Not to Know What One Knows: Some Paradoxes of Self-Deception

Jean-Pierre Dupuy

The problem of lying to, or deceiving oneself is currently one of the most debated in analytical philosophy. Now, since analytical philosophers are aware that Sartre defined "bad faith" as lying to oneself, as self-deception, and since moreover they find relatively coherent *arguments* in Sartre's text, they do not hesitate to include these arguments in their debates, if only to contest them. "To be dead is to be a prey for the living," one reads in *Being and Nothingness** (p. 695). One imagines Sartre rolling over in his grave. For this philosophy of mind is truly the Other of Sartre's philosophy. Yet, at the price of a treacherous translation, this philosophy gets something from Sartre, and perhaps gives him something in return.

In a slightly surreal, perhaps even monstrous way, I am going to make the two philosophies engage in a dialogue on the problem of lying to oneself.

1. Between Lying and Bad Faith

Sartre agrees to treat bad faith as lying to oneself. However, he immediately adds that a distinction must be made between the kind of lying involved in lying to oneself and plain lying. The latter, at least in its ideal-type, does not seem to him to require any "special ontological foundation." This is not the case with lying to oneself. If it were to be understood in the mode of lying to others, one would come up against a paradox: when I lie to myself, "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

---

*Translated by Hazel Barnes, New York, 1956. All subsequent page numbers in brackets refer to this edition.

53

truth very 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 a semblance of duality—but in the unitary structure of a single project." (p. 89) "To escape from these difficulties," Sartre notes, "people gladly have recourse to the unconscious." (p. 90) We know that Sartre will reject the Freudian evasion by attempting to show that it only moves the paradox to the level of the censor, which is also supposed to at once know and not know the truth to be repressed. His solution, as is well known, is provided by the self-transcendent structure of consciousness, or "human reality."

Sartre recognizes that there are "intermediaries between false-hood and bad faith." However, he considers them nothing more than "degenerate," "common," "popular" forms of the lie.

North American work on self-deception is unanimous in its criticism of Sartre on this point. The line drawn between lying and lying to oneself is much too hard: it is precisely in the "intermediate" zones that interesting things occur. This criticism is fundamental because the foundational case of a unique consciousness is essential to Sartre's theory of bad faith, and thus to his theory of consciousness: "bad faith does not come from outside to human reality … consciousness affects itself with bad faith." (p. 89) Showing that the division between lying to others and lying to oneself does not exist would equally be the introduction of the *mit-sein* into the structure of bad faith—but perhaps also into that of consciousness.

Enter Donald Davidson, the American philosopher whose writings on self-deception have perhaps been the most influential. He considers a man embarrassed by growing baldness who manages through various cosmetic, and especially psychological, means to deny the obvious.[1] At one and the same time, this man believes that he is bald *and* he believes that he is not bald; he manages to hide the fact that he has the first belief from himself because he wants to "see" only the second. Yet it is indeed *because* he has the first belief that a mental mechanism of the kind we call *wishful thinking* is set up and he begins to have the second: the first belief is truly the cause of the second, without, obviously, being a reason for it, since they are contradictory.

Davidson considers lying to oneself to be a particularly "hard" case of self-deception and he does not even attempt to salvage its

54

possibility. Regarding "softer" self-deception, situated somewhere, as regards its ontological implications, between lying and lying to oneself, the characteristic coexistence of two contradictory beliefs is ensured by a compartmentalization or partitioning of the mind. One might say that Sartre pre-refuted Davidson since this notion is none other than that of Freud. However, I do not want to pursue this further here.

While Davidson and Sartre differ in their analyses of bad faith, they are close to each other in their treatment of lying. Sartre notes that there is "no difficulty in holding that the liar must make the project of the lie in entire clarity and that he must possess a complete comprehension of the lie and of the truth which he is altering. It is sufficient that an over-all opacity *hide his intentions* from *the Other*[2] ... By the lie consciousness affirms that it exists by nature as *hidden from the Other*, it uses for its own profit the ontological duality of myself and the self of the Other." (pp. 88-89) Davidson refines the analysis, in the style of a philosopher of mind. Indeed, everything hangs on the play of the intentions, but Davidson distinguishes two levels. First, there is the intention to deceive; but there is also the intention to hide this intention to deceive which is, according to Davidson, the key to lying—and not, for example, the fact one says the contrary of what one knows to be true. He notes that "a liar who believes that his hearer is perverse may say the opposite of what he intends his hearer to believe."[3] One thinks of the Jewish joke reported by Freud and which Lacan so enjoyed: "Why tell me you are going to Krakow in order to make me believe you are going to Lemberg when in fact you are going to Krakow!"

