BAD FAITH AND 'LYING TO ONESELF'

In an earlier article,[1] I attempted to show some equivocations related to Sartre's claim that sincerity is a "phenomenon of bad faith" and shares its fundamental structure and goal. Although, in retrospect, I do not retract the thrust of my argument, I feel that I tended either to ignore or overlook some of the complexities and problems involved in Sartre's treatment of "bad faith." In the present article, I wish to attend to at least a few of these. While attempting to offer a close, integrative analysis of Sartre's views on bad faith, I intend to raise some basic questions concerning the adequacy of his position, as it develops. In particular—in the light of his initial characterization of bad faith—I shall raise the question as to whether, given his ontological premises and system of meanings, it is possible to lie to oneself. On Sartre's terms, of course, this question entails the question of the possibility of bad faith. Despite my persistent misgivings, I shall concede that there is, finally, a highly modified sense of "lying to oneself" in which bad faith as "lying to oneself" becomes a possibility for Sartre. But even this qualified sense, I shall show, is not without its problems.

Let us begin by reviewing Sartre's preliminary characterization of "bad faith" and elucidating the initial and gnawing question to which it gives rise.

1. *Lying in general and lying to oneself.* Sartre's discussion of bad faith follows both his discussion of the origin of negation and his introductory definition of human consciousness as "a being such that in its being, its being is in question in so far as this being implies a being other than itself" (p. 86).[2] The human being is not only the being through whom concrete negations (négatités) come into the world, but he is the being who can take negative attitudes towards himself. In order to illustrate this possibility of self-negation, he chooses to examine "one determined attitude . . . *essential to human reality* . . . which is such that consciousness instead of directing its

[1] Ronald E. Santoni, "Sartre on 'Sincerity': 'Bad Faith'? or Equivocation?," *The Personalist,* Spring 1972, pp. 150-160.

[2] Unless otherwise noted, all page references in this article are to Jean-Paul Sartre, *Being and Nothingness,* unabridged translation by Hazel E. Barnes, New York: Washington Square Press, 1969. To avoid confusion of pagination, special note should be taken of the 1969 printing. Any attempt to clarify by an appeal to the French will be done through *L'être et le neant,* Éditions Gallimard, 1943, and abbreviated by *EN* together with the page reference.

negation outward, turns it towards itself" (p. 87, italics mine). For Sartre this attitude is *bad faith.*

Sartre, in a preliminary move which remains to tantalize the rest of his discussion, allows that bad faith is a "lie to oneself" (p. 87). He does this on one proviso; namely, that lying to oneself be distinguished from "lying in general" or falsehood. The essence of the lie entails that the liar is actually and completely in possession of the truth which he is hiding: "a man does not lie about what he is ignorant of" (p. 87). Moreover, the liar *intends* to deceive and does not try either to hide this intention from himself or disguise the translucency ("translucidité") of consciousness.[3] Through the lie, consciousness affirms its existence as *hidden from the Other,* and exploits for its own ends the ontological duality between myself and myself viewed from the eyes of the Other (pp. 88, 89).

But the situation cannot be the same for "bad faith." As lying to oneself, bad faith of course hides or misrepresents truth. "But what changes everything is that in bad faith it is from myself that I conceal the truth" (*EN,* p. 87, translation mine). Here there is no ontological duality between deceiver and deceived. On the contrary, bad faith entails "the *unity* of a single consciousness." The one who lies and the one to whom the lie is told are one and the same. Consciousness "affects itself" with bad faith. Hence, it must have both an intention and project of bad faith. Moreover, given the complete "translucency" of consciousness, "that which affects itself with bad faith must be conscious (of) its bad faith" (p. 89). Otherwise, we should have a consciousness ignorant of itself, which, for Sartre, is patently absurd (p. 11). Early in *Being and Nothingness* we learn that, because of the self-referential and prereflective nature of consciousness,[4] we can wish "not to see" a specific aspect of our being only if we are acquainted with that aspect which we do not wish to see. Accordingly, our flight from anguish, for instance, in order not to know it, is in *bad faith* because we cannot avoid knowing from what we are fleeing (p. 83).

