How is Lying to Oneself Possible? The Dialetheism Reading of Sartre's Bad Faith

Nahum Brown

Abstract: One of the most important debates in Sartre scholarship today comes from the question, how is it possible to be in bad faith? In other words, how is self-deception possible, given that, in lying to ourselves, we are both the liar and the lied to at the same time? On the face of it, this sounds paradoxical, if not downright contradictory. This article aims to address this question (1) by analyzing secondary literature on Sartre that tries to prove that bad faith is not a contradictory concept and (2) by defending the "dialetheism reading" of bad faith, that is, the reading that views bad faith to be evidence for the true existence of contradictions.

Keywords: Sartre, bad faith, contradiction, dialetheism

One of the most provocative concepts of Jean-Paul Sartre's work, bad faith—the concept of lying to oneself—has sparked a number of significant debates. These debates include discussions from commentators about whether bad faith stems from an ontological claim about being or whether it is primarily an ethical concept. Is it a revision of Heidegger's analysis of inauthenticity in *Being and Time*? How does it relate to the Husserlian insight from intentionality, that consciousness is always consciousness of something? Should we always avoid bad faith whenever we can? Or is it in some respects productive to be in bad faith? What are the sources of bad faith? What triggers it and causes it to continue? How frequent is bad faith? Is it common or rare? Debates have also surfaced about the examples Sartre uses to illustrate bad faith—the woman and the waiter, in

particular, have led to controversies about what strategies and methods we use to enter into and relieve ourselves of bad faith.

Yet, perhaps the most perplexing question that emerges from Sartre's discussion of bad faith comes from the seemingly simple task of establishing how bad faith is possible at all. This article analyzes the underlying paradox of how bad faith is possible given that, by lying to ourselves, we are both the liar and the lied to. On the face of it, bad faith sounds paradoxical, if not downright contradictory. Obviously, lying in general (lying to others) is possible, since, in this case, the liar and the lied to are separate people. Lying in general is that conventional form of lying that upholds the category distinction between the deceiver and the deceived, or, in Sartre's vocabulary, upholds the "conditioning duality."[1] But Sartre claims that there is a significant distinction to be made between lying *tout court* and lying to oneself. The paradox of bad faith comes about when the agent and object of the deception are the same person. How can the liar as the source of the lie also be the one who is lied to? Bad faith thus seems to be impossible. Sartre expresses this fundamental problem when he writes: "The person to whom one is lying and the person who is lying are one and the same person, which means that I must know—in so far as I am the deceiver—the truth that is hidden from me in so far as I am deceived … How then can the lie survive, if its conditioning duality is abolished?"[2]

There is a lot of good secondary literature about this logical question of how bad faith is possible. This article has two primary purposes: (1) to catalog and critically analyze some of the most persuasive arguments from the secondary literature that argue for the possibility of bad faith, mainly by presenting bad faith as a non-contradictory concept and (2) against these readings, to defend the dialetheism interpretation of bad faith, that is, the reading that views bad faith to be evidence for the true existence of contradictions. To set up the problem, I begin by reviewing arguments that claim that bad faith is impossible or possible only in a very qualified sense. The most notable of these comes from the thesis that bad faith is overtly contradictory and therefore impossible. I then look at Sartre's objection to the Freudian distinction between the id and the ego as a "hard distinction" that would resolve the paradox of bad faith by showing that it is not contradictory. From there, I look at a series of "soft interpretations" that, in one way or another, attempt to resolve the paradox of bad faith and thereby avoid the

---

[1] Jean-Paul Sartre, *Being and Nothingness* (London: Routledge, 2018), 90.
[2] *Ibid.,* 90.

contradiction by dividing the self in subtle ways:[3] there is the Stevenson-Gordon-Hymers debate in the 1980's that focuses on the question of whether a distinction between reflection and pre-reflection saves bad faith from contradiction; there is also the translucency interpretation from Phyllis Sutton Morris and David Detmer's strategy of establishing the possibility of bad faith through the non-paradoxical structures of lying in general and lying to others. All of these readings assume that contradictions are impossible, but that if we establish either a hard or soft distinction in the self, this will resolve the paradox of being both the liar and the lied to at the same time and save bad faith from contradiction. These readings stand in contrast to the position I defend at the end of the article, which is that bad faith demonstrates the real existence of contradictions and should be embraced as thoroughly paradoxical, in the tradition of Graham Priest's dialetheism.