Sartre cannot but be in agreement with this, and he would add that in the case of lying, contrary to that of lying to oneself, the realization of this meta-intention, which is the intention to hide *from the other* that one is attempting to deceive, poses no problem. Moreover, there is no need to hide *from oneself* that one is attempting to deceive the other: "The liar intends to deceive and he does not seek to hide this intention from himself ... As for his flaunted intention of telling the truth ("I'd never want to deceive you! This is true! I swear it!")—all this, of course, is the object of an inner negation, but also it is not recognized by the liar as *his* intention." (p. 88)

55

At this point I would like to bring in an approach to the problem which is in the spirit of the work of what the French call the "Palo Alto School." Its source is Mark Anspach, a young American anthropologist established in France, who, modelling himself on Gregory Bateson, applies his experience in anthropology to a study of mental illness. Anspach asks us to consider the case in which it is the liar himself who is "perverse." This liar has the good taste to warn others that what he says is contrary to what they should believe. "Beware, I want to deceive you, what I say is false, I swear it, etc."[4] Here, the intention to deceive is unveiled, even proclaimed, instead of being hidden. This is a case of lying to others which is paradoxical, and which is so in the same way as lying to oneself. In effect, he who lies to himself reveals to himself, as a liar with access to his intention to deceive, that he has the intention to deceive himself. If I lie to myself, I say to myself: "I am lying to myself, beware, self!: what I say is false." Seen this way, lying to oneself has the familiar form of the liar's paradox.

Anspach illustrates this thesis using cases of psychotics. He has noticed, in Marcel Gauchet and Gladys Swain's valuable work, *La pratique de l'esprit humain*[5], certain observations made by Pinel's disciple Étienne Esquirol which are grist to the Palo Alto School's mill. Interpreted as a liar's paradox, the paradox of self-deception, far from being insoluble due to the contradiction it implies, provides the solution to this problem by its very paradoxical form (the "double bind"). Esquirol explains that a patient may "very well know the disorder of his intellectual faculties," try to follow and to believe what his therapist tells him, and yet "lack the force of conviction." "I know all that," says one of them, " but my idea is there and I am not cured."[6]

Another patient told him one day: "If I could believe with you that I am mad, I would soon be cured, but I cannot acquire this belief." This is an extraordinary proposition which Anspach analyzes very incisively. Literally, the patient believes he is insane to not believe he is insane. It is because he does not believe he is insane that he is insane—he believes. Inversely, however, it is because he is insane that he does not believe he is insane— he believes.

56

This differs from the Davidsonian description of self-deception in two ways. Here we are no longer dealing with the coexistence of two contradictory beliefs (the subject believing p and believing not-p), but with the co-determination of a belief and a non-belief (the subject believing p and not believing p). Furthermore, causality is no longer linear (the first belief causes the second), it has become circular: the belief is the cause of the non-belief, and vice versa. This figure is strictly paradoxical, which does not necessarily mean that it is to be rejected because it is unthinkable, but simply that it has the form of a paradox. Mark Anspach proposes visualizing it using a geometrical metaphor. What prevents the belief and the non-belief from coming into contact with each other in the mind of our insane person is not a partition as in the Davidsonian model, but a strip ... a Möbius strip. By following this strip, one passes from the belief to the non-belief, but these "mental states" nonetheless remain on opposite sides of the strip which separates them even as it links them together.

