler, *The Works of Bishop Butler* (London: Macmillan and Co. Ltd., 1900), p. 133.

25 Immanuel Kant, *The Doctrine of Virtue* (Philadelphia, PA: University of Pennsylvania Press, 1964), p. 94.

26 See Mele, *Irrationality*, especially chpt. 9.

27 Mele, *Irrationality*, p. 124.

28 In the case of self-deception, this 'person' is, of course, oneself.

29 Cf. examples in Mele, *Irrationality*, pp. 125–6.

30 Cf. examples in Mele, *Irrationality*, pp. 115–6.

31 Daryl Bem, 'Self-perception theory', *Advances in Experimental Social Psychology* 6 (1972), p. 2.

32 Daniel Gilbert and Joel Cooper, 'Social Psychological Strategies of Self-Deception' in Mike Martin (ed.), *Self-Deception and Self-Understanding* (Lawrence, KS: University Press of Kansas, 1985), p. 83.

33 Gilbert and Cooper, 'Social Psychological Strategies', p. 84.

34 See C. R. Snyder, 'Collaborative companions: the relationship of self-deception and excuse making' in Mike Martin (ed.), *Self-Deception and Self-Understanding* (Lawrence, KS: University Press of Kansas, 1985), especially pp. 36–42.

35 Cf. Abraham Tesser and Jonathan Smith, 'Some effects of task relevance and friendship on helping: you don't always help the one you like', *Journal of Experimental Social Psychology* 16 (1980).

36 Cf. Edward Jones and Steven Berglas, 'Control of attributions about the self through self-handicapping strategies: the appeal of alcohol and the role of underachievement', *Personality and Social Psychology Bulletin* 4 (1978).

37 Cf. C. R. Snyder, R. L. Higgins and R. J. Stucky, *Excuses: Masquerades in Search of Grace* (New York: Wiley, 1983), especially pp. 220f.

38 Romans 1:18–20.

39 Romans 2:8 (emphasis added).