

EURAMERICA Vol. 40, No. 3 (September 2010), 591-633  
© Institute of European and American Studies, Academia Sinica  
<http://euramerica.org>

## The Ethics of False Belief\*

Timothy Lane

Research Center for Mind, Brain, and Learning, National Chengchi University  
No. 64, Sec. 2, ZhiNan Rd., Taipei 11605, Taiwan  
E-mail: [tlan@nccu.edu.tw](mailto:tlan@nccu.edu.tw)

### Abstract

According to Allen Wood's "procedural principle" we should believe only that which can be justified by evidence, and nothing more. He argues that holding beliefs which are not justified by evidence diminishes our self-respect and corrupts us, both individually and collectively. Wood's normative and descriptive views as regards belief are of a piece with the received view which holds that beliefs aim at the truth. This view I refer to as the Truth-Tracking View (TTV). I first present a modest version of TTV, one which is sensitive to standard criticisms and one which is fully consistent with the procedural principle. I then raise some doubts about TTV by considering both anecdotal cases and empirical studies. These studies suggest that certain types of belief are designed to aim away from truth, in limited, carefully calibrated ways. Moreover, it seems to be the case that selectively aiming away from the truth is important for human well-being and performance. Beliefs that are designed

---

Invited article, Received March 9, 2010

Proofreaders: Hsin wen Fan, Chia-chi Tseng, Ying-bei Wang

\* I am grateful to Tim Bayne and Eric Schwitzgebel for their careful reading of and helpful comments on previous versions of this manuscript. Work on this manuscript was, in part, funded by a National Science Council of Taiwan research grant, 97-2410-H-004-154-MY3.

to aim away I dub “Tertullian” beliefs (t-beliefs). I then list the distinguishing characteristics of t-belief and proceed to evaluate the procedural principle in light of the evidence which suggests that t-belief plays an important role in our cognitive economy. Next I argue that t-beliefs might be essential to the maintenance of self-respect and that they do not corrupt in the way that Wood claims. Finally, I argue that the fate of Wood’s procedural principle will be determined by the results of further empirical research—sociological, psychological, and neuroscientific.

**Key Words:** belief, procedural principle, positive illusions, depressive realism, and Moore’s paradox

## I. Introduction

Allen Wood (2002, 2008) argues that the main principle governing the ethics of belief is the “procedural principle.”<sup>1</sup> According to this principle we should apportion the strength of our beliefs to the evidence. In other words we should believe “only what is justified by the evidence, and believe it to the full extent, but only to the extent, that it is justified by the evidence” (2008: 9). Wood qualifies this claim in certain respects (2008: 13, fn. 8) and, grudgingly, acknowledges the possibility of non-trivial exceptions (2002: 38). Nevertheless, he concludes that failing to adhere to this principle invariably violates our self-respect and is, as well, in other regards inevitably corrupting (2002: 36, 2008: 24). Similar sentiments have been expressed by other philosophers, such as Michael Lynch (2004: 143), who writes: “Caring about truth and believing the truth about what you care about are necessary parts of happiness by being necessary parts of integrity, authenticity, and self-respect.”

Wood is aware of the body of empirical work which suggests that people benefit from holding certain false beliefs, as well as beliefs not supported by evidence. But for various reasons he denies that this work counts against the procedural principle. For example, he (2008: 13, fn. 8) proclaims that “no one could stably hold both the belief that is supposed to benefit them and also know that it is false . . . even if illusions do benefit people’s health, it does not seem that this is justification a person could stably or self-consistently apply to their own beliefs.” Note that Wood seems to be making an empirical claim about the nature of beliefs.

Wood’s views as regards both normative and descriptive

---

<sup>1</sup> What I here refer to as the procedural principle is also known as “Clifford’s Principle” (Wood, 2002) or the “evidentialist principle” (Wood, 2008: 10). I use “procedural” rather than “Clifford’s” to emphasize that my concern is with Wood’s version, and I refrain from using “evidentialist,” because this term is more often employed by the principle’s critics than by its advocates.

aspects of belief are consistent with the received view of beliefs, viz. that they “aim at the truth” (Williams, 1973: 137-138).<sup>2</sup> Many, perhaps a majority, of late 20th and early 21st century philosophers have converged on the view that beliefs are constituted in such a way that they can be accurately characterized by this phrase. Donald Davidson emphasizes their “veridical nature” (2003: 366-367) and he (1977: 295) argues that “successful communication proves the existence of shared and largely true, view of the world;” John R. Searle (2001: 37-38, 257) claims that it is their “job” to “represent how things are;” Peter A. Railton (2003: 297) holds that belief “not only represents its propositional content as true,” it “cannot represent itself as unresponsive to—unaccountable to—their truth;” Tim Crane (2001: 103) says that “holding true” is a synonym for belief; Bernard Williams (2002: 80) claims that beliefs are “subject to a norm of truth;” and, Ralph Wedgwood (2002: 273) observes that “for every proposition *p* that one consciously considers, the best outcome is to believe *p* when *p* is true.” Wood’s view, along with this cluster of interrelated views, I refer to as the Truth-Tracking View (TTV) of belief.

I shall not concern myself with strong versions of TTV, for their vulnerabilities are conspicuous. I here consider only modest versions, of which I take J. David Velleman’s (Shah & Velleman, 2005; Velleman, 1999, 2000) to be representative. Both Wood and Velleman exemplify TTV, but Velleman provides a more detailed account. Moreover, he is in sympathy with Wood’s normative position, and is sensitive to relevant criticisms of TTV.

Velleman’s version is used below, in part, as a foil against which to develop the idea of “aiming away.” Beliefs that aim away from the truth I refer to as “Tertullian beliefs” (“t-beliefs”). Although all modest versions of TTV do qualify the sense in which

---

<sup>2</sup> Although the expression is sometimes used metaphorically, it can also be used literally, as when one is speaking of the aims of people who form beliefs or of design mechanisms that constrain the regulation of beliefs.

beliefs can be said to aim at the truth (hence, the attributive “modest”), still they lack the resources with which to account for the distinctive causal-explanatory role played by t-beliefs. The standard qualifications that are appended to modest versions of TTV, while perhaps succeeding in making them weakly compatible with instances of t-belief, also make it appear that t-beliefs are nothing but incidental, variously inconsequential, or “pernicious” (Wood, 2002: 40), features of our cognitive economy. After sketching Velleman’s account, t-belief is introduced by means of examining certain commonplace, anecdotal instances wherein behavior contravenes professed beliefs in ways that suggest belief-forming mechanisms are not responsive to evidence in the ways required by TTV. Next empirical studies of beliefs that aim away from the truth are reviewed. Then the distinctive characteristics of t-beliefs are limned, such that they can be clearly distinguished from other attitudes like desire, hope, or hypothesis. Finally, I evaluate Wood’s procedural principle in light of what we are now learning about t-beliefs. I argue, pace Wood, that the capacity to occasionally and strategically aim askew of the truth might be essential to the maintenance of self-respect and that it is not necessarily corrupting in the ways that he suggests. In a brief concluding section I suggest that whether or not Wood is correct in his uncompromising advocacy of the procedural principle will ultimately be determined by the results of empirical research—sociological, psychological, and neuroscientific.

## II. A Modest Version of TTV

On Velleman’s (2000: 255) version of TTV, belief is constituted both “by its power to cause behavioral output,” and by “its responsiveness to epistemic input.” It is not sufficient to claim that belief takes its propositional object as representing the way things are, for this alone could not distinguish it from certain other attitudes (Velleman, 1999: 198-200). What distinguishes belief

from, say, assumption or imagination is “the spirit” in which a propositional object is regarded as true: an assumption might be “tentatively” held and something imagined might be “fancifully” held, but a belief is “seriously” held. Fantasies and assumptions are not “serious” because they entail accepting a proposition as true without sensitivity to whether a person is “accepting” the truth. To believe is not to “accept” for polemical or heuristic purposes (as is the case with assuming), neither is it to “accept” for recreational or motivational purposes (as is the case with imagining); instead, to believe is to accept a proposition “with the aim of doing so if and only if it really is true” (Velleman, 1999: 200; see also Wood, 2002: 19-20).<sup>3</sup> Beliefs are regulated—formed, revised, and extinguished—in truth-conducive ways, in ways that are responsive to evidence and reasoning (Shah & Velleman, 2005: 498).<sup>4</sup>

To say that beliefs aim at the truth is not to say that the aim is to believe as many truths as possible; nor is the aim to believe as many as possible useful or valued truths; nor indeed is it to say that the aim is maximizing the proportion of truths to falsehoods among one’s beliefs (Velleman, 2000: 251-255). TTV requires only that beliefs aim at the truth in some way, while allowing that there are multiple ways in which they might do so. It further allows that belief is not exclusively governed by truth-seeking mechanisms (Shah & Velleman, 2005: 500-501; Velleman, 2000: 254): some mechanisms may cause beliefs that occasionally diverge from the truth (the adoption of better-safe-than-sorry strategies). But Velleman holds that belief is necessarily subject to mechanisms

---

<sup>3</sup> Belief and imagination can be combined though, as in cases of metaphorical belief (McGinn, 2004: 134). A simple example would be employing a simile to express one’s belief, such as “the sky is like the ocean.”