This dialectical play between belief and non-belief is in the end much closer to Sartre's theory of belief than to that of philosophy of mind, which is dominated by an "intentional realism" (a doctrine which holds that mental states in general, and beliefs in particular, have an ontological reality). Let us recall the brilliant analyses in Part Three of the chapter on bad faith, titled "The 'Faith' of Bad Faith." There we find: "belief is a being which questions its own being, which can realize itself only in its destruction, which can manifest itself to itself only by denying itself. It is a being for which to be is to appear and to appear is to deny itself. To believe is not-to-believe." And also: "To believe is to know that one believes, and to know that one believes is no longer to believe. Thus to believe is not to believe any longer because it is only to believe ... ." (p. 114) To say that "belief becomes non-belief" (p. 115) is our very paradox, and for Sartre, it is to say that "consciousness is perpetually escaping itself" (*ibid*). Like bad faith, belief is "evanescent." (p. 90)

Nonetheless, the Palo Alto School's definition differs from Sartre's on a fundamental point. For the former, lying, in lying to oneself, is not essentially different from lying to others. Certainly, this does result in the very paradox Sartre wants to avoid, but here this paradox is treated as a solution, not as a problem.

57

2. About Social Lying

We will now approach the question from a completely different angle: that of pragmatics, or the analysis of context in verbal communication. In his pioneering article of 1957, "Meaning," Paul Grice showed that what makes communication possible is the hearer's ability to recognize the speaker's intention to inform him of something. In the reformulation recently proposed by Sperber and Wilson[7], communication is defined as the production by the speaker of a certain stimulus with the double intention:

a) to inform the hearer of something;
b) to inform the hearer of his intention to inform him of something.

The distinctive feature of communication is found in intention b), which is the true communicative intention, while intention a) is simply an informative intention. A communicative intention presents itself as a second-order informative intention since its realization implies that a first-order informative intention is recognized by the hearer. This reflexiveness appears inherent to the act of communication.

Very early on, theorists in pragmatics recognized that one could not stop at this doubling, and that truly reflexive communication would imply an infinity of interlocking intentional levels. Let us consider the following counter-example.[8] Peggy's hair dryer is broken. She would like John to repair it, but she does not want to ask him to do so directly. She imagines the following scene. She takes apart the hair dryer and scatters the pieces around her, as if she were fixing it herself, but she arranges it so that John realizes that this is just an act. Her intention is to inform John that she would like his help, and, moreover, she attempts to make it manifest to John that she has this intention to so inform him. Peggy thus has a double intention, informative and communicative, and this double intention is, let us suppose, realized. According to the definition given above, one should thus accept that Peggy has *communicated* her request for help to John. Yet it is difficult to place this twisted informative strategy on the same level as a transparent communication by which Peggy would ask John to help her

directly. The difference rests in this: what makes Peggy's ploy opaque is that her second-order intention, her communicative intention, remains unknown to John. Peggy hides her communicative intention from John.

Note that what is hidden is not, as in the case of lying, an intention to deceive. *To the contrary, it is an intention to communicate.*

This situation is thus not a case of lying. What is it then? First, let us attempt to analyze Peggy's motives. She has a certain type of relationship with John. Let us say that she is a modern woman and that the two things she detests most are being rejected by John and being indebted to him. If she were to address him openly, she would run this double risk. Her staging allows her to avoid this completely. If John helps her, he does so of his own initiative: Peggy, who has asked nothing of him, owes him nothing. However, John could very well do nothing without appearing to be a boor: after all, he is not supposed to have interpreted Peggy's ploy as a request for help. This, at least, is what he believes, for he believes that Peggy does not know that he has understood that her intention was indeed to ask him for help. Peggy has given him this escape hatch: not helping would not be an unpleasant refusal, but a simple lack of attention.[9]

If this is neither lying nor, obviously, self-deception, what is it? One might speak of a kind of negative collaboration between two beings who accept, because it is convenient for them, a form of collective opacity. This could be described very precisely thanks to a concept which today plays an essential role in a whole range of disciplines, from game theory to artificial intelligence, from the philosophy of language to analytical political philosophy. This is the concept of Common Knowledge (CK). A proposition is CK in a given population if and only if this proposition is true; everyone knows it is the case; everyone knows that everyone knows it is the case; etc., *to infinity*. Truly transparent communication can then be defined as the one in which the speaker intends to make his informative intention CK between him and his hearer. The opacity which describes Peggy's ploy can then be defined as a deviation from CK.

