

---

## The Crucial Concept of Self-Deception in Presuppositional Apologetics

*By Dr. Greg Bahnsen*

That self-deception which is practiced by all unregenerate men according to the Apostle Paul's incisive description in Romans 1:18ff. is at once religiously momentous and yet philosophically enigmatic. It is also one of the focal points in continuing criticism of Cornelius Van Til's apologetic<sup>[1]</sup> and, as such, invites analysis with a view to supplementing and strengthening the saintly professor's remarkable contribution to the history of apologetics.<sup>[2]</sup>

Paul asserts that all men know God so inescapably and clearly from natural revelation that they are left with no defense for their unfaithful response to the truth about Him. In verses 19-20, Paul says "what can be known about God is plain within them because God made it plain to them... [being] clearly perceived from the created world, being intellectually apprehended from the things that have been made... so that they are without excuse." Nevertheless, even as they are categorically depicted as "knowing God" (v. 21), all men are portrayed in their unrighteousness as "holding down the truth" (v. 18). They are suppressing what God has already successfully shown them about Himself. As a result of hiding the truth from themselves, unbelievers neither glorify nor thank God, but instead become futile in their reasoning, undiscerning in their darkened hearts, and foolish in the midst of their professions of wisdom (vv. 21-22). According to God's word through Paul, then, unbelievers suppress what they very well know, confirming what Jeremiah the prophet so aptly declared, "The heart is deceitful above all things" (17:9).

The apologetical importance of such self-deception should be quite evident. Throughout the history of apologetics we find that Romans 1 has been of guiding interest to Biblically oriented apologists, and indeed the self-deceptive character of man as presented there has itself been stressed periodically by scholars of Reformed persuasion. However, no apologist has drawn more consistent attention to this characteristic of the natural man or made it more pivotal for his system of defending the Christian faith than has Dr. Van Til. It is an indispensable concept in his epistemology, as one will see in systematically studying Van Til's writings or analyzing his apologetical perspective. The point is not simply that references to the unbeliever's self-deception, as taught in Romans 1, are conspicuous and common in Van Til's books, but that this notion functions in such a crucial manner in his argumentation that without it presuppositional apologetics could be neither intellectually cogent nor personally appropriate as a method of defending the faith. A short rehearsal of a few basic points in Van Til's apologetic shows why this is so.

In *A Survey of Christian Epistemology* Van Til claims that "there can be no more fundamental question in epistemology than the question whether or not facts can be known without reference

to God... [and so] whether or not God exists."<sup>[3]</sup> That is, a metaphysical issue is the most fundamental question in epistemology. Van Til's apologetical argument for the metaphysical conclusion that God exists, however, is in turn epistemological in character. The Christian defends the faith "by claiming... he can explain... [the] amenability of fact to logic and the necessity and usefulness of rationality itself in terms of Scripture."<sup>[4]</sup> He could thus write: "it appears how intimately one's theory of being and one's theory of method are interrelated."<sup>[5]</sup> This mutual dependence of metaphysics and epistemology has always been characteristic of Van Til's apologetical position.<sup>[6]</sup>

So then, far from being a species of "fideism," as it is so often misconstrued by writers like Montgomery, Geisler or Sproul,<sup>[7]</sup> Van Til's approach to the question of God's existence offers, I believe, the strongest form of proof and rational demonstration - namely, a "transcendental" form of argument. He writes, "Now the only argument for an absolute God that holds water is a transcendental argument... [which] seeks to discover what sort of foundations the house of human knowledge must have, in order to be what it is."<sup>[8]</sup> To put it briefly, using Van Til's words, "we reason from the impossibility of the contrary."<sup>[9]</sup>

In *The Defense of the Faith*, Van Til explains that this is an indirect method of proof, whereby the believer and the unbeliever together think through the implications of each other's most basic assumptions so that the Christian may show the non-Christian how the intelligibility of his experience, the meaningfulness of logic, and the possibility of science, proof or interpretation can be maintained only on the basis of the Christian worldview (i.e., on the basis of Christian theism taken as a unit, rather than piecemeal).

The method of reasoning by presupposition may be said to be indirect rather than direct. The issue between believers and non-believers in Christian theism cannot be settled by a direct appeal to "facts" or "laws" whose nature and significance is already agreed upon by both parties to the debate. The question is rather as to what is the final reference-point required to make the "facts" and "laws" intelligible.... The Christian apologist must place himself upon the position of his opponent, assuming the correctness of his method merely for argument's sake, in order to show him that on such a position the "facts" are not facts and the "laws" are not laws. He must also ask the non-Christian to place himself upon the Christian position for argument's sake in order that he may be shown that only upon such a basis do "facts" and "laws" appear intelligible.... The method of presupposition requires the presentation of Christian theism as a unit.[\[10\]](#)

Taking Christian theism "as the presupposition which alone makes the acquisition of knowledge in any field intelligible," the apologist must conduct a critical analysis of the unbeliever's epistemological method "with the purpose of showing that its most consistent application not merely leads away from Christian theism, but in leading away from Christian theism, leads to [the] destruction of reason and science as well."[\[11\]](#) This point, which Van Til drives home persistently throughout his large corpus of publications, is expressed with these words in *A Christian Theory of Knowledge*: "Christianity can be shown to be, not 'just as good as' or even 'better than' the non-Christian position, but the only position that does not make nonsense of human experience."[\[12\]](#) Because the unbeliever's commitment to random eventuation in history (i.e., a metaphysic of "chance") renders proof impossible, predication unintelligible, and a rational/irrational dialectic unavoidable, Van Til claims repeatedly in his writings that the truth of Christianity is epistemologically indispensable.[\[13\]](#)

It is in this sense, then, that the presuppositional argument for the existence of God and the truth of the Bible is "from the impossibility of the contrary."

The argument for the existence of God and for the truth of Christianity is objectively valid.... The argument is absolutely sound. Christianity is the only reasonable position to hold. It is not merely as reasonable as other positions, or a bit more reasonable than other positions; it alone is the natural and reasonable position for man to take.[\[14\]](#)

"Christianity is proved as being the very foundation of the idea of proof itself."[\[15\]](#) Admittedly those are rather strong claims, and as I see it, they constitute the most rigorous apologetical program of intellectual defense being advanced in our time. It is, moreover, just in the all-or-nothing epistemological boldness of presuppositionalism that Van Til finds the distinctiveness of Reformed apologetics - what he calls "the basic difference" between it and other types of defense.

The Romanist-evangelical type of apologetics assumes that man can first know much about himself in the universe and afterward ask whether God exists and Christianity is true. The Reformed apologist assumes that nothing can be known by man about himself or the universe unless God exists and Christianity is true.[\[16\]](#)

Ironically, those who are uneasy with the presuppositional approach to apologetics include not only those who think that it, being fideistic, does not prove enough, but also those who (reading the claims that we have just cited) say that it proves far too much! The charge is made, you see, that presuppositionalism implies that unbelievers can know nothing at all and can make no contribution to science and scholarship since belief in God is epistemologically indispensable according to the presuppositionalist. And it is right here, right at this crucial point in the analysis, that the notion of self-deception by the unbeliever enters the picture.