<sup>4</sup> Although Wood speaks more of evidence (empirical, a priori, etc.) than of truth, it is clear that what matters for both is this—responsiveness to evidence and reasoning (2008: 10). He makes the implied conceptual relationship between evidence and truth explicit when he writes that there is “no other responsible guide to what beliefs are true than that which the evidence indicates” (2002: 70-71).

designed to make it true: “the input constraints definitive of belief are designed to yield beliefs that are true” (Velleman, 2000: 277). In other words, mechanisms that are not truth-seeking are not definitive of belief.

Because only some beliefs are caused by the goal-directed activity of persons, and many are the results of processes that do not involve agential goals or intentions, the concept of belief must include more than just the manner in which beliefs are actually regulated. A “standard of correctness,” a normative standard, must also be applied (Shah & Velleman, 2005: 498-500). Modest TTV then conjoins the descriptive and the normative: belief is regarded as “truth-regulated acceptance.” And to this a norm of truth is then applied. Norms governing belief are understood in a “biconditional” sense (Shah & Velleman, 2005: 519): although they do not require acceptance of every belief that would be correct, they do forbid the holding of a belief that would be incorrect. These norms may be lax in what they require a person to accept, but they are strict in what they prohibit—the holding of incorrect beliefs. Wood (2008: 10) expresses this prohibition thus: “beliefs not justified by the evidence are *immoral*.”

Velleman (2000: 277-279) considers the possibility that beliefs might aim at “instrumental success” or “empirical adequacy.” But he claims that while we might sometimes settle for an alternative to truth as a “second-order” aim, truth remains the “first-order” aim. Suggesting an analogy he observes that a basketball player might proclaim that his ultimate aim is to earn a salary increase, but fans don’t thereby presume that everything he does on the court is aimed at the salary increase, because the best way to achieve the salary increase is to aim at victory itself. Money might be the object for playing the game, but within the context of the game winning is adopted as the aim. “Similarly, we may enter the game of having beliefs on a particular subject because we want our motivating cognitions on that subject to yield successful actions; but success in action does not thereby become the object of the game.”

Velleman (2000: 278) is dismissive of the possibility that we might discover beliefs to be regulated so as to aim at something other than truth. He justifies this dismissal with a claim about the content of our introspections: when we discern a gap between a belief and the truth, the belief becomes unsettled and starts to change (see also Wood, 2008: 13, fn. 8). Alternatively, if the belief persists, another belief is formed to help close the gap, while the original belief is reclassified as an illusion or bias. Non-evidential considerations simply cannot be *explicitly* treated “as relevant to the question what to believe. Any influence that such considerations exert must be unacknowledged.”<sup>5</sup>

### III. Some Doubts About TTV

If, as Velleman contends, beliefs are constituted by their power to cause behavior and by their responsiveness to epistemic input, if they aim at the truth, and if they are indeed subject to mechanisms or constraints that are designed to yield true beliefs, then, at minimum, we might reasonably expect that when, on good evidence, a person categorically and sincerely asserts that a belief is untrue, that same person should not act as though it were true, especially when acting as though it were true incurs greater cost than would be incurred were one to act in accord with professed beliefs. But there seem to be clear cases in which people possess the relevant evidence, do so assert, and yet act as though they hold beliefs which they deny holding, even when doing so carries significant cost. And these are not, and are not relevantly similar to, cases of assumption or fantasy. What’s more, they at least seem to

---

<sup>5</sup> Shah and Velleman: “It is an objection to belief that it is false . . . it is a *fatal* objection, in the sense that if the person who has the belief accepts the objection, he thereby ceases to have the belief, or at least it retreats to subconscious . . .” (2005: 531, fn. 16). See also Williams (2002: 67). The received view of beliefs is that some are conscious, some not. But extreme positions do exist: Searle holds that all beliefs are conscious (1992); Crane (2001: 103-108), that none are.



allow for the possibility that non-evidential considerations are being *explicitly* treated as relevant to the question what to believe. At least it is not obvious that these considerations exert their influence in a manner that is wholly unacknowledged.

Consider the case of a person born and raised in the western world who categorically denies believing that the number 13 invites bad luck.<sup>6</sup> Moreover, this very same person is familiar with the arguments and the vast amount of evidence that demonstrate that 13 is no more lucky or unlucky than any other number. Nevertheless, given the option of choosing between a hotel room or an office space on the 13th or the 14th floor, all things being equal, he might well be disposed to choose the latter. What's more, for many people the same would likely hold true even if all things weren't equal; that is, even if avoiding 13 were to require greater cost. Doubtless there is a limit to just how much greater cost one would be willing to incur in order to avoid the 13th, but the expenditure of significant time, money, and other resources in the avoidance of 13 is not uncommon.

Were we to employ belief-desire psychology toward explaining the relevant behavior, we would likely say "Stan's belief that 13 is unlucky caused him to choose the 14th rather than the 13th floor," as part of our explanatory sketch. Here then we would have a reasonably clear case in which one has, for good reason, denied believing that 13 is unlucky, yet, a specific decision was prompted, *inter alia*, by just that very belief.<sup>7</sup> While the

---

<sup>6</sup> For those readers for whom the number 13 fails to evoke superstitious anxieties, substitute any superstition that does and construct a scenario parallel to the one sketched here; nothing hinges on this particular example. Gazzaniga (2008: 271-272), for example, cites the example of walking quickly past a cemetery at night, even though one doesn't believe in ghosts. The number 13 example is used only because it is familiar to a wide audience and because it has been demonstrated (Scanlon, Luben, Scanlon, & Singleton, 1993) to consistently and significantly affect behavior.

<sup>7</sup> Case (2000) and Case, Fitness, Cairns, and Stevenson (2004) have provided some experimental evidence to support the claim that even skeptics readily resort to superstition. For related material see Shermer (2002, 255-313), Talmont-

relevant psychology is insufficiently understood, it is by no means obvious that non-evidential considerations exert influence in a manner that is wholly unacknowledged.

Recall that, according to Wood and Velleman, if we discern a gap between a belief and the truth, the belief becomes unsettled and starts to change, or another belief is formed to help close the gap, while the original belief is reclassified as an illusion or bias. But in this case it seems plausible to claim that the person can be aware of the gap and that the belief neither becomes unsettled, nor does it require formation of a gap-closing belief. A balanced perspective should allow that post hoc introspection concerning cases of this sort is theory-laden reflection over “skittish” phenomena (Hurlburt & Schwitzgebel, 2007: 48-53). Hence, a charitable view of the Wood-Velleman position is that we are left with an introspective stand-off.

If cases of superstition strike some readers as too exotic, perhaps vanity products can serve as more compelling examples. Alleged cures and treatments for alopecia (and the many other assaults on personal vanity) are as numerous as their evidence is wanting. But just as with superstitions, there are many people who categorically deny believing that these products can promote good health or restore one to a hirsute state, yet they act as though they hold the very beliefs they deny holding, even when doing so carries significant cost.<sup>8</sup> Not only are the claims unsubstantiated, positive reasons not to believe the claims are plentiful; yet, intelligent

---

Kaminski (2008), and Vyse (2000). Talmont-Kaminski generalizes from the data to assert that superstition is a basic human trait.

<sup>8</sup> It might be thought that in contexts like this belief should be understood in a Bayesian way, i.e. as the assignment of probabilities to statements. But to do so would mislead, for the subjects express categorical denial. Moreover, on the Bayesian construal, it is perhaps more aptly said that beliefs are just “tentatively” held; therefore, it would be incompatible with TTV. Finally, although probabilities of statements can be applied in certain situations, still it would seem that those situations must then be believed to be of that type by a subject. In other words, Bayesian conceptions seem to presuppose the attribution of non-Bayesian beliefs (Nozick, 1993: 94-99).

consumers behave in ways that contravene professed beliefs. As is the case with superstition, here too the possibility of acknowledged, non-evidential considerations playing a role in belief regulation cannot be dismissed out of hand.

Some contexts, in particular those that are harrowing or life-threatening, can help to further illustrate the point about superstition and about marketing gullibility. Consider the case of a medical doctor (or scholar, or scientist) who is quite convinced on extremely good evidence that herbal treatments like echinacea cannot prevent the common cold and prayer cannot cure bone cancer.<sup>9</sup> But when the throat begins to feel raw, or while awaiting the results of the biopsy, some among these very same people are highly disposed to purchase echinacea or stop by a temple, church or synagogue. As for herbs like echinacea, since the common cold is typically just a nuisance, one might wonder why the person who categorically denies believing in its effectiveness would be so easily motivated to behave in accord with the belief that it is effective.<sup>10</sup> The threat of bone cancer is of course another matter though: desperate to cling to life one might grasp at any measures, no matter how far-fetched, and without regard to whether the person has spent a lifetime emphatically not believing in the method that is being tried. Desperation trumps justification.

Perhaps it might be argued that when confronted by life-threatening illness we abruptly adopt an assumption, an

---

<sup>9</sup> I am presupposing that few people have the intellectual courage of a John Diamond (2001), who steadfastly refused to yield to superstition or ungrounded claims, even though he was gravely ill. Instead, he devoted his time to his attempt to complete "Snake Oil," his critique of alternative medicine. Also, note that "for the most part intelligence is orthogonal to and independent of belief" (Shermer, 2002: 285), that educational level does not influence susceptibility to superstition (Case, 2000), that maintaining a dubious attitude toward a proposition requires energy (Gilbert, 1993), and that stress and uncertainty incline one to resort to superstition (Keinan, 2002).