If we go back, now, to the other side of the Atlantic, we will discover to our surprise that this configuration, described by shared

knowledge—everyone knows *p*—and an absence of CK, is not only well known, it is treated as a particular form of lying to one-self. Simply, if one can so say, the author and the victim of the lie in question are in this case the collective itself. I am referring to the notion of the social lie, or collective hypocrisy, which has played an essential role in French social sciences, as much in Durkheimian sociology as in the structuralism which has dethroned it. The example I shall take is the debate on the reciprocity of mutual exchange, which was one of the major controversies in French-style human sciences.

In his famous work *Essai sur le don* (1924)[10], Marcel Mauss notes that in a good number of archaic societies, "contracts are fulfilled and exchanges of goods are made by means of gifts. In theory such gifts are voluntary but in fact they are given and repaid under obligation." He insists the prestations have a "voluntary character, so to speak, apparently free and without cost, and yet constrained and interested. ... They are endowed nearly always with the form of a present, of a gift generously offered even when in the gesture which accompanies the transaction there is only a fiction, formalism and *social deception*, and when there is, at bottom, obligation and economic interest."

Separate acts: giving, receiving, returning, present themselves as so many gestures of generosity or cordiality, yet in fact they obey strict, inescapable imperatives. What then is the nature of this "obligation"? Once he has asked this question, Mauss adds, as if he were only repeating it in another form: "What force is there in the thing given which compels the recipient to make a return?" The native informant will rapidly convince him that "in the things exchanged ... there is a certain power which forces them to circulate, to be given away and repaid."

In his equally famous *Introduction à l'oeuvre de Marcel Mauss* (1950)[11]—a text which many consider to be the charter of French structuralism—Lévi-Strauss reproaches Mauss for allowing himself here to be "mystified by the native." Mauss's mistake, according to him, was to have remained at phenomenological apprehension, which breaks exchange down into its different moments. This thus creates the need for an operator of integration to reconstruct the whole, and it is precisely the "soul of things" which providentially

comes in to play this role. However, this is tackling the problem from the wrong end, Lévi-Strauss asserts, because "Exchange is not a complex edifice, constructed from the obligations to give, to receive and to make return with the help of an emotional and mystical cement. It is a *synthesis immediately given* to, and by, symbolic thought. … " The "underlying reality" of the exchange, he explains, is to be found in "*unconscious* mental structures," to which language can provide access.

The third step: in 1972, Pierre Bourdieu, in his *Esquisse d'une théorie de la pratique*[12], denounced the "objectivist error" of Lévi-Strauss: "Even if reciprocity is the objective truth of the discrete acts which ordinary experience knows in discrete form and calls gift exchanges, *it is not the whole truth of a practice which could not exist if it were consciously perceived in accordance with the model.*"

In effect, consider the obligation to make a return for what is received and the obligation to receive. Taken together in the theoretical schema of reciprocity, they lead to a contradiction. He who immediately returns the very object he is given refuses, in fact, to receive. The exchange of gifts can only function as such on the condition the reciprocity which would be its objective truth is *hidden*. All the space, or rather the time, of practice is needed to undo this contradiction.

Thus in Bourdieu's interpretation (as in Mauss's) there is a lie here. The natives know the truth of reciprocity, but they hide it, for this truth is lethal. From whom do they hide it? From themselves, of course. Note that the paradox is described by Bourdieu in the same terms as those used by Sartre. As we have seen, the latter remarks about lying to oneself: "I must know in my capacity as deceiver the truth which is hidden from me in my capacity as the one deceived (…) and this not at two different moments, which at a pinch would allow us to re-establish a semblance of duality—but *in the unitary structure of a single project* (…) To escape from these difficulties people gladly have recourse to the unconscious." (pp. 89-90) Lévi-Strauss resorts to the unconscious, Bourdieu to the deployment of temporality (the time of the practice), but in both cases, the subject of which one speaks, and to whom one ascribes in one case the unconscious and in the other case bad faith, is a non-subject since it is the structure or the collective. This complication is apparently

insurmountable, for Lévi-Strauss in any case, whom we leave with the collective unconsciousness structured like a language.