Arguments that Bad Faith is Impossible

There are at least two ways to argue that bad faith is impossible. The first way comes from the claim that Sartre conceives of consciousness as transparent, and that, if it is transparent, then lying to oneself is also transparent and therefore impossible. M.R. Haight argues for this position in *A Study of Self-Deception*.[4] Haight claims that because he is a follower of Descartes, Sartre views consciousness to be completely transparent. If that is right, then there is no real possibility of lying to oneself since the self sees itself and knows itself with full lucidity. It is not hard to disregard this position. Since the textual evidence of the chapter on bad faith shows that Sartre thinks of bad faith as possible (and experiential evidence also shows this to be true), it is reasonable to conclude that he departs from Descartes on this point and views the revelation of consciousness to itself to be a complex structure with barriers, thus returning us to the question of the possibility of bad faith.

---

[3] Jeffrey Gordon uses this distinction between hard and soft views of bad faith in his analysis of Leslie Stevenson's work Sartre. I broaden Gordon's distinction to include a series of readings that attempt to soften what would otherwise be the contradiction of bad faith. Jeffrey Gordon, "Bad Faith: A Dilemma," in *Philosophy*, 60 (1985), 258–262, <https://doi.org/10.1017/S0031819100051147>.

[4] Phyllis Sutton Morris interprets M.R. Haight to be arguing for this position. Phyllis Sutton Morris, "Sartre on the Self-Deceiver's Translucent Consciousness," in *Journal of the British Society for Phenomenology*, 23: 2 (1992), 104–105, <https://doi.org/10.1080/00071773.1992.11006980>. Also see M.R. Haight, *Self-Deception and Self-Understanding: New Essays in Philosophy and Psychology*, ed. by Mike W. Martin (Lawrence: University Press of Kansas, 1985), 53–54.

A more persuasive way to argue that bad faith is impossible is to claim that it is paradoxical to the point of being a full-blown contradiction. Proponents of this position start from the common-sense assumption that contradictions are always impossible, and then claim either that bad faith is contradictory and therefore impossible or that bad faith is not truly or fully an act of self-deception. This is either because the deceiver knows about the deception, in which case, there is no deception after all, or because the deceived truly does not know, in which case, the deception takes on the common form of lying in general. In either case, according to this position, it would be wrong to say that self-deception is possible.[5]

Since bad faith seems to be a common everyday experience, and since Sartre clearly views bad faith to be possible, commentators have worked to resolve the paradox of bad faith by drawing on various distinctions in the self or in the act of bad faith, which, if realized, would exonerate it of contradiction. Let's start by looking at the hard distinction that Sartre himself critically responds to, and then at softer variations, which all attempt to prove bad faith to be non-contradictory.

The Hard Distinction Between the Id and the Ego

Sartre devotes a long passage of the subchapter "Bad Faith and Lies" to a sustained discussion of Freud's distinction in the self between the "id" and the "ego." Here, Sartre claims that one of the primary strategies to avoid the problem of bad faith is to find recourse in the unconscious.[6] By projecting the theoretical division of the mind into the id and the ego, we can thereby explain away the paradox of bad faith. The self, then, becomes like two people. The barrier between the id and the ego allows the self to hide from itself and lie to itself, without, thereby, causing a contradiction or paradox. Even though it is the self that is lying to the self, the hard distinction between the two sides of the self compartmentalizes the lie so that the liar and the lied to can both share the same source of the self but also be ignorant of the other side.

Sartre's critique of Freud's divided consciousness reveals a lot about the nature and paradox of bad faith. Freud's division creates, in Sartre's analysis, an artificial difference that falsely compartmentalizes the act and

---

[5] Santoni suggests this interpretation in his article "Bad Faith and 'Lying to Oneself'" when he claims that bad faith is only possible in a very qualified sense. Ronald E. Santoni, "Bad Faith and 'Lying to Oneself,' in *Philosophy and Phenomenological Research*, 38: 3 (1978), 384–398, <https://doi.org/10.2307/2107007>.