<sup>10</sup> Bausell (2007) provides a detailed survey of complementary and alternative medicine (CAM); a preponderance of evidence shows that nearly all CAMs, including echinacea, are ineffective.

attitude that need only be “tentatively” held, like a heuristic, in order to motivate experimentation with herbs or prayer. But to characterize this attitude as a heuristic, after a life-time of deliberate, well-considered, disbelief, would be odd. Typically assumptions are adopted as a means of exploring the unfamiliar, in an attempt to gain new knowledge. But that does not seem to be a straightforward characterization of what is happening here, for by hypothesis these are cases concerning which the person previously explored the relevant claims and, for good reason, rejected them. It seems more natural to say that desperation alters belief regulation such that one becomes strongly influenced by non-evidential considerations. If this latter characterization is correct, then one explicitly treats non-evidential considerations as relevant to the question what to believe.

Velleman allows that people will sometimes choose to error on the side of caution, as when worried about potential predators. But this does not help us to explain the avoidance of 13 or the abrupt decision to behave in accord with beliefs that one rejects. Many people are familiar with the relevant evidence (we might suppose them to be avid readers of the *Skeptical Inquirer* and like material), so, unlike wilderness predation, there simply is no reason to be wary of 13 and no reason to suddenly embrace prayer. Be that as it may, people do behave in these ways.

#### IV. The Empirical Study of False Beliefs

Proper characterization of the preceding examples is contentious. They are anecdotal and their interpretations, uncertain, in part due to the vagaries of introspection. But the limitations of introspection do not imply that an interpretive stalemate is inevitable. There are some well-studied examples of belief that systematically diverge from the truth in ways which put pressure on the TTV characterization; especially worthy of note are the “positive illusions.” These have been variously described

and classified but, according to one of the better known sets of studies, they include self-aggrandizing perceptions, illusions of control, and unrealistic optimism (Taylor & Brown, 1988, 1994).<sup>11</sup>

Shelley E. Taylor and Johathon D. Brown have amassed considerable evidence to suggest that people consistently see themselves in a more positive light; others, in a negative light, relative to self. In commenting on this, the “better-than-most” effect, they observe that it is difficult if not impossible for any one to be warranted in believing that he is, for example, kinder, warmer, more humorous and more sincere than the average person. As regards illusions of control, the claim is not that people believe themselves capable of exercising control over that which clearly exceeds their reach; rather, this is a moderate distortion concerning those things over which people are in fact able to exert some control. And, there is a voluminous body of literature testifying to the claim that most people are unrealistically optimistic in believing their future will be better than can be justified on statistical grounds.

In effect, people tend to believe in a self-image that reassures. People consistently overestimate their abilities, whether in matters of leadership, getting along with others, or even just driving skills. These tendencies are not merely widespread among the poorly educated; as many as 94% of university professors assessed themselves as better at their jobs than their “average” colleagues (Cross, 1977). Moreover, most people, even when provided with accurate, relevant base rate information, tend to underestimate the likelihood that they will be stricken with cancer, be in a car accident, get divorced, and so forth.

---

<sup>11</sup> For a critical assessment, see Colvin and Block (1994). Some (Heine, 2001: 897-900) have questioned whether positive illusions are universal. The preponderance of evidence (Acker & Duck, 2008; Church et al., 2006), after allowing for some conceptual refinements and methodological tinkering, indicates that they are.

Pronin (2008, 2007; Pronin & Kugler, 2007; Pronin, Lin, & Ross, 2007) has devoted special attention to this final point. She discovered that when subjects are informed about “introspective illusions” and “bias blind spots” they, nevertheless, adjudge themselves to be less susceptible than others. Even when subjects—immediately after acting in accord with a particular bias—are presented with an explicit description of the bias, a description that indicates it is a common human tendency, they still fail to see themselves as liable. And these results are not indications of reticence, for instructions given to subjects make it clear that experimenters want to know whether bias is present and make it clear that the bias is common (Pronin, Lin, & Ross, 2002: 375).

Emily Pronin and Matthew B. Kugler (2007) have found that the only way to prompt subjects to recognize personal vulnerability is to specifically educate them concerning the epistemic failings of introspection. It remains unknown though whether such focused education can bring about efforts to compensate for bias (Pronin, 2007: 40). At the very least compensation would be difficult: recent evidence shows that the way we think about self in the present differs substantially from the ways in which we think about past or future selves. The limbic system and, consequently, affect, is much more engaged when people think about themselves in the present (Pronin, 2008: 1179-1180). This suggests that anticipations of the future or post hoc interpretations might be correctable in ways that judgments about the self-at-this-moment are not.

Does this evidence unequivocally demonstrate that Velleman is wrong in claiming that only unacknowledged non-evidential considerations can affect belief regulation? Not necessarily. Since the strategy here is not to cherry-pick results, it must be admitted that some evidence suggests, at least for individual events, after carefully being instructed concerning the frailty of introspectively based knowledge, subjects are capable of discerning a gap between a belief and the truth. The original belief might even be classified—albeit in retrospect—as the result of an illusion. But

what the evidence also shows is that treating t-belief as incidental leads us to overlook just how deeply ingrained is the tendency to aim away from the truth. If we are compelled to confront our epistemic frailty, in narrowly defined contexts, and just for the nonce, we might be able to respond in accord with TTV. But what TTV omits is the difficulty of accomplishing such a belief revision, the transience of such a revision, and an understanding of why TTV-effects are both difficult and transient. In a word, aiming at the truth can be a very unnatural act.

That TTV cognitions do not come naturally might be the result of their being detrimental in several aspects of our lives. Positive illusions can lead to higher motivation, greater persistence, and increased likelihood of success (Armor & Taylor, 2003; Taylor & Brown, 1988, 1994; Taylor & Gollwitzer, 1995)—all characteristics that can contribute to the cultivation of self-respect. Athletes, dancers, and soldiers with conviction are more likely to succeed than are those who lack conviction—albeit not nearly so likely as they believe. Positive illusions can also promote use of efficient and rapid problem-solving strategies. There is even evidence to suggest that positive illusions as regards one's children or one's partner are critical to successful parenting and to long-term relationships (Barelds-Dijkstra & Barelds, 2008; Wenger & Fowers, 2008).

Lionel Tiger (1999: 617) has argued along similar lines that “moderate” optimism is essential to overcoming our cognitive ability “to generate endlessly discouraging predictions of the pitfalls of any action.” He argues (1999: 615) that we are endowed with a “cognitive override . . . a moderate design defect of pure reason,” something that overrides “cognitive literalness,” that “biases the odds in favor of action” (1999: 619). Among many other supporting observations, he records that recent examination of the dentition of pre-hominids reveals that 3.5 million years ago our East African savannah ancestors were eating large amounts of meat when prey animals were hard to catch. Concerning this point he observes that those who woke up thinking “What a great day to

catch an ungulate' would enjoy an advantage over fellow citizens who turned off the alarm and rolled over to sleep straight through the prey's spurt or morning activity" (1999: 616, also see 1985).

But more is involved than just enhanced performance. Positive illusions can be adaptive for psychological health and well-being. Some evidence suggests (Alloy, 1995; Alloy & Abramson, 1988; Alloy & Ahrens, 1987; Taylor & Brown, 1988, 1994) that there is a group of people who accept both the good and bad about themselves: they remember both good and bad self-relevant information with equal frequency; their evaluations of self and others are congruent; their self-appraisals more nearly coincide with appraisals produced by impartial observers, and so forth. The group of people in question are those "who are low in self-esteem, moderately depressed, or both."<sup>12</sup> It is sometimes said of these people that their beliefs bespeak a "depressive realism" (Alloy, 1995; Alloy & Ahrens, 1987). When well-adjusted people process self-relevant information, they tend to be biased and partial; those who are dysphoric tend to be unbiased and balanced. Perhaps the single most distinctive finding in this regard is that depressed subjects, dramatically unlike those who are not depressed, "are consistently accurate judges of their control over events" (Alloy & Abramson, 2007: 242).

The claim is not that positive illusions are a necessary condition for mental health; rather, it is that these illusions can promote mental health (Taylor & Brown, 1994: 25). But that is not all. Positive illusions also seem to be protective of physical health (Taylor, Kemeny, Reed, Bower, & Gruenewald, 2000). For example, studies of AIDS patients reveal that those who believe they can control the disease and prevent its recurrence, those who do *not* "realistically" accept or appraise their condition, both exhibit a longer asymptomatic period and live longer, by an

---

<sup>12</sup> What seems to be true of the moderately depressed is not necessarily true of the severely depressed. As regards the latter, findings are equivocal (Alloy & Abramson, 2007; McKendree-Smith & Scogin, 2000).



average of nine months.<sup>13</sup> Studies of breast cancer and of AIDS patients also show that even the eventual disconfirmation of erroneous beliefs does not have harmful consequences. Moreover, what is true of the sick is also true of the healthy (Taylor, Lerner, Sherman, Sage, & McDowell, 2003): those with positive illusions, while undergoing stressful tests in a laboratory setting, exhibit lower cardiovascular responses, quicker recovery, and lower baseline cortisol levels.