Bourdieu's view, however, suddenly becomes much clearer when, during an analysis, he takes the example of a Kabyle worker who proclaimed the convertibility of the meal traditionally given at the end of work into money, with which he demanded to be paid instead. Bourdieu writes that here this worker was only "betraying *the best-kept and the worst-kept secret: the one that is in everyone's keeping.*"

Undoubtably a brilliant formula—we also say: "the secret is there is no secret"; Zinoviev uses the oxymoron "public secret," etc.—but it says nothing but what analytical philosophy describes as a situation with shared knowledge but without CK. Thus, social hypocrisy, collective bad faith, would be the very thing we described above as a negative collaboration between individual subjects who mean to protect a collective opacity which is convenient for all.

This picture of a state of things which is shared knowledge without being CK is in no way paradoxical, it does not escape logical analysis. The detour via the collective, which could seem to introduce a formidable increase in complexity, has perhaps put us on the right track. What if, when an individual consciousness lies to itself, there were not always this negative collaboration with another? What if the operator of reflexiveness, the self of self-deception, was Alter Ego in me? Unfortunately, we cannot pursue this line further here.

3. Choosing One's Own Past

Analytical philosophy of action has no trouble identifying with Sartre when he defines freedom in terms of choice, decision, and commitment. In spite of their true goodwill, there is a point, however, at which analytical philosophers get left behind. Too much is too much. When they read expressions such as "I am responsible for everything," (p. 710) and "the peculiar character of human-reality is that it is without excuse," (p. 709) they throw up their hands. The limit is reached when Sartre extends the range of freedom to the past, as in the following passage: "In order for us to 'have' a past, it is necessary that we maintain it in existence by our very project toward the future; we do not receive our past, but the

62

necessity of our contingency implies that *we are not able not to choose it*." (p. 639; my emphasis.)

Alain Renaut has made clear all that this Sartrian theme owes to Heidegger's analysis of the historicity of "human-reality" (*Dasein*). Renaut writes that for the author of *Sein und Seit*, "'human-reality' is first and foremost historical in that its essential property is to *choose* what on the other hand seems to it to be *destiny*"[13]; and furthermore, "what we call 'destiny' is thus the 'resolute-choice' (*Entschlossenheit*) of 'human-reality.'"[14] Sartre expresses this: "To be finite, in fact, is to choose oneself—that is, to make known to oneself what one is by projecting oneself toward one possible to the exclusion of others." (p. 698)

With this philosophical configuration, one could believe that one is at the opposite extreme of what analytical philosophy could hold. This would be a huge mistake. It so happens that rational choice theory, taking up the problems of the antinomies of reason, has developed a very interesting paradoxology, which includes, among other marvels, a paradox which one could describe using the expression: "*to choose one's predestination*." This paradox is at the origin of a veritable schism within decision theory.[15] Furthermore, the issue of self-deception finds itself playing a decisive role here.

The incarnation of this paradox which I will discuss is Max Weber's famous thesis on the "correlations" between the "Protestant ethic," or more exactly the ethical consequences of the doctrine of predestination, and the "spirit of capitalism."[16] I am interested only in the logical structure of Weber's argument, not in its empirical validity. In virtue of a divine decision taken for all eternity, each person belongs to a group, that of the elect or that of the damned, without knowing which. There is no way to affect this decree, nothing one can do to earn or merit salvation. Divine grace, however, manifests itself through *signs*. What is important is that these signs cannot be observed through introspection: they are acquired through action. The main sign is the success one obtains by putting one's faith to the test in a professional activity (*Beruf*). This test is costly. It requires one work ceaselessly, methodically, without ever resting secure in, without ever enjoying, one's wealth. "Unwillingness to work," Weber notes, "is symptomatic of the lack of grace." (p. 159)

63

The "logical consequence" of this practical problem, Weber notes again, "obviously" should have been "fatalism." Fatalism, in other words the choice of an idle life, is in effect the rational solution since, *whatever the state of the world*—here, that one is of the elect or of the damned—one has nothing to win by engaging in the costly test of professional commitment. In decision theory, one is said to be dealing with a "dominant" strategy, in the sense that it is the best one in each possible case. Weber's whole book however attempts, as we know, to explain why and how "the broad mass of ordinary men" has made the opposite choice.