But at this point we need no further details. If bad faith, in contrast to falsehood, is originally viewed as a "lie to oneself," if it unites in a single consciousness both the "deceiver" and "deceived," if it is prereflectively aware of the deceit with which it is trying to afflict its own consciousness, then a question must arise as to its possibility. Given the translucent nature

[3] For Sartre, it is to be remembered, all consciousness is consciousness (of) consciousness; consciousness is nonthetic and aware of itself.

[4] Marjorie Grene, in her noteworthy book, *Sartre* (New Viewpoints, 1973), takes issue with Sartre's self-referential view of consciousness. For a discussion of this, see Ronald E. Santoni, "Marjorie Grene's 'Sartre,' " *International Philosophical Quarterly,* December 1975.

of consciousness, is it possible for me to hide the truth from myself? Or, to put the issue more pointedly, is it possible, in the light of Sartre's characterization both of the lie and human consciousness, to lie to oneself? We must give concentrated attention to this question.

2. *Is "lying to oneself" possible?* At the end of his preliminary characterization of bad faith, Sartre himself seems troubled by this issue. Concealing the truth from myself implies knowing it, but "How then," he asks, "can the lie subsist if the duality which conditions it is suppressed?" (p. 89). If the "being of consciousness" is the "consciousness of being," he ponders, perhaps then my consciousness of bad faith is in "good faith"? But then this "whole psychic system," is done away with. He quickly avows, ". . . If I deliberately and cynically attempt to lie to myself, I fail completely in this undertaking; the lie falls back and collapses beneath my look; it is ruined *from behind* by the very consciousness of lying to myself which pitilessly constitutes itself well within my project as its very condition" (p. 89).

To be sure, Sartre acknowledges the problem and appears to have a clear awareness of it. It seems obvious that he is right in inferring that if the would-be liar "knows" (or "is in possession of") the truth which he is trying to disguise, then any conscious effort to hide that truth from himself would be in vain. Yet this acknowledgement, even of the failure or inevitable self-destructiveness of any deliberate attempt to lie to oneself, does not lead Sartre to abrogate his preliminary conception of bad faith. What we discover here, Sartre tells us, is the "evanescence" of bad faith. This is a phenomenon which fluctuates between "good faith" and "cynicism" and exists only "in and through its own differentiation."[5] Bad faith is *"métestable"*—a word which Sartre contrives to designate the mercurial, abruptly transitional nature of its psychic structures (p. 90). At this point he claims neither to reject nor understand it. Although he acknowledges its "very precarious" existence, he is still prepared to say that (note!) "it can even be the normal aspect [way?] of life for a *very great number* of people" (p. 90, italics mine).[6] Keeping in mind our question, we must attempt to understand the ambivalence of Sartre's position and, by looking at key aspects of his more developed view, come to some determination as to its salvageability and/or consequences.

a. *The 'ideal lie.'* To be noted first is that even in his preliminary

[5] *EN,* p. 88, ". . . qui n'existe que dans et par sa propre distinction."

[6] The French word "aspect" might also be translated "appearance." But in the present context, a looser translation, "way," might be better suited. See *EN,* p. 88 or *BN,* p. 90.

discussion Sartre makes a point of telling us that the lie which he has described is the "ideal lie" (p. 88). In the present context,this is important, I believe, for two reasons. First, although for the sake of relieving the problem we might wish the opposite, Sartre does not, at this point, suggest that bad faith involves anything less than an *ideal* lie.[7] On the contrary, he tells us that the "common, popular forms of the lie," in which the liar only "half persuades" himself of it, are "degenerate aspects" of the lie and represent "intermediaries between falsehood and bad faith" (*ibid.*). Hence, we cannot, *at this stage,* circumvent the problem by affirming that this project (of "lying to oneself") lacks either clarity[8] or persuasion. Second, this discussion is important for its introduction of the concept of "half-persuasion." Although Sartre cannot cash it in here, it may be the main wedge into his culminatory treatment of bad faith and his confrontation with "the true problem of bad faith" (pp. 112 ff.). We shall see later that the issue of nonpersuasion and persuasion bears significantly—though problematically—on Sartre's subsequent notion of a peculiarly *nonpersuasive* evidence (p. 113).

b. *'Facticity' and 'transcendence.'* Sartre realizes that the problem which he faces has led many to take recourse in a theory of the unconscious. For reasons generally consistent with his view of the translucency of consciousness—but which others have analyzed and found wanting[9]—Sartre rejects the Freudian psychoanalytic theory. Concluding that his problem is left "untouched" by the psychoanalytic alternative, he turns to his well-known descriptions of "patterns of bad faith." *Assuming* that these are instances of bad faith, even though he has offered no further claim to comprehend it, he asks, "What must be the being of man if he is to be capable of bad faith?" (p. 96; cf. also p. 85).