Strategically aiming away from the truth contributes to enhanced performance, a sense of well-being, and better physical health. Significantly, the findings concerning physical and mental health are further confirmed insofar as they dovetail with studies of placebo and nocebo effectiveness. These carefully studied beliefs, when coupled with the studies of positive illusion, are redolent of the example sketched above, in a way that suggests an explanation for the durability of superstition.

A placebo effect is that which follows from the administration of a pharmacologically inert substance or physiologically inactive treatment<sup>14</sup> that is coupled with the verbal suggestion of clinical benefit. Nocebo effects also follow upon administration of an inert treatment, differing from placebos in that they are accompanied by suggestion of clinical harm (Benedetti, 2008; Diederich & Goetz, 2008; Oken, 2008; Zubieta & Stohler, 2009).<sup>15</sup> Nocebos, in that

---

<sup>13</sup> One physiological factor that seems to contribute to the non-realists more robust health is their ability to maintain a higher level of CD4 T helper cells.

<sup>14</sup> A placebo can be any clinical intervention, whether words, gestures, pills, various devices, or surgery. There are even hierarchies of effectiveness: e.g. injections are more effective than pills, and incisions more effective than injections (Evans, 2004).

<sup>15</sup> Verbal suggestion is not essential; sensory stimuli in an evocative setting (e.g. the sight of a syringe while sitting in a clinic) can be sufficient to elicit the effect. But when considering the effects of placebos, one must be careful to factor out other causes, e.g. spontaneous remission, regression to the mean, and patient biases. Moreover, effectiveness can vary. This variation might be explainable in terms of functional differences in the mesolimbic dopaminergic pathway (Scott, Stohler, Egnatuk, Wang, Koeppe, & Zubieta, 2007).

their effects are adverse, bear more direct resemblance to the alleged consequences of ignoring superstitions.<sup>16</sup> What matters though is that the nocebo or placebo, despite being inert, by virtue of engaging a person's belief—in a manner that aims away from the truth—is able to bring about a measurable physiological outcome, salubrious or noxious (Benedetti, 2008: 36, 48).

Placebos have been demonstrated to have salubrious effects in the treatment of many conditions, e.g. pain, swelling, addiction, cardiovascular and respiratory problems, peptic ulcers, depression, anxiety, cancer, and Parkinson's disease (Benedetti, 2008; Evans, 2004). Some of the mechanisms<sup>17</sup> whereby belief is able to effect these changes include: the release of endogenous opioids or dopamine, the inhibition of serotonin uptake, the reduction of  $\beta$ -adrenergic heart activity, as well as the conditioning of immune receptors like lymphocytes and hormones like cortisol. Further, differentiation among types of placebo effectiveness are being teased apart: for example, conscious expectation seems to play a greater role in alleviating pain and enhancing motor performance, whereas classical conditioning can be sufficient to trigger immune and hormonal responses (Benedetti, 2008: 42; Nieme, 2009).

Once again recall that, according to Velleman, when we discern a gap between a belief and the truth, the belief becomes unsettled and starts to change. Studies of placebos, however, reveal a dissociation between different forms of belief regulation: one results from conscious expectation, the other, from classical conditioning. A natural explanation of superstition susceptibility now suggests itself. A person who, sincerely and for good reason, denies holding the belief that 13 is unlucky, might, due to analogues of classical conditioning that occur in everyday life,<sup>18</sup>

---

<sup>16</sup> Perhaps the most famous documented example of a nocebo effect is "voodoo death" (Lex, 1977).

<sup>17</sup> PET technology (Mayberg, Sliva, Brannan, Tekell, Mahurin, McGinnis, et al., 2002) has made it possible to begin teasing apart the functional neuroanatomy, even distinguishing it from the effects of pharmacologically active treatments.

<sup>18</sup> See, for example, Brunstrom (2007) and Stockhorst, Enck, and Klosterhalfen

come to behave in such a way that can best be explained by the belief that 13 is unlucky. Even if acting in that way contravenes professed beliefs, it is not obvious these non-evidential considerations (those regulated by classical conditioning) can only be influential if unacknowledged. Although this might strike some as absurd, perhaps the apparent absurdity merely reflects a design compromise that has been achieved during our evolutionary development.<sup>19</sup>

Placebo beliefs are like positive illusions in that both are false. But placebos are false in a distinctive way. To illustrate this point, compare placebo effectiveness with positive illusions that cause people to be overconfident in the extent and effectiveness of their control. What they are right about is in believing that they exercise some control; they are wrong, however, in their assessment of how much control they have. A person who asserts that he has the ability to hit a 450 foot home run might be wrong by a degree that is easily calculable. But a person who asserts that by drinking a particular potion (perhaps a mix of tap water, sugar, and food coloring) his peptic ulcer will be cured is completely wrong. There is nothing in the potion that will contribute to his cure; it is pharmacologically inert. Nevertheless, effective brain mechanisms can in this way be set in motion.

Today placebo effects are often triggered by the presence of doctors, medications, needles, even just the smell of a clinic, all things that are highly correlated with effective treatment. But this could not have been the case within which placebo mechanisms evolved. And, for both modern and antediluvian placebos, we know for a fact that any correlations which might obtain are non-causal.

Consider the candidate mechanisms cited above: the release

---

(2007).  
<sup>19</sup> McKay and Dennett (2009) side with Humphrey (2002) in treating placebo “misbeliefs” as by-product, not adaptation. My speculations in this regard differ, but nothing critical to the central thesis turns on this difference.

of opioids or dopamine, the conditioning of immune receptors, the inhibition of serotonin uptake, or the reduction of  $\beta$ -adrenergic heart activity. Presumably they are activated by means of some form of mind-body “lingua franca” (Humphrey, 2004; see also Beauregard, 2007: 233), the psychological side necessarily involving false beliefs. Also note that modern medicine is scarcely more than a century old:<sup>20</sup> we are not long past the days of blood-letting and ignorance of microscopic organisms. The mind-body lingua franca apparently evolved in an environment under which the input constraints definitive of (those) beliefs could not have been yielding true beliefs. After all, lacking alternatives, systematic examination of shamanic beliefs, rituals, and incantations, including careful consideration of instances of failure, would hardly have been worth the effort. And morbid acceptance of death and disease was likely no more helpful to individual or group esprit during the Pleistocene than it is today. So it is for good reason that the placebo effect is, as Bakan (1985: 212-213) has written, demonstrated “precisely in cases in which expectancy is falsely grounded.”

To say that these beliefs aim away from the truth is not to say that they also aim away from “instrumental success.” But recall that Velleman insists such goals are “second-order.” The contention here, by contrast, is that belief regulatory systems evolved at a time when aiming for the truth, in some areas of life, would have been as pointless as counseling Aristotle to devote more attention to the brain—millennia before the development of neuroscience. In this area of life, for good reason, instrumental success supersedes truth-conduciveness.

Reflection on the modern world suggests that communities maintain repositories of false beliefs and humans retain wells of

---

<sup>20</sup> Although if “medicine” is defined as “the provision of special care to the sick by others,” it might be very old (Evans, 2002: 459). But to establish my point it is only necessary that most medical claims concerning the cause of cures were groundless.

gullibility that can be drawn upon during times of social turmoil or personal crisis. To cite just one example, Pascal Boyer (2000: 99-100, 105-106) has found that when dealing with religion we are naturally inclined to be gullible as concerns the “odd” or the “unfamiliar.” Paradigmatic of these are spirits who are represented as intentional agents, but agents whose physical properties violate the physical qualities of embodied agents. Not only are these violations not taken as evidence that the entities aren’t real, instead, “it is precisely insofar as a certain situation violates intuitive principles and is taken as real that it may become particularly salient” (Boyer, 2000: 101). In effect it appears that appropriately structured systems of false belief remain available within society, accessible to all, and ready when needed.<sup>21</sup>

A third set of beliefs that aim from the truth are referred to as self-deception (Mele, 1997: 92, 2001: 4). These too are extremely common: paradigmatic examples include people who believe in their spouse’s fidelity, their likelihood of recovery from illness, or their child’s avoidance of illicit drugs, despite the availability of evidence so compelling that were it about the spouse, the illness, or the child of someone else, their confidence would surely be shaken. Alfred R. Mele (1997: 93-94) has noted that these self-deceptions prime other cognitive mechanisms, such that they then contribute to the production of yet more false beliefs, mechanisms that include information salience, the availability heuristic, confirmation bias, and our tendency to search for causal

---

<sup>21</sup> Some might contend that the apparent diminishing influence of institutional churches suggests diminished gullibility to beliefs that are not regulated in truth-conducive ways. Perhaps it is true that institutions of this sort are diminishing in influence—in some parts of the world—but that fact doesn’t imply a diminishing influence of such beliefs. The case of France might be instructive in this regard: it has experienced an ever dwindling supply of Roman Catholic clergy. According to tax authorities the number is now down to 36,000. But, according to those same tax authorities, France now has 40,000 professional astrologers (Kahane & Cavender, 2002: 137). If I am correct, a Conservation of Credulity principle seems to play an important role in human society.