Van Til always taught that "the absolute contrast between the Christian and the non-Christian in the field of knowledge is said to be that of principle." He draws "the distinction... between the regenerated consciousness which in principle sees the truth and the unregenerate consciousness which by its principle cannot see the truth."[\[17\]](#) If unbelievers were totally true to their espoused assumptions, then knowledge would indeed be impossible for them since they deny God. However the Christian can challenge the non-Christian approach to interpreting human experience "only if he shows the non-Christian that even in his virtual negation of God, he is still really presupposing God."[\[18\]](#) He puts the point succinctly in saying: "Anti-theism presupposes theism."[\[19\]](#) The intellectual achievements of the unbeliever, as explained in *The Defense of the Faith*, are possible only because he is "borrowing, without recognizing it, the Christian ideas of creation and providence."[\[20\]](#) The non-Christian thus "makes positive contributions to science in spite of his principles"[\[21\]](#) - because he is inconsistent. Van Til replies directly to the charge that we are now considering with these words:

The first objection that suggests itself may be expressed in the rhetorical question "Do you mean to assert that non-Christians do not discover truth by the methods they employ?" The reply is that we mean nothing so absurd as that. The implication of the method here advocated is simply that non-Christians are never able and therefore never do employ their own method consistently.... The best and only possible proof for the existence of such a God is that his existence is required for the uniformity of nature and for the coherence of all things in the world.... Thus there is absolutely certain proof for the existence of God and the truth of Christian theism. Even non-Christians presuppose its truth while they verbally reject it. They need to presuppose the truth of Christian theism in order to account for their own accomplishments.[\[22\]](#)

The sense of deity discussed by Calvin on the basis of Paul's doctrine in Romans 1 provides Van Til not only with an apologetical point of contact, but also with an account of how those who disclaim any belief in God can know much about most subjects.[\[23\]](#)

The knowledge of God which every man has as the image of God and as surrounded by God's clear revelation assures us, then, that all men are in contact with the truth.[\[24\]](#) Not even sin in its most devastating expressions can remove this knowledge, for Van Til says "sin would not be sin except for this ineradicable knowledge of God."[\[25\]](#) It is this knowledge of God, of which Paul speaks in Romans 1, that Van Til identifies as the knowledge which all men have in common, contending that such common knowledge is the guarantee that every man can contribute to the progress of science, and that some measure of unity in that task can exist between believers and unbelievers.[\[26\]](#)

Because he is convinced that self-consciousness presupposes God-consciousness,[\[27\]](#) the presuppositionalist can assert then, in the most important sense, "There are no atheists."[\[28\]](#) Van Til clearly relies very heavily on Paul in making such a surprising claim.



The apostle Paul speaks of the natural man as actually possessing the knowledge of God (Rom. 1:19-21). The greatness of his sin lies precisely in the fact that "when they knew God, they glorified him not as God." No man can escape knowing God. It is indelibly involved in his awareness of anything whatsoever.... We have at once to add Paul's further instruction to the effect that all men, due to the sin within them, always and in all relationships seek to "suppress" this knowledge of God (Rom. 1:18).... Deep down in his mind every man knows that he is the creature of God and responsible to God. Every man, at bottom, knows that he is a covenant breaker. But every man acts and talks as though this were not so. It is the one point that cannot bear mentioning in his presence.[\[29\]](#)

Van Til speaks of the unbeliever sinning against his "better knowledge" - that "it is of the greatest possible importance" to acknowledge that man knows God in some "original sense."[\[30\]](#)

Now then, just because knowledge is a category of belief (viz., justified true belief), and because it can reduce unnecessary philosophical complications throughout this discussion, we could just as well speak of the unbeliever's suppressed belief about God as we could speak of his suppressed knowledge of God. In fact, Van Til makes his point in just that way also in his writings.

To be sure, all men have faith. Unbelievers have faith as well as believers. But that is due to the fact that they too are creatures of God. Faith therefore always has content. It is against the content of faith as belief in God that man has become an unbeliever. As such he tries to suppress the content of his original faith.... And thus there is no foundation for man's knowledge of himself or of the world at all.... When this faith turns into unbelief this unbelief cannot succeed in suppressing fully the original faith in God. Man as man is inherently and inescapably a believer in God. Thus he can contribute to true knowledge in the universe.[\[31\]](#)

Our brief rehearsal of presuppositional apologetics has brought us step by step to the realization that a crucial component in Van Til's perspective, one that is necessarily contained in any credible account of its functioning, is the conviction that the non-Christian is self-deceived about God - that the one who does not believe in God actually does believe in God. The cogency of presuppositionalism is tied up with the intelligibility of this notion of self-deception. If we do not find our point of contact with the unbeliever in his suppressed knowledge of God and reason with him in such a way as to "distinguish carefully between the natural man's own conception of himself and the Biblical conception of him" - that is, if we do not proceed on the firm premise that the unbeliever is engaged in self-deception of the most significant religious kind - then, according to Van Til, we "cannot challenge his most basic epistemological assumption" that his reasoning can indeed be autonomous. And immediately Van Til adds, "on this everything hinges."[\[32\]](#)

The concept of self-deception is critical to Van Til's presuppositionalism. Everything hangs on it, according to him. If there should be something suspect or muddled about the notion of self-deception here, then the entire presuppositional system of thought is suspect and unacceptable as well. Its key argumentive thrust relies completely on the truth of the claim that unbelievers are suppressing what they believe about God the Creator. That is why I stated at the beginning that the self-deception as depicted in Romans 1 is religiously momentous and also why the unbeliever's self-deception is a pivotal notion - a *sine qua non* truth - for the presuppositional method of defending the faith.

However, as I also wrote at the outset of this essay in reference to Romans 1, the notion of self-deception is philosophically enigmatic. It is more than just a bit odd, is it not, to say that someone believes what he does not believe! Indeed, it sounds downright self-contradictory. At just the crucial point where the presuppositionalist must make reference to clear and compelling considerations in order to give a justifying and credible account of the very heart of this apologetical method, he seems to take an unsure step into philosophical perplexity. It hardly seems to the critics of presuppositionalism that its account of itself explains the unclear in terms of the clear. It appears rather to move from the unclear to the even more unclear. For now the obvious question, if not challenge, will arise: what could it mean for an *unbeliever* to simultaneously be a *believer*? Is the notion of self-deception at all coherent?

The quite enigmatic character of his conception of the unbeliever as self-deceived is confessed very plainly in Van Til's writings, where he admits that the problem of the unbeliever's knowledge "has always been a difficult point..., often the one great source of confusion on the question of faith and its relation to reason."[\[33\]](#) Van Til insists that we must do justice to the twin facts that every unbeliever knows God, and yet, that the natural man does not know God. If we do not stress these two points, following Romanist and Arminian apologists, then we will necessarily allow for a compromising apologetic.[\[34\]](#) Van Til was aware of the counter charge that was likely to be made.