Sartre's reply is prompt and definite. His phenomenological description of the erratic, transitory behavior of a coquettish woman pending imminent sexual advances appears to give him a basic insight into the mechanism of bad faith. The human being, for Sartre, is at once a *facticity* and a *transcendence*—roughly spreaking, both a "given" and the possible

[7] Of course, the difference between the *unitary* structure of the former and the ontological *duality* of the latter remains.

[8] On the following page, he says clearly (in regard to "bad faith"): "I must know the truth very exactly *in order* to conceal it more carefully" (*BN,* p. 89).

[9] See, for example, a recent article by Lee Brown and Alan Hausman, forthcoming in a *Festschrift* to Sartre in "The Library of Living Philosophers" series. I have heard only a first draft of this paper. It is not within my purposes in the present paper to deal with this disagreement.

projects associated with its freedom. The coquette uses to her advantage this "double property" of human reality (p. 98). She is aware of the possibility of "the first approach" but chooses to recognize only the "respectful" and "discreet" in her companion's attitude. She is aware of the desire she evokes but purifies it of anything humiliating by acknowledging it only as pure transcendence (p. 99). She leaves her hand to be held but at that moment "becomes" all intellect. She reduces her companion's actions to *what they are* while enjoying his desire as transcendence, i.e., as *not being what it is.* Despite what she *knows,* she attempts to transform facticity into transcendence and vice versa. She plays "see-saw" with the two aspects of her being. By doing so, she feels that she is escaping all reproaches. But she does so at the price of arresting, of "glueing down," of thingifying, her possibilities—of objectifying her transcending freedom. She is in bad faith (p. 99).

As two aspects of human reality, facticity and transcendence should be susceptible to what Sartre calls a "valid coordination." But bad faith wants neither to coordinate them nor overcome them in a synthesis. Rather "bad faith seeks to affirm their identity while preserving their differences" (p. 98). The coquette affirms her intellect or respect in the mode of being of a thing even while distinguishing between herself as object of desire (body) and herself as the intellectual or wholly respectful being which she could become. Bad faith "must affirm facticity as *being* transcendence and transcendence as *being* facticity, in such a way that at the instant when a person apprehends the one, he can find himself abruptly faced with the other" (*ibid.*). Hence, the *métestable* double-property of transcendence-facticity becomes a condition for the possibility of the *métestable* concept of bad faith.

Although this point serves to illuminate what Sartre "intends" by bad faith, it does not eliminate his dilemma(s) concerning it. For, even given the facticity-transcendence character of human reality, consciousness, it must be recalled, "affects itself with bad faith" (p. 89). And as a consequence of its prereflective nature, consciousness, for Sartre, must be aware of its intention of hiding the truth from itself—in this case of its deceptive project of "sliding back and forth"[10] from facticity to transcendence.

---

[10] I borrow this telling expression in this context from Joseph S. Catalano, *A Commentary on Jean-Paul Sartre's 'Being and Nothingness'* (New York etc.: Harper & Row, 1974), pp. 85, 86. I commend this little book as a significant addition to Sartrean scholarship. For my more extended response to this book, see my review of it in *International Philosophical Quarterly,* December 1975.

c. *Being for itself.* There is a closely related way of viewing human reality which is at the root of Sartre's facticity-transcendence analysis and is more directly responsive to his question concerning the necessary ontological prerequisite of bad faith. Anyone familiar with even a skeletal outline of Sartre's phenomenological ontology is acquainted with it. As early as in the "Introduction" (of BN), in which he offers a preliminary distinction between "being for-itself" ("l'être-pour-soi") and "being in-itself" ("l'être-en-soi"), between the being of consciousness and the being of the phenomenon, between self-referential being and being "what it is," Sartre sees the former as distinctive human reality and defines it as a being "which *is* what it is not and which *is not* what it is" (p. 28; exact words from p. 100; italics mine). Unlike being in-itself, being for-itself or consciousness is not a "what" or object or thing. That is to say, it is not "what it is"; it lacks identity or "a certain coincidence [or oneness] with itself" (p. 147); it has no "nature" or fixed essence. On the other hand, it "is what it *is not*"; it is its possibilities, its future undetermined projects, its potential for "transcendence." In different words, the *being* of human reality is not what it is and is what it is not because it is *free.* Indeed, for Sartre being for-itself or human reality *is* freedom. "The essence of the human being is suspended in his freedom," says Sartre. "What we call freedom is impossible to distinguish from the *being* of 'human reality'; . . . there is no difference between the being of man and his *being-free*" (p. 60; cf. also p. 567).