explanations. If Mele is correct, then one false belief can lead to a concatenation of further false beliefs. According to Robert Trivers (2000: 125) some self-deceptions function as do positive illusions. But he (1985, 2000) also suggests that self-deception has a unique role: it was favored by natural selection because it enhances our ability to deceive others. The idea is that if we first deceive self (e.g. a politician who says to his constituents, “I feel your pain”), then the autonomic nervous system changes that might indicate falsehood to others would not be manifest. Creatures capable of self-deception could then reap certain Machiavellian rewards.<sup>22</sup>

There is abundant evidence that sometimes beliefs are regulated so to aim away from the truth. What’s more, the mechanisms engaged in production of false belief are difficult to override. Consider, for example, that rejection of a superstition might have been regulated in truth-conducive ways. And because the rejection is so thorough, it should not be subject to truth-conducive revision or extinction. But, prior to rejection, if one has been classically conditioned, in that the superstition was learned early in life and under the proper conditions, despite having later been extinguished, self-control will still be required to resist its influence. According to the ego-depletion hypothesis, self-control is like muscle strength (Muraven & Baumeister, 2000). It is a limited resource that can be exhausted by excessive demands and, once depleted (e.g. as indicated by low levels of blood glucose), recovery is slow (Galliot et al., 2007). Accordingly, we should not be surprised to discover that when the ego is depleted,

---

<sup>22</sup> In both the empirical and the philosophical literature, there is a general consensus that “self-deception” names an important phenomenon. But most discussion of this phenomenon is contentious. These controversies are fueled by the lack of convincing laboratory demonstrations (Paulhus, 2007) and by worries pertaining to specific interpretations, e.g. self-deception does not always help with the deception of others in the way that Trivers and other evolutionary psychologists have claimed (Van Leeuwen, 2007: 335). For purposes of this paper it is sufficient that self-deception illustrates both the generation of false beliefs and that it suggests one aspect of T-belief’s capacity for causally affecting psychological and social circumstance.

as can happen when one is under great stress, people might behave in ways that contravene professed beliefs. And if one is trained to recognize the indicators of depletion, it would not be surprising to find that they are capable of acknowledging the role non-evidential considerations— e.g. low-levels of blood glucose—play in determining what is believed.

When considering whether beliefs might aim at something other than the truth, recall that Velleman invokes a basketball analogy. Regardless of whether this analogy can be cogently applied to beliefs of any kind, it certainly does not fit here. By that analogy, within the game, one plays to win (the first-order aim), because doing so will ensure salary increase (the second-order aim, the ultimate aim). To take the case of placebos as an example, clearly their ultimate—the second-order—aim is to be restored to good health. But to realize that ultimate aim, one cannot adopt a first-order goal of aiming at the truth; to do so would be self-defeating. On the court—in the game of life—it is not the case that these beliefs are truth-directed. In these contexts, truth-directedness and instrumental success are at odds with one another.

## V. The Distinctive Character of Tertullian Beliefs

What is most distinctive about the cluster of beliefs described above is that they aim *away* from the truth. I have dubbed them Tertullian beliefs, or t-beliefs.<sup>23</sup> Tertullian seems a proper eponym because it is (apocryphally) said that he proclaimed: “I believe because it is absurd.”<sup>24</sup> What he (Tertullian, 2010) actually wrote

---

<sup>23</sup> Reflection on some of the empirical evidence presented here has prompted others to wonder whether the relevant attitude is hope, not belief (Flanagan, 2002: 22, fn. 4). But hope seems ambiguous between belief and desire, not an amalgam (as will be described here). Hope in the sense of “hopeful” seems to be nothing more than belief in a better future (Breznitz, 1999: 629); and, most uses of “I hope” seem synonymous with “I desire” or “I want.”

<sup>24</sup> Tertullian’s words are a common textbook example of irrationalism (Quine &

was, “certum est, quia impossibile:” that is, “it is certain, because it is impossible.” The idea of believing because it is “absurd” or “impossible,” though a bit hyperbolic, evinces the deliberateness of aiming away. What matters is not that the beliefs are false. What matters is that they seem calibrated to be false, in a certain way, and to a certain degree.<sup>25</sup> The effectiveness of a positive illusion, a placebo, or self-deception depends upon aiming, in just the right way, a calibration which seems achievable only as the result of design.

As a first approximation the metaphor “direction-of-fit” can be employed to capture the distinction between t-beliefs and TTV beliefs (Searle, 2001: 37-38, 257). TTV beliefs exhibit a mind-to-world direction of fit; t-beliefs, on the other hand, exhibit a world-to-mind direction of fit. Mind-to-world direction-of-fit implies that it is the purpose of TTV beliefs to change so that they match the world. Ordinarily world-to-mind direction of fit—representation not of how things are, but of how we would like things to be—is used to characterize the attitude “desire.” Of desire it can be said that its purpose is to change the world to match its content.

T-beliefs, like desire, aim to change the world.<sup>26</sup> In this respect, they are unlike TTV beliefs. Nevertheless, they are asserted in such a way as to imply that they represent how things are—13 is unlucky, I am better-than-average, the waters of Lourdes have curative powers, or I am sincere. T-beliefs are not regarded as true in a fanciful, polemical, or heuristic way. The spirit in which the propositional object is regarded as true is serious.

---

Ullian, 1978: 60).

<sup>25</sup> A balance must be maintained: too much positive illusion, and one will be disinclined to seek available help. Too little, and one might despair. Daniel Gilbert (2005: 177-178) refers to this as a “psychological immune system” that must, like its physiological counterpart, maintain a balance between hypo and hyper activity.

<sup>26</sup> Other amalgams of belief and desire have been proposed, e.g. “besire” (Blackburn, 1998: 97-100).



T-beliefs require a blurring of the usual belief-desire distinction. Direction-of-fit can help elucidate this relationship, but it remains just a metaphor (cf. Sobel & Copp, 2001). Fortunately direction-of-fit can be further explicated in terms of causal connectedness. Consider again self-confidence, self-healing, and self-deception. The beliefs associated with these phenomena are generated by mechanisms that aim away from the truth. Not only that, like desire they conspire to change the world to match their content. When they succeed, it is not by accident. Rather their success results from their being about the same part of the world (the body) that they inhabit, either intra-cranially or interpersonally. Positive illusions and placebos can be effective intra-cranially via the appropriate mind-body lingua franca, and self-deceptions can be effective by shutting down autonomic reactions that would otherwise be detectable to those one wants to persuade. This corner of the world—the intra-cranial and the interpersonal—is, so to speak, within striking range of belief. There is a causal link between these beliefs and that portion of the world that they target.<sup>27</sup>

Note that t-belief is not reducible to desire. People who merely “want” to be better-than-average, to recover good health, or to be interpersonally successful don’t succeed in the way that those who t-believe do. A clear distinction between mere wanting and believing remains: the moderately depressed want to perform at a higher level they just don’t believe that they will. And most who fall ill want to recover. But only t-belief improves chances of recovery by means of the intra-cranial causal nexus. Simple desire, mere wanting, doesn’t cut it.

Another way to approach this distinctive blend of belief and desire is to note that it can help to diminish the puzzlement of a

---

<sup>27</sup> Tamar Szabó Gendler (2008) has recently introduced the concept “alief.” Like t-belief alief is claimed to “govern all sorts of belief-discordant behavior” (2008: 663). But t-belief differs in several respects, including its ability to change the world to match its content.

philosophical curio, Moore's Paradox. Consider that denouncing a superstition but being influenced to act in accord with that superstition seems rather like an instance of "p but I don't believe that p." In other words, it is suggestive of what has come to be called Moore's Paradox, which is just a paradox in the informal sense for "p" and "I don't believe that p" might both be true. Nevertheless, since asserting "p" seems to imply the belief that p, this is typically regarded as an utterance of a type that I cannot sensibly assert of myself.<sup>28</sup> In the superstition case we have apparently contradictory expressions, both the assertion "I don't believe that p" and behavior which seems best explainable by attributing the belief "p." Typically it is claimed that one could not self-ascribe both.<sup>29</sup> But when one is aware of what is implied by one's behavior, such self-ascription is possible.

Jeanette Kennett and Cordelia Fine (2008: 176-177) have found that psychopaths and sociopathic delinquents produce many statements that are "Moorean paradoxical" (cf. Joyce, 2007: 51-57). As a typical example, consider: "John is an honest person. Of course, he has been involved in some shady deals!" As with Moorean paradoxes generally, when treated as a whole, the statement seems to make no sense. Kennett and Fine treat this paradox as a measure by which to determine whether the psychopaths or sociopaths grasp what is implied by evaluative terms.

What I am suggesting is that healthy people are capable of Moorean paradoxical expressions in that their professions of belief are contravened by behaviors whose implications are recognizable to the subject. This tension between what one professes and how one behaves reflects a design compromise, one which for most

---

<sup>28</sup> Moore's concern was to illustrate the distinction between what is asserted and what is implied; Wittgenstein bestowed the name "Moore's paradox" (Baldwin, 1990: 226).

<sup>29</sup> Note that the same could not be said of unalloyed desire. "I want it to be the case that p" and "not-p" are not inconsistent (cf. Crane, 2001: 102- 105).

people is salubrious. If this view is correct, we can reasonably expect that those who are mildly depressed, should be more sensitive to the implications of Moore's paradox; therefore, they would be less likely to behave in ways that contradict their professed beliefs.<sup>30</sup>

One might wonder why this aspect of belief has yet to be duly recognized. After all the relevant scientific studies can now be traced back to well over two decades. Perhaps it is that one of the institutions which consistently gives these beliefs pride of place, religion, is not taken seriously.<sup>31</sup> Perhaps as well we live in a world with so many dangerous false beliefs that we fail to appreciate non-TTV forms of belief regulation (Bennett & Hacker, 2003: 172-174). A further factor that causes neglect of t-belief might derive from an under-appreciation of what Wallace Arthur (2004) calls "internal adaptations."