The popular Calvinist doctrine held it "to be an absolute duty to consider oneself chosen, and to combat all doubts as temptations of the devil, since lack of self-confidence is the result of insufficient faith, hence of imperfect grace."(p.111) "Intense worldly activity" was what allowed one to obtain this self-confidence, the means to assure oneself of one's state of grace.

The debate which pitted Lutherans against Calvinists is still extremely interesting. The former accused the latter of holding views which amounted to the doctrine of "salvation by works." This greatly distressed the Calvinists, who were outraged that one could identify their doctrine with the doctrine they most scorned: that of the Catholics. This accusation boiled down to saying that he who has chosen to pay the full price for the signs of grace reasons *as if* these signs were the cause of his salvation—this is magical behaviour, insist the accusors, since it consists in taking the sign for the thing (divine election). Now, this accusation is none other than that which partisans of orthodox decision theory—those who, when faced with a problem with this structure, defend the dominant strategy—make of their current adversaries, the heterodox theorists who defend the rationality of the Calvinist choice.

Weber dubs these Calvinists "saints overflowing with self-confidence" and "Self-Proclaimed Saints." At this point, philosophy of action asks: were they also in bad faith, were they lying to themselves?

Orthodox decision theorists answer yes. Their argument can be schematized in the following way. Propositions (1) and (2), applied to the situation under study, are both true:

64

(1) The Calvinists believe they have placed themselves among the elect by choosing to acquire the signs of grace;

(2) the Calvinists believe they have not placed themselves among the elect.

(1) and (2) express contradictory beliefs. Furthermore, one can suppose that:

• The Calvinists find a way to hide (1) from themselves
• because they want to believe they were chosen by God.

If one further postulates that the first belief is the *cause* of the second, without obviously being able to constitute a reason for it, one obtains a case of self-deception, as described by Donald Davidson.

There is no denying that this is an acceptable interpretation of the Calvinist choice. I would simply like to propose another interpretation which has the effect of revealing the *rationality* of the Calvinist choice. In effect, I entered this debate fashioning my approach after the work of the analytical theologian Alvin Plantinga.[17]

This other interpretation attributes the Calvinists with the two following beliefs, which are not (necessarily) incompatible:

(3) The Calvinists believe they have not placed themselves among the elect because they believe God has chosen them;

(4) The Calvinists believe they were free to make the opposite choice when they chose.

Under the orthodox interpretation, the irrationality is found in proposition (2): the Calvinists set themselves on the belief that they did not place themselves among the elect because "at the bottom of their hearts," they know very well that they have acted to give themselves the signs of their election and that they want to hide this truth from themselves. According to the heterodox interpretation, the Calvinists believe that they have not proclaimed themselves saints simply because they take seriously the givens of the problem as they were submitted to them, or as they have internalized them: God has proclaimed them to be as they are. Nonetheless, they must face a serious problem: they must consider it not incoherent to believe at once that God has chosen for them (proposition [3]) and that they were free to choose (proposition [4]). In other words, in order for them and us to be able to take

Max Weber's problem seriously, we must first convince ourselves that it is reasonable to be a *"compatibilist"*: to believe in the compatibility of (here, causal) determinism and free will.

I am obliged here to pass directly to the conclusion of a complex analysis. Being a compatibilist implies reasoning thus. Just as "when Adam took the apple it would have been *possible* for him not to take it" (p. 602), when the Calvinist makes the Calvinist choice, it would have been possible for him to make the opposite choice. Just as there would then have been another Adam, there would have been another Calvinist: instead of being chosen, he would have been damned. The example of Adam and the apple is, as we recall, the one that Sartre takes in order to distinguish his position from that of Leibniz. According to the latter, Adam's essence is not chosen by Adam, but by God. His freedom is thus only illusory. According to Sartre, to the contrary, Adam's existence precedes his essence. Free Adam chooses himself: his existence determines his essence, "henceforth what makes his *person* known to him is the future and not the past; he chooses to learn what he is by means of ends toward which he projects himself." (p. 603) Under the heterodox interpretation, the free Calvinist follows Leibniz and Sartre at the same time. His essence determines his existence, but, since he is free to choose his existence, he can determine his essence. He has, literally, the power to choose his predestination. However, as Plantinga insists, this power is not causal—which would make it inconceivable since causality would then fly counter to the arrow of time. It is a *"counterfactual* power over the past."