[6] Sartre, *Being and Nothingness*, 91.

object of consciousness, essentially dividing the self into an other, and making a dualism where there is not one. Sartre claims here that, although the deceiver and the deceived are one and the same person, we are not able to release ourselves from this unity by creating a hard difference in the self. To divide the self in this way turns bad faith into a variation of lying in general, which, according to Sartre, covers over the issue and leads to even deeper expressions of bad faith.

Some commentators have pointed out that Sartre's reading of Freud in the Bad Faith chapter is not that sophisticated. According to Jonathan Webber, Sartre over-emphasizes the "resistance" concept in psychoanalysis, which Webber defines as "the purported phenomenon of a psychoanalytic patient engaging in a variety of strategies to prevent the analyst from getting to the truth."[7] But even if Sartre has misinterpreted Freud and his reading should not be viewed as a fair treatment or effective guide of Freud's work, knocking down the Freudian strawman nevertheless leads to revealing insights about the nature of bad faith: it cannot be resolved by dividing the mind and attempting to demonstrate non-contradiction through the compartmentalizing of the id and the ego, nor by any other opposition of the mind to itself.

The question a number of commentators then ask is whether a softer division of the self can resolve the paradox of bad faith. Obviously, Sartre is opposed to any hard division of the mind that creates a separation and barrier in the self, such that the self appears as the other of itself. But is there a softer temporal, modal, or thetic distinction that can be used to clear bad faith of contradiction?

The Stevenson-Gordon-Hymers Debate

The journal *Philosophy* published a series of response articles in the 1980's about how to resolve the paradox lurking in bad faith. This series of articles is known as the Stevenson-Gordon-Hymers debate. In the first of these articles from 1983, "Sartre on Bad Faith," Leslie Stevenson proposes a distinction between pre-reflective and reflective acts that demonstrates how, even though the liar and the lied to are one and the same person, bad faith is

---

[7] Jonathan Webber, *The Existentialism of Jean-Paul Sartre* (New York: Routledge, 2009), 89. For other discussions that focus on whether Sartre's reading is fair to Freud, see Jerome Neu, "Divided Minds: Sartre's 'Bad Faith' Critique of Freud," in *The Review of Metaphysics*, 42: 1 (1988), 79–101, <https://www.jstor.org/stable/20128695>. Also see, Mary L. Edwards, *Sartre's Existential Psychoanalysis: Knowing Others* (London: Bloomsbury Academic, 2023), 55–60.

nevertheless possible and non-contradictory.[8] Stevenson's reading is intuitive enough: We may very well know the truth of the matter in our pre-reflective consciousness and at the same time lie to ourselves as we reflect on the matter. Because the truth is posited in our pre-reflective consciousness and the lie occurs in the act of reflection, there is nothing inherently contradictory about lying to oneself.

In the second article of the series, "Bad Faith: A Dilemma" (published in 1985), Jeffrey Gordon repositions Stevenson's analysis by proposing that there is a hard and a soft interpretation of Stevenson's distinction.[9] Even though Stevenson uses analytically precise language—such as, "for all conscious predicates F, and all people x, x is F if and only if x is *pre-reflectively* aware that he is F"[10]—Gordon objects that there is an ambiguity in Stevenson's analysis. It is not clear to what extent Stevenson is proposing a distinction in the self. If the difference between reflection and pre-reflection leads to a hard distinction in the self, then Gordon thinks that Stevenson has devised a similar solution to the ego-id distinction, which would face all of the same objections that Sartre leveled against Freud.[11] But, according to Gordon, if we think of pre-reflective consciousness and reflective consciousness as merely the difference between two modes of the self, rather than as a full distinction, then the division does not do enough work to solve the paradox of bad faith.[12] Thus, Stevenson's dilemma, as Gordon sees it, is that no matter whether he goes along with the hard or soft interpretation of reflection, bad faith does not work out conceptually. Gordon thus knocks down Stevenson's position, while claiming that even though the reading aligns well with Sartre's intended concept of bad faith,[13] it is not a sustainable reading. Gordon does not go on to propose an alternative reading that would clarify how bad faith is possible, leading his reader to assume that his underlying conclusion is that bad faith is impossible.