It is ambiguous or meaningless, says the Arminian, to talk about the natural man as knowing God and yet not truly knowing God. Knowing is knowing. A man either knows or he does not know. He may know less or more, but if he does not "truly" know, he knows not at all.... In reply to this the Calvinist insists that... the natural man does not know God. But to be thus without knowledge, without living, loving, true knowledge of God, he must be one who knows God in the sense of having the sense of deity (Romans 1).[35]

As we can see, Van Til was appropriately sensitive to the charge of self-contradiction. Accordingly he wanted to draw some kind of distinction which would indicate that he, with Paul, was not taking away with one assertion what he gives in another. Thus he qualified his statements. "Non-Christians know after a fashion, as Paul tells us in Romans." [36] Elsewhere he writes that "there is a sense in which all men have faith and all men know God. All contribute to science." [37] Therefore he taught "there are two senses to the word 'knowledge' used in Scripture." [38]

A common way in which Van Til denominates those two senses, and the difference between them, is by saying that unbelievers know God but "not according to the truth," or they do not "truly" know him, or they do not have "true knowledge." [39] How is this to be construed? Unbelievers presuppose (and hence believe) the truth of God and of Christianity "while they verbally reject it." The non-Christian "acts and talks as though this were not so," for he cannot bear the mentioning of his knowledge of God. [40] Why not? Van Til says all sinners "have an ax to grind and do not want to keep God in remembrance. They keep under the knowledge of God that is within them. That is they try as best they can to keep under this knowledge for fear they should look into the face of their judge." [41] Being troubled in conscience, the unbeliever must make an effort "to hide the facts from himself," somewhat like a cancer victim who, in distress, keeps the awareness of the truth at a distance from himself. [42] Some students of presuppositionalism have made, I think, the hasty error of conceiving of this situation as a simple matter of lying. The unbeliever, it is thought, knows God, but simply says that he does not know God. However, Van Til did not take this artificial and simplistic route. He recognized that the unbeliever's situation is epistemologically strange and hard to describe accurately (unlike the lying scenario). On the one hand, Van Til portrayed the unbeliever as holding this knowledge of God "subconsciously." The non-Christian is said to borrow Christian ideas "without recognizing it." [43] "He knows deep down in his heart" or "deep down in his mind," [44] so that the natural man's knowledge of God is taken as "beneath the threshold of his working consciousness." [45] And yet on the other hand Van Til wanted to contend unequivocally for the sinful guilt of men who suppress the knowledge of God. Thus they are also portrayed by him as somehow conscious of what they are doing. Knowing that it cannot successfully be done, says Van Til, the unbeliever pursues the impossible dream of moral and epistemological autonomy, seeking to suppress what he knows about God. [46] Van Til writes, "He knows he is a 'liar' all the time," [47] and accordingly his denying of the truth is a self-conscious act. And yet in saying this, Van Til immediately felt the need to place a qualification on his claim. Notice that the word 'liar' in the preceding quotation is placed conspicuously in quotes. Van Til wants to say it with some measure of reservation. Elsewhere he explained that the unbeliever's hostility is not "wholly self-conscious." [48] To his qualitative distinction (knowledge/true

knowledge), and to his spatial distinction (knowing/knowing deep down), he now adds a quantitative distinction (wholly self-conscious/partially self-conscious).

Again it must be borne in mind that when we say that fallen man knows God and suppresses that knowledge so that he, as it were, sins self-consciously, this too needs qualification. Taken as a generality and in view of the fact that all men were represented in Adam at the beginning of history, we must say that men sin against better knowledge and also self-consciously. But this is not to deny that when men are said to be without God in the world they are ignorant.... There is therefore a gradation of those who sin more and those who sin less, self-consciously.[\[49\]](#)

One way or another, however, Van Til teaches that the natural man is "ethically responsible" for his suppressing of the truth.<sup>[50]</sup> He states that "the Scriptures continue to hold man responsible for his blindness,"<sup>[51]</sup> and he calls the result of the unbeliever's self-deceptive effort "culpable ignorance."<sup>[52]</sup> The reason for his failure to recognize God as he should "lies exclusively in himself," says Van Til; it is nothing less than "willful transgression" which accounts for his refusal.<sup>[53]</sup> So again, Van Til has indicated how awkward it is to speak of the unbeliever as self-deceived. On the one hand, the unregenerate's knowledge is considered sub-conscious, and he does not recognize his utilizing of it. And yet on the other hand, the unregenerate is portrayed as actively seeking to suppress it, and in some measure he consciously and willfully works to hide it from himself. Van Til runs his reader from pole to pole. On the one hand he does not want to say that the unbeliever is a bare liar, and yet on the other hand he does want to say that the unbeliever is fully culpable, just like any liar would be.

Given this short review of Van Til's discussion of the apologetical situation, we have learned (1) that a recognition of the unbeliever's self-deception is indispensable to presuppositional apologetics, and yet (2) that its recognition is fraught with obscurity. As long as the notion of self-deception appears uncertain, awkward, or unclear, the cogency of the presuppositional method will remain in the balance. We must say in conformity to Romans 1 that in some sense the non-Christian knows and does not know God. In some sense, he believes, but disbelieves in God. In some sense, he is unconscious of suppressing the truth and still responsibly conscious of doing so. So then, what might prove especially beneficial would be for us to give some sense to these apparent paradoxes. If we can do so, the philosophy of presuppositionalism will be noticeably advanced and more readily presentable to struggling defenders of the faith who need it so desperately.<sup>[54]</sup>



## An Enigmatic yet Familiar Notion

In working toward a solution to the problem of self-deception, we should pause at the outset to observe that while Paul's (and Van Til's) use of that concept may be perplexing, the concept itself has certainly not been unfamiliar. Portraying men as self-deceived has been a virtual commonplace in Western literature, and thus the apparently paradoxical nature of the concept cannot be thought to be a uniquely religious matter.

Popular, cynical platitudes about man's proclivity to self-deception have been published continually by men from Demosthenes to Benjamin Franklin, who once quipped, "who has deceived thee so often as thyself?" The Puritan preacher, Daniel Dyke, wrote a four-hundred page treatise published in 1617, entitled *The Mystery of Self-Deceiving*. A century later, the Anglican apologist, Bishop Butler, included his famous sermon "Upon Self-Deceit" in a published collection of his sermons. In it he correctly recognized, "A man may be entirely possessed of this unfairness of mind, without having the least speculative notion what the thing is."<sup>[55]</sup> It has been common to make mention of self-deception, even though it may be uncommonly difficult to explain philosophically just what it is.

Yet even among philosophers the notion has been common stock. From what was said about it by Plato, Rousseau, Goethe, Schopenhauer, and Nietzsche, one would learn how dubious a view it is that men really want the truth when the truth happens to be uncomfortable for them. Special attention is given to the concept of self-deception in Hegel's theory of "unhappy consciousness," in Kierkegaard's discussion of "purity of heart," and Sartre's view of "bad faith." According to Sartre, men evade responsibility for their existential freedom through intentional ignorance of the human reality.