I have shown that the Calvinist choice, interpreted as free and rational, defined a temporality, if not to say a historicity, in which the past is interpreted according to the present choice. We are not very far from Sartre, and yet we are a world away since what allows this reversal with respect to the flow of physical phenomena—the domain of the "in-itself"—is a determinism combined with free will. To use the classic metaphor, everything is already "written." The subject acts according to a previously prepared scenario, but since he is free, he can raise himself to the level at which this scenario is written and exercise a form of power over it—the power which Plantinga calls counterfactual.

Calvinist bad faith and rationality have the same structure, and this structure shares many points with the structure Sartre sees as shared by consciousness and bad faith. Calvinist faith strangely resembles what Sartre calls "the faith of bad faith." (p. 112) The fact that bad faith is faith, in other words, belief, he asserts, is precisely what distinguishes it from lying. "How can we believe by bad faith in the concepts which we forge expressly to persuade ourselves?" he asks. (p. 112) This rhetorical question seems to fit the Calvinist choice like a glove. Sartre's answer is: "the project of bad faith must itself be in bad faith." With respect to the "disposition" I take to per-suade myself—and here one obviously thinks of the Calvinist *action*, performed *in order* to believe—he specifies: "For me to have represented it to myself as bad faith would have been cynicism; to believe it sincerely would have been in good faith." (pp. 112-113) Sartre is looking for an improbable intermediary position, of which the choice of the Calvinist, who believes himself to be neither entirely foreign to his election nor totally responsible for it, perfectly illustrates the coherence.

This parallel, I repeat, cannot be pushed too far. The heterodox position in rational choice theory results in these unusual con-figurations only thanks to determinism's helping hand. The Calvin-ist certainly endorses determinism, but it remains no less radically the Other in him. Sartre, however, intends to hold to what he postu-lated at the beginning of his analysis: "bad faith does not come from outside to human reality." (p. 84) Is this not the postulate which con-demns him to impotence? For finally, when he concludes with: "The decision to be in bad faith does not dare to speak its name; it believes itself and does not believe itself in bad faith" (p. 113), he only comes back to his point of departure, the paradox of the simultaneous pres-ence of the belief and the non-belief, the paradox which he none-theless aims to allow us to escape.

Notes

1. Donald Davidson, "Deception and Division," in Jon Elster (ed.), *The Multiple Self*, Cambridge, 1986.
2. Sartre's emphasis is in italics; except for 'hide his intentions' which is mine.

67

3. Donald Davidson "Deception and Division," loc. cit, p. 88.
4. Mark Anspach, "Madness and the Divided Self," communication at the Self-Deception Symposium, Stanford University, February 1993.
5. Paris, 1980.
6. Quotations from Gauchet and Swain by Anspach, loc. cit.
7. Dan Sperber and Deirdre Wilson, *Relevance*, 1986.
8. *Ibid.*, p. 30.
9. *Ibid.*, p. 62.
10. *Année Sociologique*, 2ème série, 1923-1924, vol. I; reprinted in *Sociologie et anthropologie*, PUF, 1973. *The Gift: Forms and functions of exchange in archaic societies*, trans. Ian Cunnison, New York: Norton, 1967.
11. In *Sociologie et anthropologie*, op. cit.
12. Droz, Geneva, *Outline of a Theory of Practice*, Cambridge, 1977.
13. Alain Renaut, *Sartre, le dernier philosophe*, Paris, 1993, p. 49.
14. *Ibid.*, p. 48.
15. See J.-P. Dupuy, "Temps et rationalité" in *Sciences sociales et sciences cognitives*, Cours de l'École polytechnique, 1993-1994.
16. Max Weber, *The Protestant Ethic and the Spirit of Capitalism*, New York, 1985.
17. J.-P. Dupuy, "Temps et rationalité," loc. cit.; Alvin Plantinga, "On Ockham's Way Out," in *Faith and Philosophy*, 3, 1986.