The connection between this paradoxical manner of viewing human reality and Sartre's alleged relation of "facticity-transcendence" to "bad faith" should now become more apparent. "Bad faith"—*assuming* for the moment that it is possible—could not affirm facticity as *being* transcendence and transcendence as *being* facticity, in the abruptly evanescent manner which Sartre suggests, unless human reality were a being "which is what it is not and which is not what it is." This game could not be played between facticity and transcendence unless, "within" human reality itself, there were the possibility of free interplay between "what it is" and "what it is not": unless—to put it more radically—there were a "built-in" "tension" or dialectical "to-and-fro" between the "what it is" and "what it is not" aspects of human reality. It is because the human being's "esssence" is *suspended* in his freedom, because human *being*, as consciousness, exists "at a distance from itself" (if I may be permitted later terminology, p. 125) that it can move freely and elusively between facticity and transcendence and *attempt* to treat each in the mode of being of the other. It is because I both am and am not what I have been, both am and am not *only*

my past, for example, that I can try to escape from my past by appealing to the "perpetual re-creation" which my freedom entails (p. 100).

A brief return to Sartre's coquettish woman might add clarification here. It is because she "*is* what she is not" and "*is not* what she is," qua human reality, that she can try to convince herself of the *possibility* that she is in intellectual discussion with her companion when in fact she is being "advanced to" sexually. After all, she is not just body, and she is more than her present emotions or behavior. Involvement in intellectual conversations is among her possibilities. Her freedom cannot foreclose this. So, given the situation, she shifts disingenuously to what she is not. In so doing, she not only denies that she is mere body but arrests her transcendent intellectual capability and turns it into the mode of being of a thing or in-itself (facticity). She is said to be in "bad faith" for she cannot be "all intellect" in the manner that an inkwell is an inkwell: she is a being who *is not* what she is and *is* what she is not: as human reality, she is free and without identity. To borrow a passage from the last part of *Being and Nothingness*: "To the extent that the for-itself wishes to hide its own nothingness from itself and to incorporate the in-itself as its true mode of being, it is trying also to hide its freedom from itself" (p. 567).

We now, then, have the more basic answer to Sartre's question as to what must be the *being* of man if he/she is to be capable of bad faith. Sartre puts it directly: "The condition of the possibility for bad faith, in its most immediate being, in the intra-structure of the prereflective *cogito,* must be what it is not and not be what it is" (p. 112). "If man is *what he is,* bad faith is forever impossible" (p. 101, italics mine);[11] bad faith "requires that . . . there be an imponderable difference separating being from non-being in the mode of being of human reality" (p. 111). Were human reality not so constituted, the fluctuating interchange between our transcendence and facticity could not be attempted, and we could not, in "bad faith," be involved in the self-destroying enterprise of treating ourselves as identities, of fleeing our freedom.

But, despite this elucidation and the interconnections we have observed, we must ask whether we are brought any further toward accepting the feasibility of "lying to oneself," on Sartre's terms. I submit not. However directly Sartre speaks to the ontological requirements needed for human *being* to be capable of "bad faith," he has not shown how that *being* of

[11] In a relevant supportive statement, Sartre says, "If I were sad or cowardly in the way in which this inkwell is an inkwell, the possibility of bad faith could not even be conceived" (p. 110).

human reality permits lying to oneself. However well he shows that "a being which *is* what it is not and which *is not* what it is" makes possible the game of reciprocal metamorphosis between facticity and transcendence, he has not shown how this game involves lying to oneself, in his initially defined sense of lying. No matter how sensitively he unveils the "double property" of human reality and describes its inner workings, he cannot get away from his own view, as we have noted, of the "total translucency" of consciousness. It would appear that no interworkings between the "two aspects" of human reality (=consciousness), however deceptive in intent, can escape this translucency.[12] To suggest that they do would not only violate a critical dimension of his theory of consciousness but appear to break the "psychic unity" on which he insists against Freud and the psychoanalysts.