Apart from the obscure works of the philosophers, however, self-deception is also one of those human realities on which great works of Western literature have been richly sustained over many years. One thinks of the classic portrayal of it in Sophocles' *Oedipus Rex* or Shakespeare's *The Tragedy of King Lear*. We remember the soliloquy on self-swindling in Dickens' *Great Expectations*, Emma's intrigues with lovers in Flaubert's *Madame Bovary*, or Strether's efforts to remain oblivious to unwanted evidence in Henry James' *The Ambassadors*. The tragic condition of self-deception is discussed and depicted in great Russian literature of the past - such as Dostoevsky's *Notes from Underground*, Tolstoy's *Anna Karenina*, *Father Sergius*, and *The Death of Ivan Ilych*. Indeed, one of the most graphically accurate depictions of self-deception is found in Tolstoy's *War and Peace*, when Count Rostov returns home from a business trip to discover that something has happened to his daughter. We read:

The Count saw clearly that something had gone wrong during his absence; but it was so terrible for him to imagine anything discreditable occurring in connection with his beloved daughter, and he so prized his own cheerful tranquility, that he avoided asking questions and did his best to persuade himself that there was nothing very much wrong or out of the way....[\[56\]](#)

The illustrations from literature could be multiplied many times over. We could mention O'Neill's *The Iceman Cometh*, or Andre Gide's *Pastoral Symphony*, or Camus' *The Fall*, or any number of other entertaining, perplexing accounts.

We still would not be fully aware of how common the notion of self-deception has been in human thought until we supplemented the survey with those sociological and psychological approaches to man which have so profoundly affected Western culture in the last century. One thinks here, of course, of the discussion by Marx of "false consciousness" and collective illusion, causing an entire social class to obscure the motives of its thought from itself. We recall the sociology of knowledge presented by Karl Mannheim, who pointed to the tenacity of commitment to theoretical formulations which, although impractical, have been acquired in the cooperative process of group life. Finally, we cannot overlook Freud's psychoanalytic study of subconscious maneuvers and defense mechanisms by which men cling to their cherished illusions.

So whether we turn to works in religion, philosophy, literature, sociology or psychology, we cannot come to the conclusion that the notion of self-deception is somehow an unfamiliar one. We have ample evidence that men identify something in their experience as self-deception. The notion is readily utilized in everyday conversation, not simply in published works of scholars. The vocabulary of self-deception is recognizable (even by children), mastered by people, and taught to others. And so, when the son of Mrs. Jones has been caught red-handed stealing lunch money out of students' desks at school, and Mrs. Jones continues to protest her son's innocence - despite this being the third time such an incident has taken place, despite her discomfort and red face when the subject of dishonesty comes up in casual conversations, despite the fact that she does not trust her son around her purse any longer - and she continues to explain his innocence with strange explanations (like the school officials have a vendetta against little Johnny, they were framing him, etc.) nobody finds it awkward to say the poor lady "is deceiving herself." You see, self-deception is part of our common experience, and familiarity with it breeds acceptance of it as a genuine reality of life.

## The Apparent Paradox and Search for a Solution

Our ready acceptance of the phenomenon of self-deception, however, has been challenged over the last thirty-five years; philosophical attention has been given to conceptual questions about self-deception which arise in both the theory of knowledge and the philosophy of mind.

The analytical-epistemological approach to the subject was somewhat anticipated in Bertrand Russell's critique of Freud in *The Analysis of Mind* (1921) and in Gilbert Ryle's criticism of mind-body dualism in *The Concept of Mind* (1949). Russell spoke of desire-motivated beliefs (or wishful thinking), and Ryle pointed out that the practice of self-deception challenges the common dualist assumption that man has some direct introspective knowledge of the workings of his own mind, a knowledge free from illusion and doubt. However critical, intense and thorough philosophical scrutiny of the notion of self-deception was inaugurated in 1960 by Raphael Demos in his pioneering article entitled "Lying to Oneself." [57] A long series of reactions and counter-proposals has developed in the philosophical journals since that time. Now inquiry was made into just what self-deception must involve to qualify as such, and into whether it is a feat which can literally be accomplished. Analyses of the notion always seemed headed for some form of paradox.

You see, the natural thing to do is to model self-deception on the well-known activity of other-deception. Deceiving oneself is thought of as a version of deceiving someone else. A problem here, of course, is that in other-deception the roles of deceiver and deceived are incompatible; yet in self-deception a person is thought to play both of these incompatible roles himself! Sartre put the matter plainly in his book *Being and Nothingness*.

It follows first that the one to whom the lie is told and the one who lies are one in the same person, which means that I must know in my capacity as deceiver the truth which is hidden from me in my capacity as the one deceived. Better yet I must know the truth exactly in order to conceal it more carefully - and this not at two different moments, which at a pinch would allow us to re-establish the semblance of duality - but in the unitary structure of a single project. How then can the lie subsist if the duality which conditions it is suppressed?[\[58\]](#)

Let us stop and analyze the situation. In a case of other-deception, Jones is aware that some proposition is false, but Jones intends to make Smith believe that it is true - and he succeeds. If we take Smith out of the picture and substitute in Jones, so as to gain "*self*-deception," we end up saying "Jones, aware that p is false, intends to make himself believe that p is true, and succeeds in making himself believe that p is true."[\[59\]](#) Such a statement is surely puzzling, for it suggests, "that somebody could try to make, and succeed in making, himself believe something which he, *ex hypothesi*, at the same time believes not to be true."[\[60\]](#) It would be easy to conclude, then, that self-deception is an incoherent project that cannot be fulfilled.

So we are forced to ask whether there actually is such a thing as perpetrating a deception on oneself. How could it occur in practice? How could it be described without contradiction? How can someone, after all, as deceived, *believe* p, yet as deceiver *disbelieve* p? It now appears that self-deception, despite the familiarity of the notion, is about as difficult to do as presiding over one's own funeral. When we introduce the element of mendacity (dishonesty, lying) into the picture, the problem is even further complicated. Here we move from epistemic notions about belief into the philosophy of mind with questions about consciousness, purpose, and intention. There have been "weak models" of self-deception proposed by some philosophers, intending to take the sting out of the paradox by maintaining that an agent does not know what he is up to in self-deception.[\[61\]](#) In "strong" self-deception the enterprise is purposeful and not so innocent. And it is this *strong* version of self-deception which is usually thought necessary for moral *culpability* in self-deception. This approach, however, only intensifies the philosophical perplexity involved in the notion, for the kind of thought that goes into planning and executing what you are doing in purposefully deceiving someone else, makes doing it to yourself seem impossible. "Self-deception is not a matter of mere stupidity or carelessness in thinking. It is a craftily engineered project, and this is why it seems pointless and self-contradictory."[\[62\]](#)

So then, the analytical-epistemological approach to the literature on self-deception in recent years makes us hesitant to speak of it confidently and clearly. And the maze of philosophical treatments given to the paradoxical notion only intensifies our confusion. Herbert Fingarette, in the first full book published on the subject, summarizes the problem nicely:

Were a portrait of man to be drawn, one in which there would be highlighted whatever it is most human, be it noble or ignoble, we should surely place well in the foreground man's enormous capacity for self-deception. The task of representing this most intimate, secret gesture would not be much easier were we to turn to what the philosophers have said. Philosophical attempts to elucidate the concept of self-deception have ended in paradox - or in loss from sight of the elusive phenomenon itself. . . . We are beset by confusion when once we grant that the person himself is in self-deception. For as deceiver one is insincere, guilty; whereas as genuinely deceived one is an innocent victim. What, then, should we make of the self-deceiver, the one who is both the doer and the sufferer? Our fundamental categories are placed squarely at odds with another. . . . 'The one who lies with sincerity,' who convinces himself of what he even knows is not so, who lies to himself and to others and believes his own lie though in his heart he knows that it is a lie - the phenomenon is so familiar, the task so easy, that we nod our heads and say, 'of course.' Yet when we examine what we have said with respect to our inner coherency, we are tempted to dismiss such a description as nonsense.[\[63\]](#)

At this juncture we can take the route of denying the reality of self-deception or the route of resolving the apparent contradiction involved in the notion. My procedure will be to take self-deception as a *datum*, and thus I am committed to saying that at best it is only apparently self-contradictory. While it is not inconceivable that those many people who have made use of the notion of self-deception over the centuries have been unwittingly contradicting themselves, it is still not very likely. We resist the conclusion that self-deception is actually impossible because we know that people do not merely play at self-deception. They engage in it in tragic ways, and very often they later come to realize the fact (for instance, think here of that devastating book by Albert Speers, *Inside the Third Reich*). Given Paul's teaching in Romans 1 - not to mention the actual use of the phrase 'to deceive oneself' in James 1:26 and 1 John 1:8 - the Christian especially will want to resist dismissing self-deception as an incoherent impossibility. Most people, then, will be more sure that self-deception occurs than they would be of any explanation which renders it only apparent. So whenever we confront an account of self-deception which makes it appear self-contradictory, our assumption should be that the confusion lies not in the notion of self-deception but in the person's philosophical account of it. Accordingly our work is cut out for us: as elusive as it may be, we are committed to finding an adequate and coherent analysis of self-deception.

What will be required of us if we are going to succeed? The basic requirement for an acceptable analysis of self-deception is simply that it must "save the phenomenon," while at the same time respecting the law of contradiction. Thus our account must be descriptively accurate - true to paradigm examples of self-deception. It is useful here to recall Wittgenstein's warnings against a reductionistic "craving for generality" which is "contemptuous of the particular case." We must admit at the outset that the many and varied uses for the term 'self-deception' bear a "family resemblance" to each other.<sup>[64]</sup> Doubtless there will be borderline cases, where ambiguous evidence makes it difficult to tell if all of the usual elements of self-deception are present. There will be extreme cases where some element of self-deception is accentuated out of proportion - even as the colloquial exclamation "That's insane!" is an exaggeration of the literal and proper use of the concept of insanity. There will be analogous cases, deficient cases, peculiar cases, and on and on. Nevertheless, there are typical or paradigmatic cases from which we learn to use the expression "self-deception" and apply it to further, diverse cases. Our use of this vocabulary is not so ad hoc as to preclude the possibility of our picking out genuine cases of self-deception. So I will aim to give necessary and sufficient conditions for the truth of the assertion, "S deceived himself into believing that p," as it is taken in the full-fledged and paradigmatic sense.

In order to be descriptively correct, our analysis must not radically depart from ordinary language. Nor must it confuse or merge self-deception with related and similar phenomena in human experience (e.g., ignorance, wishful thinking, change of belief). Beyond being accurate and exact, our account must also be completely rid of any incoherence, which requires using clearly defined notions in the analysis so that self-contradiction (or its absence) is detectable. We do not want to explain self-deception, moreover, by appealing to concepts which are even less clear than the one we are attempting to understand - for example, by an ambiguous and misconceived distinction between "psychological knowing" and "epistemological knowing," which is easily faulted as obscure, if not simply wrong. Yet on the other hand, we do not want to make the analysis so pat and easy that the perplexing element in self-deception is dismissed altogether, causing us to wonder why it should ever have appeared problematic to begin with (for instance, by drawing a trivial distinction between what someone ought to know and what he actually does know - a



strategy which brings self-deception down to the level of any mundane oversight in one's thinking, such as not knowing your father's age).[65]

Within the guidelines we have rehearsed here, we need to formulate an adequate analysis of self-deception. While existentialist treatments (e.g., Sartre, Fingarette) affirm the contradiction found in self-deception as an experienced reality, the analytic tradition has offered various avenues for removing the apparent logical difficulties. In the philosophical journals, you will notice three basic strategies for resolving the paradox.

The first strategy is to deny that there is a parallel between self-deception and other-deception. Some maintain that deception is inherently other-regarding, and thus the skeptical conclusion is advanced that there actually is no such thing as self-deception. What is commonly called "self-deception" needs to be given a more accurate description.[66] Others say that words like "deceive," "know," or "believe" are used in a non-standard fashion in accounts of self-deception, not having the same intended sense as in descriptions of other-deception.[67] Finally, others who deny the other-deception parallel recommend that we "look and see" what conditions actually hold when self-deception locutions are utilized, in which case we will notice that self-deception situations do not involve two incompatible beliefs (as in other-deception), but rather only a particular kind of single belief entertained under peculiar circumstances. Thus we speak of "self-deception" when we want to reprimand irresponsible holding of an unwarranted belief,[68] or self-deceived beliefs are taken as those held in belief-adverse circumstances,[69] or where there is an irrational refusal to look at evidence,[70] or where one simply desires to hold the belief,[71] or where weak-willed dishonesty permits desire-generated blindness,[72] or some emotion has irrationally obscured the contrary evidence.[73]

The second strategy is to accept the other-deception model (the reality of perpetrating a deception upon oneself) and maintain that self-deception is a conflict state of holding incompatible beliefs, but then resolving the paradox of believing contrary things by introducing various kinds of distinctions. Some distinguish between knowledge and "as-it-were-knowledge,"[74] or between full belief and "half-belief,"[75] contending that the different senses for this epistemic vocabulary in analyses of self-deception render the paradox only apparent. Other philosophers treat self-deception as a literal case of other-deception, positing some kind of duality (e.g., levels of consciousness, split personality) within the self-deceived person himself.[76] Another approach is to draw a temporal distinction between S-the-deceiver and (later) S-the-deceived.[77] Finally, many writers have attempted to give a coherent account of self-deception as a conflict state of incompatible beliefs by drawing some kind of distinction regarding consciousness - for instance, distinguishing two levels of awareness,[78] or between general and explicit consciousness,[79] or between general awareness and detailed awareness,[80] or between conscious purpose and unreflective purpose,[81] or between conscious and unconscious knowledge,[82] or between strong and weak consciousness.[83]

The third strategy proposes to utilize an altogether different model for self-deception which avoids appeal to such epistemic terms as "knowledge" or "belief," using instead a volition-action model wherein one fails to "spell-out" for himself his engagements in the world. In this way it is thought we can preserve the purposiveness and culpability essential to any adequate account of the

phenomenon, yet avoiding the paradoxes which have proved inherent in the epistemic accounts of self-deception.[\[84\]](#)