Arthur (2004: 117-127) points out that "ecological" adaptations, adaptations to the (external) physical environment, receive most attention in biology; internal adaptations or "coadaptations," adaptations among body parts, tend to be neglected. An example of the former is the adaptation of forest flies to higher ambient temperatures (Arthur, 2004: 122-123): flies must struggle to stave off desiccation. The hotter it gets, the faster they lose water. Because the larger one is, the smaller one's surface area is relative to volume, and because water loss occurs at the body's surface, in a dry, hot environment being bigger is better. In accord with selective pressure, the average body size of the fly population will increase. Because the fitness difference is clearly produced by the external environment, this counts as an external adaptation.

---

<sup>30</sup> This is a testable implication of the t-belief hypothesis.

<sup>31</sup> Richard Dawkins (1993) stakes out an especially uncompromising position: he regards religious beliefs as either marks of cowardice or as "pernicious," symptoms of disease such that those who hold them should be regarded as "patients."

But suppose that along with the difference in body size, these flies also differ in the way their wings are connected to their thorax. Suppose as well that this variation slightly affects their ability to fly. Under such circumstances, the population will evolve toward better integrated joints. Here though selection is unrelated to the forest's change in ambient temperature; it isn't even related to the forest. Although flight occurs in environments, good flying ability is generally advantageous for flies, no matter what environment they inhabit. Accordingly, these fitness differences are "quasi-environment-independent." Internal selection, in an important sense, "travels with the organism wherever it goes."<sup>32</sup>

Relating this distinction to t-belief, we might say that most philosophical attention to belief has concerned "ecological adaptations." Understandable though this might be, the "internal" environment is also part of that to which we must adapt.<sup>33</sup> And because t-belief is so critical to internal adaptations, it warrants more attention than it receives from within the TTV conceptual framework.

W. V. O. Quine (1994: 66) famously wrote: "Creatures inveterately wrong in their inductions have a pathetic but praise-worthy tendency to die before reproducing their kind." This clever turn-of-phrase strikes many as necessarily true. But it misleads. Sometimes we are wrong for good reason. And, to grasp what counts as a good reason, we should attend to internal adaptations.<sup>34</sup> A balance must be struck between the external and

---

<sup>32</sup> "External" and "internal" are best understood as occupying opposite ends of a continuous spectrum.

<sup>33</sup> A relevant example of this is the "tragedy of cognition" (Atran, 2003): we can meta-represent self and others, project the future, and envision the demise of all we care about. These too are part of the environment that we must adapt to.

<sup>34</sup> I do claim that t-beliefs can be adaptive, in that they enhance fitness. Whether or not they count as biological adaptations (i.e. whether or not we have inherited them because they enhanced the fitness of our ancestors) is not something that needs to be dealt with here (cf. Buller, 2005: 35). But because false beliefs can seem so non-functional (cf. Konner, 2002: 15), and because maintaining a proper balance and calibration seems very complex (cf. Buller, 2005: 31-37), I

the internal, between TTV and t-belief. Accuracy of inductions concerning the external world is not enough. Neglect of internal adaptations can also lead to pathetic but praise-worthy ends.

## VI. Wood's Procedural Principle and T-Beliefs

Recall that Wood (2008: 13, fn. 8) presupposes that no one can stably hold a belief they know to be false. But there appear to be counter-examples to this claim. In the anecdotal cases, people behave in accord with superstitions and purchase vanity products or resort to miracle cures, even when they can fairly be said to know that recourse to these strategies or products is grounded in false beliefs.

As regards empirical studies of t-beliefs, even when subjects are presented with explicit description of cognitive biases immediately after acting in accord with those biases, they exhibit no evidence of belief instability. If the notion of “instability” is unpacked in the way proposed by Velleman—i.e. gaps between belief and truth are reconciled by either adding new beliefs or changing those originally held—it seems that at most the instability of believing is narrowly circumscribed. Only if compelled to confront evidence in constrained experimental settings might one evince the predicted adjustments. And even these meager findings might not be ecologically valid. In sum, there is no evidence, independent of claims based upon contentious introspective reports, that belief instability is a natural disposition.

As regards the ethics of belief, Wood (2002: 38-40) acknowledges the possibility of exceptions to the procedural principle. But he regards criticisms of it that are based upon this possibility as “cheap” and “wrongheaded.” Even should a person determine that the principle need be violated, Wood counsels that

---

suspect biological evolution may have played some role. Nevertheless, since no precision can as yet be given to the claims of usefulness here, it is better to allow that t-belief mechanisms are labile with the environment.

the person should “feel squeamish and conflicted.”

Wood’s (2002: 33) harsh judgment in this regard is motivated by his belief that violations of the procedural principle are “shameful in something of the same way that telling lies is shameful.” When we believe that which is “comfortable to believe,” we show contempt for self and perform a disservice to others. That is, we fail to respect ourselves as rational beings and we deprive others of honest evaluations that they might need.

Doubtless Wood (2008: 13) is correct that one need not look long or far to find innumerable examples of “shameless evasions.” But if we apply a principle of psychological realism to our moral theories,<sup>35</sup> then there are reasons to be dubious of the procedural principle. First, there is good reason to believe that gaps between belief and truth do not necessarily precipitate instability. We seem to be designed in such a way as to allow for these inconsistencies, without the untoward spill-over effects that concern Wood. A stable compromise has been forged between internal and external adaptations. The more closely one examines instability claims, the more they seem to be artifact derived from unwarranted philosophical expectations of consistency.

Second, if we were to feel squeamish and conflicted each time we acted in accord with a positive illusion or with the distribution of a placebo, the benefits of illusions and placebos would not be attainable. As regards self, for the positive illusions to contribute to our well-being (and not, say, exacerbate depressive realism), we should not feel squeamish or conflicted. As regards our treatment of others, for the placebo to be effective, likewise, we should not feel squeamish or conflicted, as these would be evident to the patient.<sup>36</sup> And if the benefits of t-belief were not forthcoming, the

---

<sup>35</sup> According to Owen Flanagan’s (1991: 32) Principle of Minimal Psychological Realism, our moral theories should not require of us that which is not possible for creatures like us (also see Doris, 2002: 112).

<sup>36</sup> Note that placebo induction seems to be the only deliberate induction of a false belief that Wood (2008: 14) finds acceptable, albeit grudgingly: “To lie paternalistically to people may sometimes help them (for instance, to overcome a

results would include diminished health, performance, motivation, and well-being.

In addition to Wood's skepticism that there are legitimate exceptions to the procedural principle,<sup>37</sup> he believes that violations are essentially "corrupting" (2002: 36). He believes that people are inclined to take unacceptable liberties, allowing for both unrestrained rationalizations of personal behavior and dishonesty in public discourse. But t-beliefs seem to be legitimate exceptions, and there is no empirical evidence that the tendency to act in accord with t-beliefs leads to the corrupting tendencies that are the object of Wood's concern. It might be the case that Wood is correct in his assessment of other beliefs. But it is not difficult to conceive of people who hold positive illusions, self deceive in the standard circumstances, and react to placebos in ways that enable them to be effective, while not allowing for these breaches of the procedural principle to adversely affect other aspects of their lives or of public discourse. T-beliefs seem designed so as to be insulated from the rest of our beliefs.

What Wood's advocacy of an uncompromising adherence to the procedural principle requires is evidence of a particular sort. For example, if it turns out to be the case that depressive realists are less inclined to corruption and more respectful of self than are the majority of people, then Wood's views could be said to be rightly affirmed. But there is no evidence of this sort; none, whatsoever.

It might be said that Wood's view reflects a strictly normative position, and that empirical evidence, be it anecdotal or scientific, is of no relevance. But, just as a matter of fact, Wood justifies his uncompromising position by making specific empirical claims

---

life-threatening illness), but . . . it shows a lack of respect . . . and seems justifiable only temporarily, under very special conditions."

<sup>37</sup> "There are no matters about which we do not owe it both to ourselves and to others to maintain our intellectual integrity by forming our beliefs according to the evidence" (Wood, 2002: 33).

about the nature of belief as well as about the tendencies of people who fail to act in accord with the procedural principle. To show that these empirical assumptions are dubious as regards t-belief then is to weaken support for Wood's version of this principle.

Furthermore, the evidence suggests that, contrary to what Wood maintains, t-beliefs might be critical to—rather than detrimental to—the maintenance of self-respect. What is neglected is the compromise between internal and external adaptations, as well as the causal role that t-beliefs can play. TTV tends to treat t-beliefs as incidental; Wood takes this view a step further and treats them as “pernicious.” But when properly calibrated, they can help alleviate depression, improve health, enhance motivation, and improve performance. Depression, ill health, indolence, and failure are not contributors to self-respect. They are obstacles. What both anecdotal and empirical evidence suggest is that an appropriate dose of the right kind of false beliefs might be a necessary condition for the development of self-respect. Sometimes it pays to be Moorean paradoxical.