We are left to wonder if Sartre's view of bad faith as "lying to oneself" is salvageable. Perhaps, as his position develops, "lying to oneself" takes on a modified sense. We must probe further.

d. *Faith, belief and 'nonpersuasive' evidence.* In an important and demanding section entitled "The 'Faith' of Bad Faith" (pp. 112-116), Sartre affirms that the "true problem" of bad faith originates in the fact that bad faith is "faith." We are not dealing here, he maintains, with certainty, if "certainty" is to be understood as "the intuitive possession of the object."[13] However, if "belief" is taken to mean "the adherence ["l'adhesion"] of being to its object" when the object is either not given or given indistinctly, then bad faith is *belief* and the fundamental problem of bad faith is a problem of *belief* (p. 112). Our question concerning the possibility of bad faith as *lying to oneself* then becomes a question, roughly speaking, of whether and how it is possible, given one's consciousness of one's attempts to lie, to believe (have faith in) one's own would-be lies to oneself.

At this point, so much depends on the nature of "faith," the relation of belief to evidence, and—perhaps most importantly—the manner in which one "puts oneself" in (or "affects oneself with") bad faith. All of these issues come together here. Sartre now tells us that in its original pro-

[12] Sartre himself, in a statement which at this point seems problematic, says, ". . . There exists an infinity of types of behavior in bad faith which explicitly reject this kind of explanation [by means of the unconscious] because their essence implies that they can appear only in the translucency of consciousness" (p. 96).

[13] I have some difficulty with Hazel Barnes' translation of "evidence" as "certainty" (*EN*, p. 108: . . . "si l'évidence est la possession intuitive de l'objet"). Yet given the context and Sartre's later discussion of knowledge (e.g., *BN*, p. 295), I accept it as adequate. Note, however, that Barnes translates the same word on the next page (*EN*, p. 109) as "evidence."

ject and in its coming "into the world," bad faith makes a conscious decision regarding the precise nature of its requirements. For it realizes that "faith is decision and that after each intuition, it must decide and will *what it is.*"[14] It thus stands forth "in the firm resolution *not to demand too much,* to count itself satisfied when it is *barely* persuaded, to force itself in decisions to adhere to uncertain truths" (p. 113, italicizing of "barely" is mine). This means that, although bad faith perceives evidence, it resigns itself in advance to not being "fulfilled" by the evidence. In other words, it commits itself ahead of time to a "*nonpersuasive* evidence" which *it* designs as nonpersuasive. Prima facie, this bears resemblance to the common lies of "half-persuasion" which Sartre has earlier dismissed as "degenerate aspects of the lie" and has distinguished from bad faith (p. 88).

In any case, what we can now infer, according to Sartre, is that the primitive project of bad faith is itself in bad faith. Or, to put the matter in a slightly different way, the initial enterprise of bad faith is to be seen as a decision made in bad faith about the nature of faith. Because faith is not certainty, the consciousness of bad faith (or "bad faith consciousness") decides to be content with an insufficiency of evidence; to determine arbitrarily the amount of evidence by which it will be "persuaded" while "knowing" that the amount it requires is not sufficient to persuade *fully.* In a passage which is germane to this point, Sartre says the following:

> I am not only in bad faith at the end of my effort when I have constructed my two-faced concepts ["concepts amphiboliques"] and when I have persuaded myself. In truth, I have not persuaded myself: to the extent that I could be so persuaded, I have always been so. And at the very moment when I was disposed to put myself in bad faith, I of necessity was in bad faith with respect to this same disposition ["ces dispositions mémes"] (p. 112; *EN,* p. 108).

And he adds poignantly:

> 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; *it believes itself and does not believe* itself in good faith (p. 113, italics mine).