My evaluation is that none of these three major strategies for resolving the apparent paradox will pass the tests of adequacy prescribed above. In some cases we find necessary, but not sufficient, conditions for self-deception set forth (e.g., adverse evidence, the influence of desire on human belief). In other cases necessary conditions are dismissed altogether (e.g., belief, incompatible beliefs). Some proposals merely state all over again the need for a resolution to the problem (e.g., those using new senses for the epistemic vocabulary), or else they reintroduce the paradox at a different point (e.g., having a policy of not spelling-out an engagement in the world). Some suggestions end up reducing self-deception to something else (e.g., reducing it to a change of belief, ignorance, cognitive error, or pretending) and thereby render the notion dispensable. Another group of attempted solutions rely on notions which are even more obscure or problematic than self-deception itself (e.g., diverse kinds of consciousness), escaping the appearance of paradox at the price of equivocating on just what the self-believer believes he is aware of. Other analyses confuse or merge self-deception with one of many related states or actions (e.g., with wishful thinking, delusion, simple trust, vacillation of opinion, obstinacy, or motivated belief). Virtually all of the authors who have written on the subject have contributed some helpful insights into the difficult issue of self-deception, and I will draw from many of them in my own proposed resolution to the apparent paradox. However, I am not convinced that these writers have been fully true to the phenomenon or have escaped paradox.

## Belief and Its Characteristics

There is something of a cognitive mess at the core of our lives. We are inconsistent in our choices, incoherent in our convictions, persuaded where we ought not to be, and deluded that we know ourselves transparently. The concept of belief shows up in all of these kinds of personal failures, and it should seem obvious that it does as well in the kind of cognitive error we call "deception." Deceived people have been misled, deluded, beguiled or somehow mistaken in what they think and expect to be the case; they engage in false believing. There are few (if any) plausible grounds for disputing the claim that self-deception involves holding one or more false beliefs. Ordinarily in everyday thinking we construe self-deception in terms of belief (of some variety, under some circumstance, etc.).

Fingarette, however, proposes as an alternative analysis a volitional account of self-deception which, stressing the element of intentional ignorance, takes it to be a kind of action rather than a kind of belief. Consciousness is an active and vocal power (rather than, as traditionally thought, passive and visual), and a person becomes explicitly conscious of something through an intentional act of "spelling out his engagements in the world." Sometimes, though, there are overriding reasons for a person to avoid spelling out these engagements, as when doing so would be destructive of his self-conception or the personal identity he has achieved. Lest the effort to avoid spelling out the engagement itself reveal the engagement, one must avoid spelling out that effort as well. Self-deception thus involves adopting an avoidance policy whereby one purposefully chooses to stay ignorant of some engagement in the world.

This is an inadequate alternative to belief-analyses of self-deception, in the first place, because the troublesome concept of self-deception is explained at the price of even greater obscurity (the unfamiliar metaphor of "spelling out an engagement in the world"). Secondly, the volition-action family of terms (which Fingarette prefers for explaining self-deception) is itself heavily laden with notions involving cognitive or epistemic terms like "belief," "knowledge," "perception," etc. A further difficulty is that Fingarette's analysis overlooks completely those cases of self-deception which involve an artificial and misleading overdoing of spelling out one's engagements in the world with an inappropriate emotional detachment - the very opposite of Fingarette's avoidance policy. Finally, Fingarette's alternative account does not rid the notion of self-deception of paradox, but simply restates the paradox in new terms. The effort to avoid spelling out one's (preceding) effort to avoid spelling out a distressful engagement in the world makes one conscious of making oneself unconscious.

Others use the word "deceive" in a way which does not seem to make believing false propositions essential to the act. Freudian psychologists speak of the self-deceived person as being in the grip of unconscious motivations (without mention of cognitive processes). Kierkegaard spoke of a person's failure to be true to himself and ethically consistent as self-deception. However, Freudian and existentialist uses of "deception" are either figurative language or implicitly employ the cognitive sense of believing. If we are unable to cash in talk of unconscious motives and true selves into descriptions of ourselves which can be believed, it makes little sense to say we are "being false" to ourselves or "living a lie." Even when we say the husband who is unfaithful to a knowing wife (they do not speak to each other of his indiscretions) has "deceived" her, we mean he has

violated her expectations- in which case the cognitive sense of "deceive" is again waiting in the wings.

There is simply no good reason to omit reference to belief in a proper analysis of self-deception. More particularly, what is essential in self-deception is that people hold a false belief - not simply an unwarranted belief (e.g., the patient who chooses to disbelieve his doctor's report of cancer, only to turn out right in his wishful thinking), and not simply the absence of expected belief (e.g., the cuckold who literally thinks nothing about his wife's infidelity, although the neighborhood is loud with rumors and she has too many shady late-night excuses). Even where people deceive themselves about their attitudes, hopes, emotions, etc. (e.g., false security, false pride), the objects of self-deception themselves have a cognitive core. The parent who is inappropriately proud of his child's report card experiences a certain emotion only by believing something about the marks on the card. About the colleague who shows false sorrow over a fellow worker's firing we say, "He may think that he is sorry, but he knows quite well he is delighted over this turn of events."

I would maintain, then, that self-deception, as a form of deception, involves believing false propositions. Further, the mistaken believing which is involved is fully genuine believing. We do not here speak of "belief" in some odd, defective, or "twilight" sense. The self-deceiver is not merely feigning ignorance or being an obvious hypocrite. He is concerned with the truth and makes efforts, albeit strained, to sustain his false belief as rational. He is aware of the weight and relevance of the evidence contrary to his belief, so he distorts the evidence through pseudo-rational treatment of it. He is not simply pretending. Although his twisting of the evidence shows that he is trying to convince himself of something unlikely, he still behaves in ways which rely upon the truth of what he says about his (false) belief. He must say that he really believes the false proposition, or else he would not be "deceived" after all. This is not simply half-belief or near-belief, for that proposal would reduce self-deception to mere vacillation, lack of confidence, or insincerity. There is no lack of evidence for the self-deceiver's full-fledged believing; it is just that we have too many beliefs of his for which there is adequate evidence - beliefs which are incompatible. Moreover, the self-deceiver's false belief is not simply performatory in character (an avowal which initiates a commitment about which he will not follow through), for that would reduce self-deception to personal determination, striving, hoping contrary to fact, or wishful thinking.

We must turn attention, then, to the concept of belief if we would hope to analyze self-deception adequately. This is a safe and promising move because the concept of belief is familiar to everyone (despite notorious philosophical questions which can nettle one's understanding of it). Of course "belief" could be defined in such a way as to preclude the possibility of self-deception, but philosophers who have done so have paid the price of implausibility. In the history of epistemology belief is sometimes artificially restricted to an ideal philosophical notion where people never believe contradictory propositions - which might better be termed "rational belief."[\[85\]](#) This will hardly do as an account of belief itself, for human nature is capable of more things and stranger than common-sense philosophers suppose or than rationalistic philosophers impose on the world in Procrustean fashion. One has a far smaller opportunity to rid the world of irrationality if he takes the short-cut of defining unreasonable or incoherent thinking out of existence. Accordingly, I would suggest that the adequacy of one's conception of belief and of one's conception of self-deception will probably need to be judged jointly. To give a satisfactory account of one while being untrue to the other is to fail to do justice to the full range of human reality.