Michael Hymers offers a further layer of critical analysis in his 1989 response article "Bad Faith." He claims that neither Stevenson nor Gordon have adequately considered Sartre's discussion of how lying in general differs from lying to oneself,[14] and that because of this oversight, their

---

[8] Leslie Stevenson, "Sartre on Bad Faith," in *Philosophy*, 58 (1983), 254–256, <https://doi.org/10.1017/S0031819100068741>.

[9] Gordon, "Bad Faith: A Dilemma," 258–259.

[10] Stevenson, "Sartre on Bad Faith," 254.

[11] Gordon, "Bad Faith: A Dilemma," 259–260.

[12] *Ibid.*, 260-261.

[13] *Ibid.*, 262.

[14] Michael Hymers, "Bad Faith" in *Philosophy*, 64 (1989), 397, <https://doi.org/10.1017/S0031819100044740>.

readings are problematic. Hymers thinks that because Stevenson confuses the unique structure of bad faith with lying in general, he ends up with too much of a division in the self; even the soft version of Stevenson's reading is too strong. This oversight also leads Gordon to the erroneous conclusion that Stevenson's reading aligns well with Sartre's intentions about bad faith, causing Gordon to dismiss bad faith as impossible. After outlining these criticisms, Hymers offers a positive reading of bad faith by claiming that its character is marked by ambiguity, double-negation, and half-truths, and that the reason why it is possible is because it relies on a level-distinction between implicitly and explicitly knowing something.

The Translucency Reading

In addition to positing a distinction between reflection and pre-reflection, there are a number of other soft strategies that attempt to mitigate the puzzle of self-deception. In "Sartre on the Self-Deceiver's Translucent Consciousness," Phyllis Sutton Morris presents a translucency reading of bad faith. Morris claims that translucency is not the same as transparency.[15] Sartre's view of consciousness is sometimes misunderstood as a transparency theory of consciousness,[16] where the self has unconditional access to itself. Morris claims, instead, that Sartre's view of consciousness is translucent, which means that our mental life is public, that consciousness is out there in the world,[17] but also that our access to ourselves is opaque, like "looking through frosted window glass."[18]

The concept of translucency is multi-layered. On the one hand, translucency means that the self is all of one piece and that it cannot be divided into strict compartments, as the hard distinction between the id and the ego purports to do. But, on the other hand, Morris' thesis that translucency is not transparency aims to resolve the seemingly paradoxical nature of bad faith by exploring a series of opacities, which enable the self to lie to itself. These opacities can be viewed as low-level barriers, processes of organization, or even marginally divided compartments of the self. By Morris' estimation, they include (1) the distinction between thetic and non-

---

[15] Morris, "Sartre on the Self-Deceiver's Translucent Consciousness," 105-106.

[16] For example, Haight, *Self-Deception and Self-Understanding: New Essays in Philosophy and Psychology*, 53–54.

[17] Sartre discusses the idea that consciousness is out there in the world in terms of the concept of intentionality in Husserl. See Jean-Paul Sartre, "Intentionality: A Fundamental Idea of Husserl's Phenomenology," trans. by Joseph P. Fell, in *Journal of the British Society for Phenomenology*, 1: 2 (1970), 4–5, <https://doi.org/10.1080/00071773.1970.11006118>.

[18] Morris, "Sartre on the Self-Deceiver's Translucent Consciousness," 105.

thetic experience, (2) the soft version of Stevenson's distinction between reflection and pre-reflection, as well as (3) the temporality of the self, and (4) being-for-itself in contrast to being-for-others.[19] (1) thetic versus non-thetic positioning organizes human experience through the lens of attention, thereby softly dividing the translucent intentionality of consciousness and the world into foreground and background structures, which leads to enough opacity in the self to lie to oneself. The other three structures work in similar ways: (2) reflection allows the self to re-interpret and thereby to deceive pre-reflective consciousness; (3) temporality allows the self to undermine the responsibility that we have towards the past with the projection of the future; and (4) Sartre's well-known distinction between being-for-itself and being-for-others opens paths for bad faith through the difference between first and third person perspectives. Together, these four translucent structures work in synergy to demonstrate the underlying conditions that allow bad faith to be possible and prevalent in our lives. Morris' main argument is that, because of these four structures, the seemingly paradoxical nature of bad faith is resolved and the contradiction of being both the liar and the lied to is averted.