What we see, then, is that from its very inception bad faith is aware of its structure and attempts to exploit the nature of faith by setting up weak requirements for the acceptance of *nonpersuasive* evidence. Bad faith begins with an awareness that human reality is *métestable,* that it *is* what it is not and *is not* what it is. On this "intuition," it sets out to forge the instruments of its own (non) self-persuasion. It "decides" that "nonpersuasion is the

---

[14] Sartre says (*EN,* p. 109): "elle n'ignore pas, dit elle, que la foi est decision," etc. I am baffled as to why Barnes translates this: "it is not ignorant . . . that" etc., instead of "it does not ignore . . . that" etc. See *BN,* p. 113.

structure of all convictions." As consciousness is a being which is always "in question," which always has nothingness at its heart, so belief must be a being which constantly "questions its own being." This means that belief can realize itself only at the price of self-destruction, that it can make itself known to itself only by simultaneously denying itself. This is what Sartre is getting at when he says, "To believe is to know that one believes, and to know that one believes is no longer to believe" (p. 114.). To believe is not-to-believe, so "pure belief" is impossible.[15] This allows him to say also that it is the impossibility of perfect "good faith" (i.e., of totally believing what one believes) that makes bad faith possible. The reason for him is evident: consciousness is evanescent; it is constantly fleeing from identity with itself; it is "perpetually escaping itself." Hence in its being, the immediate "becomes" mediate, the absolute becomes relative, and belief becomes nonbelief (p. 115). In other words, because of the nature of consciousness, *everything* in it must be "in question" or *métestable;* otherwise, the translucency of consciousness would be spoiled by the opaqueness of the *en-soi.* As a consequence, every belief "falls short" ["n'est pas assez croyance" (*EN,* p. 110)], can never believe enough ["ne peut jamais assez croire" (*ibid.*)], and one never fully believes what one believes. Belief then becomes "impossible belief." So my failure to *believe* that I am courageous, for example, does not dispirit me, since believing always comes short of believing. This applies to "good faith"[16] as well as to "bad faith."

Hence we see that for Sartre the primitive project of "bad faith" turns out to be essentially the making use of the constant autodestruction of "the fact of consciousness" ("*fait de conscience*"). As we have observed, because of the alleged nature of consciousness and of "faith," we are able, in bad faith, to set conditions for our being "persuaded" even when we are not persuaded. Similarly, the stage is set for "believing" even when our belief comes short of believing. For Sartre, it follows, of course, that if "bad faith" is "faith" and it includes its negation in its primitive project, then at the outset a faith which wants itself to be "not quite convinced" ("mal convaincue") has to be possible. The faith of the coquette, for instance, in "deciding" in advance not to be fully convinced in order to convince itself that she is not what she is (i.e., a flirt) is in *bad* faith from the start.

[15] I borrow this term from Catalano, *op. cit.,* p. 88. I acknowledge some insights of his here.

[16] I deliberately steer clear, here, of an extended discussion of "good faith." The issue of its meaning and possibility also involves problems, but is a topic for a separate paper.

We are advanced in understanding the project and intentions of bad faith. But, again, we must not forget our critical question as to the possibility of lying to oneself, on Sartre's terms. In reconsidering this question in the light of what has immediately preceded, we also bring to a head Sartre's position regarding bad faith.

Sartre appears to offer a direct reply to the question. He says: "I shall not be able to conceal from myself that I believe so as not to believe and that I do not believe *so as to* believe" (*EN,* p. 110, translation is mine). How then can one be in "bad faith" if bad faith involves "lying to oneself" and "lying to oneself" involves concealing the truth from oneself (p. 89)? In fact, does not this admission totally annihilate "bad faith"? Sartre concedes that "knowing" that I am a coward, for instance, at the moment that I wish to believe myself courageous serves eventually to destroy my belief. But he is quick to make three qualifications, the third of which is most pertinent here. First, considered in terms of the mode of being of the in-itself, I *am* not any more courageous than cowardly. Second, "I do not *know* that I am courageous," for any such view of myself must be accompanied by belief (for it goes beyond reflexive certainty). Third, it *is* the case that bad faith "does *not* bring itself to believe" (*ne s'arrive pas à croire," EN,* p. 111) what it wants to believe. However, *"it is precisely as the acceptance of not believing what it believes that it is bad faith"* (p. 115, italics mine).[17] In contrast to "good faith," which attempts to perform the converse, bad faith flees being by finding refuge in "not-believing-what-one believes." By exploiting the "identity-less" nature of consciousness, by including its negation in its primitive project, bad faith disarms in advance all beliefs—both those it would like to grasp and those that it wishes to flee. That is to say, in "willing" the self-destruction of belief, bad faith at the same time spoils beliefs which are opposed to it—namely, those "which reveal themselves as *being only* belief." For given that the being of consciousness as consciousness is to exist "at a distance from itself" (p. 125) beliefs which are apprehended as beliefs can no longer be "only belief": they are already no longer belief; they are what Sartre later calls "troubled belief" (p. 121).