The term 'believe' has received analysis as a "parenthetical verb," a performative utterance, an expression denoting an occurrent mental event or denoting a personal disposition to act in certain ways under certain conditions.<sup>[86]</sup> Each analysis has its advantages and drawbacks, and in the end we are probably unable to provide a genuine "analysis" of belief just because it appears to be a notion which is primitive or fundamental in the explanation of the wide range of concepts in epistemology and philosophy of mind. Belief cannot be traditionally defined in terms of anything more basic than itself. Nevertheless, nothing prevents us from offering a general characterization of the ordinary notion of belief (without claiming completeness).

Belief is a positive, intellectual, propositional attitude which is expressed in a large variety of symptoms (some of which are subject to degrees of strength). To believe something is to have a favorable attitude toward a proposition - an attitude of the intellectual (rather than merely conative or affectional) kind. It is to take the proposition as true in a virtually automatic response to the evidence as it is perceived by the person. Thus to believe *p* is to see it as evidenced, to regard *p* as reliable. In the sense that belief is controlled by and informed by the way evidence is construed by the believer, belief is often said to be "constrained" - and some propositions are popularly said to be "beyond belief." Even seemingly unreasonable beliefs (cf. "blind faith") will turn out upon exploration to rest on something which is regarded by the believer anyway as a warrant, calling for the belief in question. Although belief is a positive propositional attitude informed by the evidence, that evidence can (and often is) misconstrued, misperceived, and approached with myopia of mind and senses. On this characterization, belief by no means precludes believing false propositions.

We can attempt a more precise characterization of belief here, one which with a modicum of judicious philosophical industry can survive whatever problems may remain to be worked out elsewhere.<sup>[87]</sup> The proposed way of speaking of belief shows initial plausibility, has been defended by respected scholars, and is bolstered by our common understanding of the concept of belief (even though it may not be a completely systematic account or analysis). At base belief is an action-guiding state of mind; it is a map-like mental state that is a potential cause of particular action (mental, verbal, or bodily). Specifically, belief is a persisting, intentional, mental state (made up of ideas which give a determinate character to the state corresponding to the proposition believed) with a stimulus-independent causal capacity to affect or guide one's theoretical and practical behavior, under suitable circumstances, in a wide variety of manifestations. In what follows, then, the expression "S believes that *p*" will be understood as true if and only if S relies upon *p* (sometimes, intermittently, or continuously) in his theoretical inferences and/or practical actions and plans.<sup>[88]</sup>

The grounds for saying that someone is self-deceived will coincide with or include the grounds for saying that he believes some proposition. If S did not take *p* as evidenced - that is, if S did not have a positive attitude or mental state such that *p* was relied upon in his theoretical or practical inferences - then we could not distinguish self-deception from mere ignorance of, or dislike for, *p*. It is just because S unavoidably looks upon some evidence as supporting *p* - and is thereby in the mental state of relying upon *p* in his inferences (practical and/or theoretical) - that his desire to avoid or manipulate that evidence in "self-deception" is meaningful. S does not wish to have his mind "in-formed" by the evidence in this fashion; he does not want to believe what he does believe. He would rather forget or hide the unpleasant truth that has gripped him, that is, to make covert

that he relies upon p in his theoretical inferences and/or practical actions and plans. His negative emotional response to p leads him to try and escape his uncontrived way of seeing things.

There are certain further points regarding belief about which we should make special mention. First, the bases for ascribing a belief to someone (the marks by which we discern a belief) are provided by both occurrent and dispositional accounts of belief. We consider the person's outward assertion of p (or inward, if oneself), and the way in which he behaves, reasons, gestures, feels, etc.; we take into account his decisions, emotions, habits, and even inaction. Of course neither a person's actions nor his utterances are infallible signs of belief, but they do offer fairly reliable correlations. The various kinds of indicators for belief should be used to supplement and qualify each other. One's own avowals of belief have a presumptive authority in determining what he believes, but those avowals can be defeated by cautious and relatively thorough observation of his other behavioral indicators. To put it simply: over time, actions will speak louder than words.

Second, not all of our beliefs are formed consciously, rationally, and with the giving of internal or external assent. To give assent to a proposition is explicitly to spell out (inwardly or outwardly) how one stands in respect to that proposition, thereby bringing one's belief to a conscious level of experience. However, there is no special logical or conceptual connection between beliefs and their linguistic expression. Holding a belief is not logically dependent upon a willingness or competence to express that belief verbally to oneself or others. Assent is not necessary to the mental state of belief. The cognitive and affective aspects of belief can sometimes be separated in a person and even be at odds with each other (e.g., hoping for what cannot be, fearing what you know does not hurt, failing to feel conviction in the face of strong proof). Accordingly we can easily imagine situations where most of the affective manifestations of a belief that p occur in S, and yet S does not assent to p, even when the proposition is attended to in his mind. He does not notice that his actions, emotions, assumptions, inferences, etc. are such as would be expected symptoms of someone who accepts p. It is a false picture we entertain of intelligent beings if we think of them as incessantly talking to themselves internally and always making explicit (or reporting on) their mental states and acts. A person's condition can be quite obviously belief-like, even when the (usual) assent-symptom of belief is absent; most, if not all, of the other symptoms of belief are evident. His behavior can hardly be explained without postulating in him a belief that p. It would be an artificial imposition to erect a terminological rule at this point, prohibiting us from saying that "S believes p" under such circumstances.

That would only screen off the complexity of human nature and behavior from us. We can certainly imagine, if we have not actually encountered, people who would protest that they do not hold beliefs about the inferior human dignity of people from other races - and yet who evidence just such an attitude in their social behavior nonetheless. The fact that belief can be divorced from explicit assent shows us, then, that there can be beliefs held by a person of which he is not aware - not consciously entertaining in his mind by introspection. A person can rely upon a proposition in his theoretical inferences and/or practical plans (e.g., "There is sufficient gas in the car's tank") without entertaining that proposition in mind; the proposition may not come to mind until something goes wrong (e.g., when he ends up stranded down the road). When I am surprised by meeting my previously vacationing neighbor at the mall, it is hardly because I had consciously inferred or entertained the proposition that he would not yet be back from his travels. The fact is that our set of beliefs is expanded and diminished throughout our waking moments (through sense

experience, casual reflection, etc.), and thus beliefs can be adopted without concentrating on the adoption procedure or even being aware of its results. Furthermore, it is quite clear that not everything that a person believes can be simultaneously attended to by him in thought. We must conclude that introspection and assent do not invariably accompany a person's each and every mental state or action.

Third, we must add that self-ascriptions of belief by way of assent - just like disavowals of belief - are not incorrigible (i.e., there can be overriding reasons to think them false) and therefore not infallible (i.e., such reports can be mistaken). A person can be held to believe something from which he dissents, and can be found not to believe something to which he assents. To some appreciable extent we can be mistaken about our own beliefs. This may seem surprising, but there are after all limits on our self-knowledge, even though our own reports about our beliefs (or pains, or perceptions, etc.) have a presumptive authority and are granted a degree of accuracy.