The Deceiving-Others-Makes-Deceiving-Oneself-Possible Interpretation

In *Sartre Explained: From Bad Faith to Authenticity*, David Detmer proposes a different strategy from the soft reflection and translucency readings that Stevenson and Morris propose. Detmer's solution to the paradox of bad faith is to recognize that lying in general has a specific structure and tactics that make it possible—namely misleading statements and the discordance of ambiguity and vagueness—and that just as these tropes make lying in general possible, likewise they make it possible to conceive of lying to oneself. In the section of his book called "Bad Faith,"[20] Detmer follows a two-step process of diagnosing the fundamental categories of lying in general and then extending these categories to self-deception. Since lying to others is clearly possible and commonplace, the advantage of Detmer's strategy is that he can rely on simple inductive arguments by pointing to the experiential evidence of conventional lying in general and then establish how the self can turn these tactics of deception on itself.

Detmer claims that there is a whole array of tropes that makes deception and misdirection possible, from simply misleading the other to the

---

[19] *Ibid.*, 107–110.

[20] David Detmer, *Sartre Explained: From Bad Faith to Authenticity* (Chicago: Open Court, 2008), 75–89.

more complex exploitations of the double- and multiple-meanings of ambiguity and vagueness. These tactics are commonly directed towards others in the form of political rhetoric, courtroom guile, and deceitful salesmanship. However, these same tactics can be folded back on the self to initiate bad faith. Detmer points to our common propensity to advocate for one or another side of a debate, while pretending, at the same time, that we are also being receptive and open-minded to the other sides of the debate, as an instance of bad faith. We also have a propensity to let ourselves be persuaded by weak evidence,[21] and we also commonly direct fallacies of distraction, such as the red herring, on ourselves as often as we do on others. Each of these instances demonstrate, on Detmer's reading, how the typical shapes of lying to others equally enable self-deception to be possible. Through the exploitation and manipulation of ambiguity and vagueness, the self becomes able to play both roles—both the liar and the lied to—simultaneously without causing a full-blown contradiction.

There is a notable disadvantage to Detmer's approach, however. Similar to Hymers' criticism of Stevenson and Gordon, one can object that Detmer conflates bad faith with lying in general. To make his reading work, Detmer has to argue that bad faith is simply a carry-over of lying in general, even though Sartre clearly wants to uphold a category distinction between the two structures. What if lying to oneself requires a different set of conditions? What if it cannot be explained in the way that the conventional tactics of exploiting ambiguity and vagueness explain lying in general? What if bad faith requires the contradictory unity of opposites, where the same person is both the liar and the lied to and this cannot be explained away by projecting and othering the self? Furthermore, Gordon's objection to the soft version of Stevenson's pre-reflection strategy—that it does not do enough work to save bad faith from contradiction—can also be applied to Morris' more complicated synergy of soft distinctions of translucency.

The Dialethism Reading

What if bad faith is genuinely contradictory and the point is not that we should avoid or explain away the paradox, as most Sartre commentators have assumed? What if the point, instead, is that we should embrace it as contradictory, not so that we can live in bad faith and saturate ourselves in it, but so that we can diagnose its paradoxical nature as one of the fundamental structures of reality and consciousness? This type of questioning leads to the dialetheist reading of bad faith. Most commentators (and Sartre himself at

---

[21] *Ibid.*, 84.

some places in the text)[22] assume that the seemingly paradoxical nature of bad faith needs to be rooted out and resolved, that if bad faith is contradictory, then it is impossible. But to approach bad faith in this way is to undermine Sartre's overall position about the self, which is that it cannot be divided in the way that other people are divided from the self. Moreover, to approach bad faith in this way is to reduce lying to the self to an extended mode of lying in general. The proponent of the dialetheist reading claims that, far from needing to resolve the seemingly paradoxical nature of bad faith, its perplexity should be embraced. This proponent believes that bad faith is contradictory but that this does not make bad faith impossible. Let's briefly analyze some of the main ideas and examples of dialetheism, as Priest states them generally, and then draw up insights about bad faith based on the reading that it is a productive contradiction of consciousness and reality.