3. *A modified sense of 'lying to oneself' and the possibility of 'bad faith.'* Two inferences may now, I believe, be drawn from these observations. First, given the totally translucent nature of consciousness and the troubled nature of belief, we must still say that, in the strict sense of successfully and completely hiding a truth from oneself, within the unity of a

[17] I acknowledge following Sartre closely on this page (115).

single consciousness, "lying to oneself" is not possible for Sartre. The consciousness (of) belief cannot miss "questioning" its own belief; the "bad faith" consciousness cannot escape an awareness of its own self-destroying and disarming "project"; the consciousness (of) belief, or disbelief, cannot be ignorant of its believing so as not to believe, or its not believing so as to believe; consciousness cannot flee an awareness of what it is trying to hide, flee, or negate. If consciousness is self-conscious, if it involves a prereflective awareness of itself, it cannot alter or hide anything from itself.

But to say that "lying to oneself" is not possible in Sartre's strictest sense of this expression is not *necessarily* to say that bad faith is impossible in Sartre. It is only to say that it is precluded by his initial, strict sense of "lying to oneself." But weaker, modified senses are possible. I submit that the substance of Sartre's developed position on the phenomenon of bad faith—as expressed particularly in his section on "The 'Faith' of Bad Faith"—allows us to infer that he finally adopts a qualified sense of the expression. We have seen much evidence for this. In its original project, "bad faith" (as Sartre perhaps prematurely calls it) structures itself to accept *nonpersuasive* evidence, to count itself satisfied when it has only met minimal requirements for persuasion, to regard itself as persuaded when it is only partially or "half" persuaded. It constructs "two-faced," i.e., ambiguous, concepts (e.g., persuasion, belief) by which this can be done. Bad faith *decides* that nonpersuasion is constitutive of all convictions. It exploits the nature of belief (belief always "falls short" of belief; "to believe is not-to-believe," given the nature of consciousness) by "pretending" to believe what it cannot *fully* persuade itself to believe, given the translucency of consciousness. And to repeat: "it is precisely as the acceptance of not believing what it believes that it is bad faith" (p. 115).

This statement is surely an important key to understanding both Sartre's modified sense of "lying to oneself" and his developed view regarding the possibility of bad faith. I submit that "the *acceptance* of not believing what one believes" (italics mine) represents one articulation of the qualified sense in which Sartre finally understands "lying to oneself." To revert to the language of "persuasion," we can perhaps also say that "lying to oneself" turns out to have the modified sense of "the acceptance of not being *fulfillingly* persuaded even when one is "persuaded." In any case, one can accept "not-believing-what-one-believes" not only because of the dialectical, métestable nature of belief but because, as we have stressed, the primitive project of bad faith sets weak evidential requirements for persuasion. One can meet these requirements, be "persuaded" on that count, and

yet be aware that the "required" evidence is inadequate for full ("fulfilled") persuasion. One can, in effect, not believe while "believing" and presumably believe while "not believing."

If this qualified sense of "lying to oneself" is endorsed, and if bad faith is, as we saw from the start, "lying to oneself," then it follows that *bad faith, in this qualified sense of 'lying to oneself,' is possible.* By presupposing and acknowledging the translucency of consciousness—its total "openness"—this qualified usage answers our repeated question about the possibility and manner of lying to oneself in the full translucency of consciousness. This modified meaning of the expression also allows a viable reply to Sartre's own question: "How can we believe by bad faith in the concepts which we forge expressly to persuade ourselves?" (p. 112). In addition, it has other advantages: I mention only as few. It sees bad faith precisely in terms of what Sartre ultimately views the problem of bad faith—namely, in terms of *belief* (p. 112). It is grounded on the assumption, thoroughly compatible with the translucency of consciousness, that bad faith does not succeed in believing what it wishes to believe. In attaching itself to the phenomenon of "not-believing-what-one-believes," it testifies anew to the "distance" which exists at the core of the *being* of consciousness, and it elucidates further the métestable "double property" (facticity-transcendence) of human reality.[18] To refer to one concrete illustration, it allows us to see how, finally, the bad faith of the coquettish woman precisely involved her *acceptance* of her not believing (because of the translucency of her consciousness) what she "believed" (namely, that she was involved in intellectual discussion, not bodily flirtation).