We have seen that normally first-person, present-tense, occurrent mental state beliefs are direct, far more reliable than the counterpart beliefs about others, excellent evidence for the presence of the states they "report," ... but they are like our beliefs about others in being fallible, dubitable, corrigible, and testable.[\[89\]](#)



People may have the best word on what they believe, but they do not logically have the last word (as in the example of racial prejudice above). It is not hard to find examples in ordinary experience of someone believing something, but yet withholding, avoiding or suppressing internal and external assent to it. We also have ready examples of someone believing that he believes something, although in fact he does not believe it. Such examples can only be explained away or recategorized by the *ex post facto* imposition of artificial conditions upon what we call "belief." People can and do sometimes come to realize, on the evidence in their behavior, that their previous avowals (or disavowals) of a belief were mistaken.

Fourth, the last thing about belief which calls for special mention is its voluntariness. This may seem strange since we have above spoken of belief as a propositional attitude which is "constrained" by the evidence as seen by the person in question. The seeing of the evidence as this or that - the taking of it in a particular way - constrains one to believe as he does. Since I see myself as right-handed, I cannot voluntarily and on the spot believe (genuinely) that I am left-handed. Nobody can believe contrary to the way in which he sees the evidence, to be sure. However, one can exercise some control over the way in which he sees that evidence - directing his attention, giving prominence to some matters over others, suppressing what he does not wish to encounter, re-evaluating the significance of past considerations, etc. If belief is like "seeing-as," then we must also recognize that seeing-as is somewhat subject to one's will. A person is free to ignore the grounds for a belief, in which case that belief is not compelled (in an absolute sense) after all. A person cannot choose voluntarily and arbitrarily to believe whatever he wishes, but he can nevertheless freely doubt propositions, suspend judgment about them, voluntarily inhibit extending inferences based on them, etc. Directing our thoughts is a kind of doing, and by the directing of our attention we can encourage or thwart our propensity to believe things. People are thus free to fortify or undermine beliefs they have by voluntarily concentrating on certain lines of evidence, ignoring others, misconstruing yet others, etc. In such ways we can deliberately cultivate a belief (whether about some matter or about ourselves and our beliefs) which turns out contrary to the facts.

Everyone knows the experience of weighing or deliberating about the options and then "taking the plunge" of assenting to one over the other. We ordinarily take responsibility - and are held responsible - for our beliefs. They are assessed as though we had some control over them; our beliefs are evaluated as more or less reasonable, justifiable, and even moral. We at times hear people declare "I cannot believe that" (e.g., a close relative has been convicted of a heinous crime), but we all realize that the "cannot" here should be interpreted as "will not" - because one does not want it to be true, cannot emotionally afford to admit it, thinks it is his duty to resist it, or lacks the intellectual energy to rise to the occasion. In many ways, then, we recognize the voluntary aspect of belief.

Given the preceding explanation of belief as such, and with the salient features of belief just enumerated in mind, we can proceed to explicate a non-paradoxical account of self-deception.

## Incompatible Beliefs, Motivated Rationalization, and Self-Covering Intention

We should maintain the appropriateness of modeling self-deception on other-deception, contending that there is a common sense for the word "deception" in both cases. This does not commit us to going to the extreme of making self-deception a literal case of other-deception (the same in every detail), as though we were dealing with a split personality. Rather self-deception should be seen as a general parallel to other-deception in certain specifiable ways. For instance, elements of deception which are shared by both self-deception and other-deception are the deceiver's responsibility for causing the deceived to believe falsely, the deceived holds (at least implicitly) an erroneous belief about the deceiver's beliefs, and the rationalization maneuvers taken in the face of evidence brought to the attention of the deceiver by others.

Given the other-deception model, incompatible beliefs need to be attributed to the self-deceiver on the basis of his behavior. Self-deception is a conflict state in which S holds incompatible beliefs, but the nature of this incompatibility needs to be noted. The self-deceived person holds a first-order belief (viz., that p) which is not a matter of personal indifference to himself, but somehow distressing; he has a personal stake in (or against) p. Thus it is a special kind of belief: one which S dreads, cannot face up to, or wishes were otherwise since it brings some unpleasant truth before him. Accordingly, S brings himself to deny that belief - not only to deny p (about the distressing issue in question) but more significantly to deny something about himself (namely, his believing p). Thus the analysis of self-deception involves reference to iterated beliefs (i.e., beliefs about one's beliefs). While believing p, S comes to hold additionally a (false) second-order belief about that belief - namely, that S does not believe p. A person may believe that dogs are dangerous (first-order), and may also believe (second-order) that this belief concerning dogs is quite reasonable. A person may believe (first-order) that members of other races are inferior and yet (second-order) believe about himself that he does not believe in racial inferiority.[\[90\]](#)

It is important to note that the behavioral symptoms of believing p overlap extensively with the behavioral symptoms of believing that you believe p. In the examination of one's actions, emotions, words, etc. it will be found that they can easily be taken as indicators of both the first-order and the second-order belief. Likewise, the behavioral indicators for S not believing p readily shade back and forth into the behavioral indicators for S believing (about himself) that he does not believe p. A man who believes that dogs are dangerous engages in most of the same inferences, reactions, emotions and behavior as a man who believes that he believes dogs are dangerous. This helps us to understand that the nature of the incompatibility of beliefs in self-deception is not logical in nature, but behavioral and practical. The first-order and second-order beliefs are not formally contradictory, but the inferential and behavioral effects of the two beliefs are in conflict with each other. The self-deceiver believes something (which causes him distress) and gives evidence of believing it; however, he brings himself to believe that he does not believe it (which brings a measure of relief) and gives evidence that he does not think of himself as believing it. S believes p, but his assent to it is blocked by acquiring the (false) second-order belief that S does not believe p. The incompatibility between these two beliefs is thus practical in nature. They call for conflicting kinds of intellectual, verbal, and behavioral responses.

Now S has an obvious interest at stake in maintaining the rationality of his second-order belief (which brings him into a conflict state with his first-order belief). This analysis of self-deception

holds that it comes about when, in the face of evidence adverse to his cherished second-order belief (about himself), S engages in contrived and pseudo-rational treatment of the evidence. That is, he manipulates, suppresses, and rationalizes the evidence so as to support a belief which is incompatible with his believing that p. He ignores the obvious, focuses away from undesirable indicators, twists the significance of evidence, goes to extreme measures to enforce his policy of hiding his belief that p from himself and others. If he looked at himself as others see him, he would have all the evidence he needs to conclude that he believes that p, but he strains and strains to convince himself that he does not believe that p.

This rationalizing activity, in order to count as self-deception and not something else (e.g., a cavalier disagreement), must be given a motivational explanation. S distorts the evidence in order to satisfy a desire - namely, the desire to avoid the discomfort, distress, or pain associated with believing that p. By means of it he enters into and maintains self-deception, believing that he does not believe that p. Actions or reactions which have the effect of achieving the special state of incompatible beliefs traced above are referred