Recall Tiger's speculation concerning the lot of a pre-hominid who lacked the capacity for t-believing. It is simply too easy “to generate endlessly discouraging predictions of the pitfalls of any action” (Tiger, 1999: 617). We seem to need an antidote to “cognitive literalness,” something that moves us to action. T-beliefs, in right measure, just are that antidote. Without them we are less inclined to taking action in a whole host of ways that are essential for self-respect.

Wood (2008: 19) emphasizes that self-respect requires the apportioning of belief strictly in accord with the evidence. But those who best adhere to this requirement as regards beliefs about self tend to have low self-esteem (Alloy, 1995; Alloy & Ahrens, 1987). The claim advanced here is that a certain measure of self-esteem is a precondition for self-respect. Those who are without t-beliefs seem to lack the minimum esprit necessary for the maintenance of that which Wood values so highly.



## VII. Conclusion

Wood (2002: 8, 2008: 9) emphasizes that we are responsible for the processes of belief formation and maintenance. Just as we would be blameworthy for killing someone in a drunken rage, so too we are blameworthy for acting in accord with beliefs that are not properly formed or maintained. In the former case, we should have known not to get drunk. In the same way, we behave irresponsibly when we allow cognitive biases to lead away from the truth. We are obliged to be proactive.

Might Wood have a point here? Perhaps the way things are with beliefs is blinding us from the way things could be. Perhaps we, individually and collectively, need to be weaned from t-beliefs. And perhaps this would be a good thing. But the formal investigations of Pronin (2007) suggest that weaning is not an option.

Less formally, it seems to be the case that when progress is made toward reducing the effects of cognitive biasing on one front, those biases reemerge on another. Above I noted that, in France, as the number of Roman Catholic clergy decrease in numbers, the number of professional astrologers increase. A Conservation of Credulity Principle seems to be in effect.

Whether or not individuals or societies can be weaned from t-belief in such a way as to manage proper alignment with the procedural principle is an empirical issue. Whether or not we should be weaned is an ethical issue. For Wood, it would seem, the two become relevant to one another when we assess the cost of attempts at weaning. If success brings about enhanced self-respect and no collateral, corrupting effects, then it is a good. If it brings about diminished self-respect and an increase in corrupting effects, then it is not. If the latter, then even on Wood's terms, the procedural principle should be compromised.

## References

- Acker, D., & Duck, N. (2008). Cross-cultural confidence and biased self-attribution. *The Journal of Social-Economics*, 37, 5: 1815-1824.
- Alloy, L. B. (1995). Depressive realism: Sadder but wiser? *The Harvard Mental Health Letter*, 11, 10: 4-5.
- Alloy, L. B., & Abramson, L. Y. (1988). Depressive realism: Four theoretical perspectives. In L. B. Alloy (Ed.), *Cognitive processes in depression* (pp. 223-265). New York: The Guilford Press.
- Alloy, L. B., & Abramson, L. Y. (2007). Depressive realism. In R. F. Baumeister & K. D. Vohs (Eds.), *Encyclopedia of social psychology* (pp. 242-243). Los Angeles: Sage.
- Alloy, L. B., & Ahrens, A. (1987). Depression and pessimism for the future: Biased use of statistically relevant information in predictions for self versus others. *Journal of Personality and Social Psychology*, 52, 2: 366-378.
- Armor, D., & Taylor, S. (2003). The effects of mindset on behavior: Self-regulation in deliberative and implemental frames of mind. *Personality and Social Psychology Bulletin*, 29, 1: 86-95.
- Arthur, W. (2004). *Biased embryos and evolution*. Cambridge, UK: Cambridge University Press.
- Atran, S. (2003). The neuropsychology of religion. In R. Joseph (Ed.), *Neurotheology: Brain, science, spirituality, and religious experience* (pp. 147-166). Berkeley, CA: University Press.
- Bakan, D. (1985). The apprehension of the placebo phenomenon. In L. White, B. Tursky, & G. E. Schwartz (Eds.), *Placebo: Theory, research, and mechanisms* (pp. 211-214). New York: The Guilford Press.
- Baldwin, T. (1990). *G. E. Moore*. London: Routledge.
- Bausell, R. (2007). *Snake oil science: The truth about complementary and alternative medicine*. New York:

- Oxford University Press.
- Benedetti, F. (2008). Mechanisms of placebo and placebo-related effects across diseases and treatments. *The Annual Review of Pharmacology and Toxicology*, 48: 33-60.
- Bennett, M. R., & Hacker, P. M. S. (2003). *Philosophical foundations of neuroscience*. Oxford: Blackwell.
- Beauregard, M. (2007). Mind does really matter: Evidence from neuroimaging studies of emotional self-regulation, psychotherapy, and placebo effect. *Progress in Neurobiology*, 81, 4: 218-236.
- Blackburn, S. (1998). *Ruling passions*. New York: Oxford University Press.
- Boyer, P. (2000). Evolution of the modern mind and the origins of culture: Religious concepts as a limiting case. In P. Carruthers & A. Chamberlain (Eds.), *Evolution and the human mind: Modularity, language, and meta-cognition* (pp. 93-112). New York: Cambridge University Press.
- Breznitz, S. (1999). The effect of hope on pain tolerance. *Social Research*, 66, 2: 629-652.
- Brunstrom, J. (2007). Associative learning and the control of human dietary behavior. *Appetite*, 49, 1: 268-271.
- Buller, D. J. (2005). *Adapting minds: Evolutionary psychology and the persistent quest for human nature*. Cambridge, MA: The MIT Press.
- Case, T. (2000). *The need to believe: Motivation, disposition, and superstition*. Unpublished doctoral dissertation, Macquarie University, Australia.
- Case, T. I., Fitness, J., Cairns, D. R., Stevenson, R. J. (2004). Coping with uncertainty: superstitious strategies and secondary control. *Journal of Applied Social Psychology*, 34, 4: 848-871.
- Church, A. T., Katigbak, M. S., del Prado, A. M., Valdez-Medina, J. L., Miramontes, L. G., & Ortiz, F. A. (2006). A cross-cultural study of trait self-enhancement, explanatory variables, and adjustment. *Journal of Research in Personality*,

40, 6: 1169-1201.

- Colvin, C., & Block, J. (1994). Do positive illusions foster mental health? An examination of the Taylor and Brown formulation. *Psychological Bulletin*, 116, 1: 3-20.
- Crane, T. (2001). *Elements of mind: An introduction to the philosophy of mind*. New York: Oxford University Press.
- Cross, K. P. (1997). Not can, but will college teaching be improved? *New Directions for Higher Education*, 17: 1-15.
- Davidson, D. (1977). The method of truth in metaphysics. In P. A. Finch, T. E. Uehling, & H. K. Wettstein (Eds.), *Midwest studies in philosophy II: Contemporary perspectives in the philosophy of language*. Minneapolis, MN: University of Minnesota Press.
- Davidson, D. (2003). Thought and talk. In T. O'Connor & D. Robb (Eds.), *Philosophy of mind: Contemporary readings* (pp. 355-369). London: Routledge.
- Dawkins, R. (1993). Viruses of the mind. In B. Dahlbom (Ed.), *Dennett and his critics: Demystifying mind* (pp. 13-27). Oxford, MA: Blackwell.
- Diamond, J. (2001). *Snake oil and other preoccupations*. New York: Vintage.
- Diederich, N., & Goetz, C. (2008). The placebo treatments in neurosciences: New insights from clinical and neuroimaging studies. *Neurology*, 71: 677-684.
- Doris, J. (2002). *Lack of character*. Cambridge, UK: Cambridge University Press.
- Evans, D. (2002). Pain, evolution, and the placebo response. *Behavioral and Brain Sciences*, 25, 4: 459-460.
- Evans, D. (2004). *Mind over matter in modern medicine*. London: HarperCollins.
- Flanagan, O. (1991). *Varieties of moral personality: Ethics and psychological realism*. Cambridge, MA: Harvard University Press.
- Flanagan, O. (2002). *The problem of the soul: Two visions of mind and how to reconcile them*. New York: Basic Books.

- Galliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., Tice, D. M., et al. (2007). Self-control relies on glucose as a limited energy source: Willpower is more than a metaphor. *Journal of Personality and Social Psychology*, 92, 2: 325-336.
- Gazzaniga, M. (2008). *Human: The science behind what makes us unique*. New York: HarperCollins.
- Gendler, T. S. (2008). Alief and belief. *The Journal of Philosophy*, 105, 10: 634-663.
- Gilbert, D. (1993). The assent of man: Mental representations and the control of belief. In D. Wegner & J. Pennebaker (Eds.), *Handbook of mental control* (pp. 57-87). Englewood-Cliffs, NJ: Prentice-Hall.
- Gilbert, D. (2005). *Stumbling on happiness*. New York: Vintage Books.
- Heine, S. (2001). Self as cultural product: An examination of East Asian and North American selves. *Journal of Personality*, 69, 6: 881-906.
- Humphrey, N. (2002). *The mind made flesh: Essays from the frontiers of psychology and evolution*. New York: Oxford University Press.
- Humphrey, N. (2004). Placebo effect. In R. L. Gregory (Ed.), *The Oxford companion to the mind* (pp. 735-736). New York: Oxford University Press.
- Hurlburt, R., & Schwitzgebel, E. (2007). *Describing inner experience: Proponent meets skeptic*. Cambridge, MA: The MIT Press.
- Joyce, R. (2007). *The evolution of morality*. Cambridge, MA: The MIT Press.
- Kahane, H., & Cavender, N. (2002). *Logic and contemporary rhetoric* (9th ed.). Belmont, CA: Wadsworth.
- Keinan, G. (2002). The effects of stress and desire for control on superstitious behavior. *Personality and Social Psychology Bulletin*, 28, 1: 102-108.
- Kennett, J., & Fine, C. (2008). Internalism and the evidence

- from psychopaths and “acquired sociopaths.” In W. Sinnott-Armstrong (Ed.), *Moral psychology: Vol. 3. The neuroscience of morality: Emotion, brain disorders, and development* (pp. 173-190). Cambridge, MA: The MIT Press.
- Konner, M. (2002). *The tangled wing: Biological constraints on the human spirit* (2nd ed.). New York: Henry Holt.
- Lex, B. W. (1977). Voodoo death: New thoughts on an old explanation. In D. Landy (Ed.), *Culture, disease, and healing studies in medical anthropology* (pp. 327-331). New York: Macmillan.
- Lynch, M. (2004). *True to life: Why truth matters*. Cambridge, MA: The MIT Press.
- Mayberg, H. S., Silva, J. A., Brannan, S. K., TeKell, J. L., Mahurin, R. K., & McGinnis, S., et al. (2002). The functional neuroanatomy of the placebo effect. *The American Journal of Psychiatry*, 159, 5: 728-737.
- McGinn, C. (2004). *Mindsight: Image, dream, meaning*. Cambridge, MA: Harvard University Press.
- McKay, R., & Dennett, D. (2009). The evolution of misbelief. *Behavioral and Brain Sciences*, 32, 6: 493-561.
- McKendree-Smith, N., & Scogin, F. (2000). Depressive realism: Effects of depression severity and interpretation time. *Journal of Clinical Psychology*, 56, 12: 1601-1608.
- Mele, A. R. (1997). Real self-deception. *Behavioral and Brain Sciences*, 20: 91-136.
- Mele, A. R. (2001). *Self-deception unmasked*. Princeton, NJ: Princeton University Press.
- Muraven, M., & Baumeister, R. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin*, 126, 2: 247-259.
- Niemi, M. (2009). Cure in the mind. *Scientific American Mind*, 20, 1: 42-49.
- Nozick, R. (1993). *The nature of rationality*. Princeton, NJ: Princeton University Press.

- Oken, B. (2008). Placebo effects: Clinical aspects and neurology. *Brain*, 131, 11: 2812-2823.
- Paulhus, D. L. (2007). Self-deception. In R. F. Baumeister & K. D. Vohs (Eds.), *Encyclopedia of social psychology* (pp. 189-190). Los Angeles: Sage.
- Pronin, E. (2007). Perception and misperception of bias in human judgment. *TRENDS in Cognitive Science*, 11, 1: 37-43.
- Pronin, E. (2008). How we see ourselves and how we see others. *Science*, 320, 5880: 1177-1180.
- Pronin, E., & Kugler, M. (2007). Valuing thoughts, ignoring behavior: The introspection illusion as a source of the bias blind spot. *Journal of Experimental Social Psychology*, 43: 565-578.
- Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, 28, 3: 369-381.
- Quine, W. V., & Ullian, J. S. (1978). *The web of belief* (2nd ed.). New York: McGraw-Hill.
- Quine, W. V. O. (1994). Natural kinds. In H. Kornblith (Ed.), *Naturalizing epistemology* (pp. 57-76). Cambridge, MA: MIT Press.
- Railton, P. (2003). *Facts, values, and norms: Essays toward a morality of consequence*. Cambridge, UK: Cambridge University Press.
- Scanlon, T. J., Luben, R. N., Scanlon, F. L., & Singleton, N. (1993). Is Friday the 13th bad for your health? *British Medical Journal*, 307, 6919: 1584-1586.
- Scott, D. J., Stohler, C. S., Egnatuk, C. M., Wang, H., Koeppe, R. A., & Zubieta, J. K. (2007). Individual differences in reward responding explain placebo-induced expectations and effects. *Neuron*, 55, 2: 325-336.
- Searle, J. (1992). *The rediscovery of mind*. Cambridge, MA: The MIT Press.
- Searle, J. (2001). *Rationality in action*. Cambridge, MA: The

MIT Press.

- Shah, N., & Velleman, J. (2005). Doxastic deliberation. *The Philosophical Review*, 114, 4: 497-534.
- Shermer, M. (2002). *Why people believe weird things: Pseudoscience, superstition, and other confusions of our time* (Rev. ed.). New York: Henry Holt.
- Sobel, D., & Copp, D. (2001). Against direction of fit accounts of belief and desire. *Analysis*, 61, 1: 44-53.
- Stockhorst, U., Enck, P., & Klosterhalfen, S. (2007). Role of classical conditioning in learning gastrointestinal symptoms. *World Journal of Gastroenterology*, 13, 25: 3430-3437.
- Talmont-Kaminski, K. (2008). In a mirror, darkly: Does superstition reflect rationality? *Skeptical Inquirer*, 32, 4: 48-51.
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, 103, 2: 193-210.
- Taylor, S. E., & Brown, J. D. (1994). Positive illusions and well-being revisited: Separating fact from fiction. *Psychological Bulletin*, 116, 1: 21-27.
- Taylor, S. E., & Gollwitzer, P. M. (1995). The effects of mindset on positive illusions. *Journal of Personality and Social Psychology*, 69, 2: 213-226.
- Taylor, S. E., Kemeny, M. E., Reed, G. M., Bower, J. E., & Gruenewald, T. L. (2000). Psychological resources, positive illusions, and health. *American Psychologist*, 55, 1: 99-109.
- Taylor, S. E., Lerner, J. S., Sherman, D. K., Sage, R. M., & McDowell, N. K. (2003). Are self-enhancing cognitions associated with healthy or unhealthy biological profiles? *Journal of Personality and Social Psychology*, 85, 4: 605-615.
- Tertullianus, Quintus Septimius Florens (2010). Tertullian project: "De Carne Christi," chapter 5, verse 4. Retrieved August 7, 2010, from <http://www.tertullian.org/>
- Tiger, L. (1985). *Optimism: The biology of hope* (2nd ed.). New



- York: Kodansha.
- Tiger, L. (1999). Hope springs internal. *Social Research*, 66, 2: 611-623.
- Trivers, R. (1985). *Social evolution*. Menlo Park, CA: Benjamin/Cummings.
- Trivers, R. (2000). The elements of a scientific theory of self-deception. *Annals of the New York Academy of Sciences*, 907: 114-131.
- Van Leeuwen, D. (2007). The spandrels of self-deception: Prospects for a biological theory of a mental phenomenon. *Philosophical Psychology*, 20, 3: 329-348.
- Velleman, J. (1999). The possibility of practical reason. In R. Jay Wallace (Ed.), *Reason, emotion, and will* (pp. 185-218). Brookfield, VT: Ashgate.
- Velleman, J. (2000). *The possibility of practical reason*. New York: Oxford University Press.
- Vyse, S. (2000). *Believing in magic: The psychology of superstition*. New York: Oxford University Press.
- Wegner, D. (2002). *The illusion of conscious will*. Cambridge, MA: The MIT Press.
- Wedgwood, R. (2002). The aim of belief. *Philosophical Perspectives*, 16: 267-297.
- Wenger, A., & Fowers, B. (2008). Positive illusions in parenting: Every child is above average. *Journal of Applied Social Psychology*, 38, 3: 611-634.
- Williams, B. (1973). *Problems of the self*. Cambridge, UK: Cambridge University Press.
- Williams, B. (2002). *Truth and truthfulness: An essay in genealogy*. Princeton, NJ: Princeton University Press.
- Wood, A. (2002). *Unsettling obligations: Essays on reason, reality, and the ethics of belief*. Stanford, CA: Center for the Study of Language and Information.
- Wood, A. (2008). The duty to believe according to the evidence. In E. T. Long & P. Horn (Eds.), *Ethics of belief: Essays in tribute to D. Z. Phillips* (pp. 7-24). Dordrecht, the

Netherlands: Springer.

Zubieta, J., & Stohler, C. (2009). Neurobiological mechanisms of placebo responses. *The New York Academy of Science*, 1156: 198-210.

## 錯誤信念之倫理學

藍亭

國立政治大學心智、大腦與學習研究中心  
11605 台北市指南路二段64號  
E-mail: tlan@nccu.edu.tw

### 摘要

根據艾倫·伍德的「程序原則」，我們應該只接受可以被證成的信念。他認為，無論是個人或群體，接受不為證據所證成的信念會失去自尊、造成墮落。伍德的觀點和大部分分析「信念」概念的哲學家是一致的：信念應該瞄準真理，本文將說明此觀點並提出質疑，我將指出某些實證研究顯示，在特定情況下不瞄準真理反而更能維持自尊、避免墮落，這些不瞄準真理的信念我稱之為T信念。本文將進一步說明T信念的特色，以解釋在何種情況下程序原則應該被違背。

**關鍵詞：**信念、倫理、正向錯覺、低調實在論、穆爾弔詭