What shall we say then? Has Sartre met what he calls his "true problem of bad faith?"(p. 112). I believe the answer to this is a heavily qualified, problematic "yes." Insofar as he has given us, within the context of *faith* and *belief,* a sense in which a "lie to oneself" is possible on his own ontological premises (e.g., that consciousness is totally translucent and that there is an idiosyncratic unitary structure to "single consciousness"), I believe that he has moved towards resolving both his "essential" problem (p. 116.) and our central difficulty with his initial definition of bad faith. Certainly, he has provided us with a sense in which the "lie to oneself," because of bad faith's primitive resolution in respect to evidence, need not necessarily collapse or "fall back,"[19] in the face of the translucency of consciousness.

[18] We must recall that early in his analysis of bad faith, Sartre offers this "double-property" as a condition for the possibility of the métestable concept of bad faith.

[19] One may remember that this was one of Sartre's own worries about lying to oneself.

Nonetheless, there are at least two qualifications or difficulties which must be expressed. *First,* as I already anticipated, because Sartre's qualified sense of "lying to oneself" presupposes a conscious decision to be content with evidence that does not fully persuade, it appears to be a "lie" in a weak sense of that word. As such, it seems to be a lie of "half-persuasion," which Sartre earlier characterizes as a common, degenerate form of the lie and rejects as an instance of bad faith (p. 88). Either there is an inconsistency here, or Sartre, without having specifically expressed it, has altered his approach to "half-persuasion" while developing his account of bad faith. *Second,* given again his view of the primitive project of bad faith, there is no question that the modified sense of "lying to oneself" involves a *deliberate* resolve to accept as persuasive what is not fully persuasive, and to "believe" (according to criteria which one "knows" are inadequate for *full* persuasion) what one consciously does not fully believe. This suggests that there is a cynical element even in Sartre's modified sense of bad faith. But Sartre clearly states in a number of places that bad faith does *not* involve a *cynical* lie (e.g., pp. 112,115). To be sure, Sartre early describes the "cynical consciousness" as one which, though affirming truth within itself, is involved in the "doubly negative attitude" of both denying it in words and "denying that negation as such" (p. 87). Although the qualified sense of bad faith which I have been attributing to Sartre is not so overtly cynical—it seems more like a game of "pretend" which consciousness plays with itself in respect to persuasion and belief!—it does seem, to make "knowing preparations" (*EN,* p. 111, "préparation savante") for offering to oneself as "truth" a "pleasing untruth," (p. 89), etc. Moreover, insofar as bad faith for Sartre finally *denies* the "inner disintegration" (*désagrégation*)[20] or freedom of human reality, which it *both* attempts to *flee* in the human search for identity and to *utilize* in its wish to escape the in-itself, it appears to share—at least partly—in the cynical consciousness. In any case, if this qualified sense of "bad faith" is to hold, Sartre must differentiate it more clearly from both the "cynical lie" and "cynicism."[21]

In conclusion, one point merits explicit affirmation. However adequately this *acceptance* of not-believing-what-one-believes might qualify as

[20] This refers, of course, to the fact that for Sartre human reality is a being which is not what it is and is what it is not; that nothingness therefore lies at the heart of (human) being.

[21] One must not forget that, in the early part of his treatment of bad faith, Sartre has told us that the " 'evanescence' of bad faith . . . vacillates continually between good faith and cynicism" (p. 90). This at least suggests the importance of not *identifying* "cynicism" with the "cynical lie." Among other differences, the former is clearly the broader concept.

a weakened version of lying to oneself, it is a far cry from the view of bad faith (suggested in the early part of Sartre's analysis) which requires a *successful concealment* of truth from consciousness in the full translucency of consciousness.

RONALD E. SANTONI.

DENISON UNIVERSITY.