

3-25-2011

On Bad Faith as a Means of Adding Clarity to the Internalist Debate

Cal Margulis
Macalester College

Follow this and additional works at: <http://digitalcommons.macalester.edu/philo>

Recommended Citation

Margulis, Cal (2011) "On Bad Faith as a Means of Adding Clarity to the Internalist Debate," *Macalester Journal of Philosophy*: Vol. 8: Iss. 1, Article 13.
Available at: <http://digitalcommons.macalester.edu/philo/vol8/iss1/13>

This Article is brought to you for free and open access by the Philosophy Department at DigitalCommons@Macalester College. It has been accepted for inclusion in Macalester Journal of Philosophy by an authorized administrator of DigitalCommons@Macalester College. For more information, please contact scholarpub@macalester.edu.

On Bad Faith as a Means of Adding Clarity to the Internalist Debate

Cal Margulis

One of the most obvious characteristics of moral belief as distinct from non-moral belief is the tendency for such belief to be action-guiding. Few people could both hear someone announce his holding of a particular moral belief and see him act in a manner contrary to that belief without coming to the conclusion that something was amiss. For millennia, philosophers have debated the implications of this action-guiding characteristic, unsurprisingly reaching many different and often contradictory conclusions.

Today, such debate largely centers around the many descriptive moral theories that belong within the rather vague category of "internalism." Understood in its most general form, internalism is the position that there exists some sort of inherent connection between morality and action. Such a simple definition, however, belies the tremendous amount of disagreement and controversy that exists among those philosophers who are intimately concerned with the issue. For the complexities of the internalist debate to be more fully understood, one must analyze the conceptual distinctions made by internalists and externalists (those opposed to internalism) alike.

One of the two most significant areas of debate surrounding modern internalism concerns the general theoretical nature of the connection between morality and the actions of a moral agent. Some internalists hold the belief that moral concerns provide a person with *motivation* to act based upon those concerns. Thus, it is impossible for a moral agent to possess a moral belief without being to *at least some degree* moved, urged, or otherwise compelled to act in a manner consistent with that belief. Gilbert Harman states the matter rather succinctly in his book, *The Nature of Morality*, when he says, "To think that you ought to do something is to be motivated to do it. To think that it would be wrong to do something is to be motivated not to do it" (Harman, 33).

Other internalists believe that moral concerns provide only *reason* to act, and not motivation. Kant's categorical imperative is one example of this. To Kant, the existence of certain moral truths inherently provided moral agents with sufficient reason to act upon them, regardless of the agents' awareness of truths.

This distinction between motive and reason alone, however, does not provide us with much information. For such a basic outline to be even vaguely helpful in our attempt to understand internalism about reasons, we must first establish exactly what is meant by the term "reason." Almost all English speakers would likely be able to agree on a workable definition of the word "motive," since

it is a rather specific, extremely demonstrable psychological fact. The word "reason," however, is a far more vague term, possessing a multitude of different denotations and connotations. David O. Brink understands this lack of clarity of meaning and attempts to make the issue somewhat more clear by analyzing two distinct meanings of the word "reason."

Brink first establishes one sense of the phrase "reasons," that of "the considerations that *motivate* the agent and so *explain* her actions." He finds this usage extremely inadequate for the purposes of the internalist debate, however, because of the extremely relative and subjective nature of such "reasons." As he states, "... we often think that an agent can have explanatory reasons for action without having *good* or *justifying* reasons." He then gives the example of a person who believes that light bulbs are nutritious and thus eats them to remain healthy. He quite accurately states that, although this person does have in one sense (the purely explanatory sense) a *reason* to eat light bulbs, since light bulbs are obviously *not* nutritious, this person's reason for light-bulb consumption is neither good nor justified (Brink, 39).

Indeed, such usage does not seem to capture what is usually meant when the word "reason" is used within the context of ethical discourse. Also, this usage *certainly* cannot be used to assist the internalist question. If one's "reasons" for action consist merely of the psychological factors that motivated one to act, where is the distinction between "reason" and "motive"? Brink therefore concludes, "internalism about reasons concerns the connection between moral considerations and good or justifying reasons for action" (*Ibid.*).

As convenient as this definition is, however, it does raise the question of exactly what it means for a particular reason to be "justified" or not. Since it lies almost completely outside of the domain of morality, Brink's example of the light-bulb eater does very little to help answer the question. This is probably just as well, since it is unlikely that anyone could create a criteria for "justness" that would be accepted by all, or even most, reason internalists. Some would base their distinctions between "just" and "unjust" upon traditional Christian morality, some upon Confucian ethics, and so on. It would therefore seem to be the responsibility of any particular reason internalist to explain the theoretical structures that she believes underlie her definition of a "just reason."