Central to the project of dialetheism is the controversial claim that some contradictions genuinely exist. It is important to note, however, that for Priest only *some* contradictions exist, not *all* contradictions exist.[23] One of the reasons why we tend to react strongly against dialetheism is because we conflate *some* with *all*. Priest agrees that the claim "all contradictions exist" is problematic; nevertheless, he thinks that there are cases of contradiction, i.e., states of affairs where both something and its opposite occur as one unity. Proponents of the dialethism reading of bad faith think that lying to oneself is one of these contradictions that genuinely exist. Against a long tradition of Sartre commentators who attempt to resolve the contradiction of bad faith, the dialetheists uphold bad faith as a prime example of *a* contradiction that exists. Let's look at three different explanations Priest gives of dialetheism and decide whether, or to what extent, bad faith aligns with these examples.

First, in his full book-length study of dialetheism, *Beyond the Limits of Thought*, Priest delineates the field of dialetheism in purely formal terms by presenting an exhaustive list of the logical possibilities for the truth function of a given entity, x. There are four permutations in total: (1) x is true and not false; (2) x is false and not true; (3) x is neither true nor false; and (4) x is both true and false.[24] Because of the prejudices of traditional logic and our commonsense abhorrence of contradiction, the standard assumption is that only (1) and (2) are acceptable. Based on the conventional definition of bivalence, x is either true or false, but not neither and certainly not both. Priest

---

[22] For example, Sartre says phrases like "if we wish to resolve this difficulty [of the paradox of bad faith]…" Sartre, *Being and Nothingness*, 97.

[23] Graham Priest, "What is so Bad about Contradictions?," in *The Journal of Philosophy*, 95: 8 (1998), 410, <https://www.jstor.org/stable/i323997>.

[24] Graham Priest, *Beyond the Limits of Thought* (Oxford: Clarendon Press, 2002), 4.

claims that logicians who are more liberal might go along with (3), but accepting (4), that x is both true and false, is definitely not something traditional logicians would entertain. This tradition can be traced back to Aristotle, who articulates the law of non-contradiction in the *Metaphysics* to the effect that something cannot both be itself and the opposite of itself at the same time, manner, and place.[25]

Priest often expresses dialetheism in purely formal and logical terms, as he does when he presents the four possibilities of truth function. In this way, he both challenges traditional logicians by exposing longstanding commonsense prejudices about the law of non-contradiction but also remains faithful to traditional logic by upholding the project of formalizing reason. He challenges the conventions of logic by demonstrating that there clearly are other possibilities of truth function than those produced from the limitations of bivalence, and that, moreover, there is no good justification for assuming that (1) and (2) are acceptable while (3) and definitely (4) are unacceptable.

The proponent of the dialetheism reading of Sartre's bad faith believes that self-deception can be interpreted as a uniquely non-formal expression of the real existence of contradiction as it occurs through the negativity of the self. In *Beyond the Limits of Thought*, Priest's discussion of contemporary philosophy post-Hegel is mostly concerned with formal articulations of dialetheism, preferring to analyze debates about the nature of contradiction in contemporary philosophy of language and logic.[26] One of the values of the dialetheist reading of bad faith is that we become able to demonstrate contradiction from the insights of descriptive, non-formal phenomenology and existentialism. A short chapter on Sartre's bad faith would have made for a worthy addition to Priest's contemporary account of dialetheism in *Beyond the Limits of Thought*. Such a chapter would have fit well alongside Priest's discussion of Heidegger and Derrida and would have helped to develop avenues for the analysis of non-formal, experiential accounts of contradiction in terms of the self.

Let's look at two more of Priest's examples of dialetheism, which are more friendly to non-formal articulations of contradiction, and which align well with the dialetheism reading of bad faith. In one of his most lucid expositions of the basic principles of dialetheism to date, "What is so Bad

---

[25] For Aristotle's discussion of the law of non-contradiction. See Aristotle, *Metaphysics*, in *The Complete Works of Aristotle*, in two volumes, ed. by Jonathan Barnes (Princeton: Princeton University Press, 1984), Book IV, Sections 3-6, 1005a20–1011b22. Also see Graham Priest, "To be and not to be - that is the answer. On Aristotle on the Law of Non-Contradiction," in *History of Philosophy & Logical Analysis*, 1 (1998), 91–130, <https://doi.org/10.30965/26664275-00101007>.

[26] Two notable exceptions to this are Priest's discussions of Derrida and Heidegger. See Priest, *Beyond the Limits of Thought*, 219–222 and 235–248.

about Contradictions?," Priest discusses the existence of contradiction as a paradox of self-reference, such as the lair's paradox.[27] The liar's paradox can be triggered from the simple statement, "I am lying," which, if true, is false, and if false, is true. Priest suggests that our attempts to mitigate such paradoxes of self-reference are futile and that we would be better off accepting these paradoxes as paradoxes.

Lying to oneself is a paradox of self-reference, similar to the liar's paradox, in the sense that if we deliberately lie to ourselves and acknowledge this, then we are telling the truth and not lying. Thus, it seems to be impossible to catch oneself in bad faith. Or, as Sartre writes, "if I deliberately and cynically attempt to lie to myself, I must completely fail in this undertaking: the lie recedes and collapses before our eyes; it is ruined, *from behind*, by the very consciousness of lying to myself that constitutes itself before my project, as its very condition."[28] The liar's paradox can be set in formal terms, as the truth function of statements, but it can also be set in more general non-formal terms as well, especially in terms of lying to oneself. This comparison helps us to understand why Sartre claims that bad faith constantly evades us whenever we try to articulate that we are definitely doing this to ourselves.

Priest also mentions another explanation of dialetheism in "What is so Bad about Contradictions?" in terms of standing inside and outside of a room:

> I walk out of the room; for an instant, I am symmetrically poised, one foot in, one foot out, my center of gravity lying on the vertical plane containing the center of gravity of the door. Am I in or not in the room?[29]

If you are standing both inside and outside of the room, then it is both true and false that you are inside the room and both true and false that you are outside of the room. Different from a self-reference paradox, this example illustrates the existence of a contradiction in our everyday sphere of experience. We are often faced with double-positions, ambiguities, and vagueness in our everyday lives. According to the dialetheist, there is nothing wrong with these common experiences, and we should not flee from them or dress them up as non-contradictory.

---

[27] Priest, "What is so Bad about Contradictions?," 415.

[28] Sartre, *Being and Nothingness*, 90.

[29] *Ibid.*, 415.

Priest's example of standing both inside and outside of a room presents us with a significantly different way of perceiving these seemingly mundane everyday spatial relations. We perceive movement and position as containing the possibility of being both something and its opposite as one unity. This leads to statements such as that it is both true and false that I am standing inside the room. But this also leads to dialetheist insights about movement, perception, as well as a re-conception of the identity of things in the world. Bad faith has a similar effect. According to Sartre, because the self is constituted by the concept of negativity, the self therefore contains, as one of its fundamental modes of being, the act of self-deception, not as a division of the self, nor even as a folding of the self upon itself, but as an immediate pure act of its being. To conceive of this requires a significantly different way of perceiving the identity of being a self, such that the self has the capacity to lie to itself without being an other to itself.

Conclusion

There is a lot of good secondary literature about Sartre that focuses on the fundamental question of how bad faith is possible, given that in lying to ourselves, we are simultaneously the liar and the lied to. An effective way to organize and catalog this secondary literature is to recognize that, in one way or another, most commentators aim to save bad faith from its seemingly paradoxical nature by demonstrating why bad faith is not contradictory and therefore impossible. The Stevenson-Gordon-Hymers debate does this by discussing the distinction between reflection and pre-reflection in the self; Morris does this by proposing a translucency reading that establishes a series of soft divisions in the self; Detmer does this by claiming that lying to oneself is possible because it follows similar patterns to the tactics people use to lie to each other. All of these voices contribute, in one way or another, to the conclusion that bad faith only seems to be paradoxical but that it is, in truth, a non-contradictory concept.

However, in this article, I outline an alternative reading of bad faith that embraces bad faith as genuinely contradictory rather than a reading that attempts to save it from paradox. To this end, I discuss some of the main principles and examples Priest offers in his account of dialetheism. I ultimately draw the conclusion that this alternative reading of bad faith converges with Priest's project of dialetheism in the sense that both claim that contradictions exist and can be productive. But the reading of bad faith diverges in terms of Priest's project of formalizing dialetheism. Following the practice of descriptive phenomenology and existentialism, Sartre presents

bad faith in non-formal, experiential terms. Formalizing dialetheism is an abstraction derived from the more primary structures of existence. These structures appear first through the experience of our everyday lives, through what the phenomenologists call "the givenness of the phenomena in intuition." If bad faith can be included as a moment in Priest's historical account of dialetheism, it belongs to a unique branch of dialetheism that resists formalization.

Department of Philosophy, Chiang Mai University, Thailand

References

Aristotle, *The Complete Works of Aristotle*, in two volumes, ed. by Jonathan Barnes (Princeton: Princeton University Press, 1984).

Detmer, David, *Sartre Explained: From Bad Faith to Authenticity* (Chicago: Open Court, 2008).

Edwards, Mary L., *Sartre's Existential Psychoanalysis: Knowing Others* (London: Bloomsbury Academic, 2023).

Gordon, Jeffrey, "Bad Faith: A Dilemma," in *Philosophy*, 60 (1985), 258–262, <https:// doi.org/10.1017/S0031819100051147>.

Haight, M. R., *Self-Deception and Self-Understanding: New Essays in Philosophy and Psychology*, ed. by Mike W. Martin (Lawrence: University Press of Kansas, 1985).

Hymers, Michael, "Bad Faith," in *Philosophy*, 64 (1989), 397–402, <https://doi.org/10.1017/S0031819100044740>.

Morris, Phyllis Sutton, "Sartre on the Self-Deceiver's Translucent Consciousness," in *Journal of the British Society for Phenomenology*, 23: 2 (1992), 103–109, <https://doi.org/10.1080/00071773.1992.11006980>.

Neu, Jerome, "Divided Minds: Sartre's 'Bad Faith' Critique of Freud," in *The Review of Metaphysics*, 42: 1 (1988), <https://www.jstor.org/stable/20128695>.

Priest, Graham, "To be and not to be—that is the answer. On Aristotle on the Law of Non- Contradiction" in *History of Philosophy & Logical Analysis,* 1 (1998).

Priest, Graham, "What is so Bad about Contradictions?," in *The Journal of Philosophy,* 95:8 (1998).

Priest, Graham, *Beyond the Limits of Thought* (Oxford: Clarendon Press, 2002).

_____, "To be and not to be - that is the answer. On Aristotle on the Law of Non-Contradiction," in *History of Philosophy & Logical Analysis*, 1 (1998), 91–130, <https://doi.org/10.30965/26664275-00101007>.

_____, "What is so Bad about Contradictions?," in *The Journal of Philosophy*, 95: 8 (1998), 410–426, <https://www.jstor.org/stable/i323997>.

Santoni, Ronald E., "Bad Faith and 'Lying to Oneself'," in *Philosophy and Phenomenological Research*, 38: 3 (1978), 384–398, <https://doi.org/10.2307/2107007>.

Sartre, Jean-Paul, "Intentionality: A Fundamental Idea of Husserl's Phenomenology," trans. by Joseph P. Fell, in *Journal of the British Society for Phenomenology*, 1: 2 (1970), 4–5, <https://doi.org/10.1080/00071773.1970.11006118>.

_____, *Being and Nothingness*, trans. by Sarah Richmond (London: Routledge, 2018).

_____, *L'Être et le Néant* (Paris: Éditions Gallimard, 1943).

Stevenson, Leslie, "Sartre on Bad Faith," in *Philosophy*, 58 (1983), 253-258, <https://doi.org/10.1017/S0031819100068741>.

Webber, Jonathan, *The Existentialism of Jean-Paul Sartre* (New York: Routledge, 2009).