The difference between reason and motive is, however, but one of the distinctions made within internalist moral theory. As defined by Brink, all forms of internalism, both motive- and reason-based, can be shown to belong to one of three categories. These categories are defined largely based upon the nature of the morality that is being considered, specifically the relative objectivity or subjectivity that it possesses.

Agent internalism, for example, is essentially the belief that the existence of objective moral obligations inherently provide a rational agent with either reason

or motivation to act according to such obligations. The knowledge that the agent has of the moral obligation is unimportant: the very existence of such obligations is sufficient to provide her with reason or motivation. One can find a prime example of this within the traditional theology of Catholicism. According to this theology, all human beings have an inherent need to accept the truth regarding Jesus' life, death, and resurrection, a need not obviated by a lack of knowledge. The woman who heard the Gospel and yet turned her back on Christ has just as much of a reason to accept his divinity as the man who lived his entire life never hearing Jesus' name.

The second category, agent internalism's more subjective counterpart, is called by Brink *appraiser internalism*. According to this theory, the *possession of moral beliefs* inherently implies that the possessor has either motive or reason to act based upon these beliefs. Thus, according to the motive-based appraiser internalist, independent of any sort of objectively defined morality, if I believe that it is fundamentally wrong to kill someone in self-defense, it would be impossible for me to be in a situation in which such an action were an option without having motive to avoid it.

It is with this specific brand of internalism that I am particularly concerned. Specifically, I am struck by the divisions that exist within such a seemingly unified area of philosophy: for there is not just one, but two distinct types of appraiser internalism about motives. The weak version claims that the possession of moral belief inherently provides one with a certain degree of motivation to act. The strong version claims that the possession of moral belief provides one with *sufficient* motivation to act. My intention is to demonstrate that, in light of Jean-Paul Sartre's analysis of "bad faith," the strong, motive-based appraiser internalist description of morality is not only flawed, but blatantly incoherent. Because of humanity's somewhat inherent possibility for self-denial, it is extremely possible to both truly hold a moral belief and yet act in a manner contrary to the spirit of this belief.

Sartre's major focus in the first chapter of *Being and Nothingness* is the existence of non-being as a counterpart to being. Taking his cue from Heidegger, he questions the being of humanity or, as he defines it, the For-itself as it relates to its existence within the world. He also, as Heidegger before him, analyzes the questioning process in terms of its three fundamental elements: the being who questions, the being who is questioned, and the subject about which the questioning takes place. Upon his delineation of the questioning process, however, he is struck by one particular aspect, the essential duality of response of the questioning being faced. This dichotomy between positive and negative reply, between existence and non-existence, between being and nothingness, forms the basis of a great deal of his philosophical constructs, most particularly his concept of bad faith.

Sartre's response to this duality is to accept non-being not as the mere absence of being, but as a real possibility in and of itself—a true counterpart to being. In regard to the questioning process, Sartre states that the "*permanent possibility of non-being*, outside us and within, conditions our questions about being" (Sartre, 36; italics mine). With the possibility of negativity firmly established, Sartre next seeks to determine the source of this negation.

This consideration brings him to what he feels to be the true source of non-being, the For-itself. Sartre's separation of the universe into the categories of the Being-for-itself and the Being-in-itself is based largely upon humanity's ability to judge, to question, to posit non-being as a concrete possibility. Beings-in-themselves, non-conscious beings, have no such possibility. Since a stone, for example, lacks consciousness, it also lacks the possibility to question, either itself or other beings. Therefore, it is perpetually what it is: a stone.

The For-itself, however, because it possesses the ability to question, is not simply what it is. Unlike a stone, a human being has no labels that can be affixed to it which completely encompass the entirety of its being. A woman, for example, can become President of the United States. However, Sartre believes that the verb "is" in the phrases "Jane is President" and "this is a stone" must be interpreted in distinctly different ways. A stone is quite certainly a stone, and has no possibility of being anything else. However, because the word "President" is a theoretical construct, all Jane can do is try as hard as she can to act like a President. The word "President" can never truly encompass her being in the same way that the words "stone," "table," and "pen" can encompass the beings to which they refer.

Thus, according to Sartre, the most fundamental quality of mankind is the ability to choose. Unlike all other beings, the For-itself alone has the capability of determining what is and what is not, of accepting or negating possibilities. Sartre gives the example of the concept of "destruction." It is only through the For-itself, he claims, that destruction comes into the world. An earthquake, a tornado, a supernova all exert very distinct changes upon the environment in which they exist. However, it is only because of the For-itself's capability to judge that such change is labeled "destruction." It is due to this unique capability that Sartre calls the For-itself "*a being by which nothingness comes to things*" (*Ibid.*, 57).

After his analysis of the "Origin of Negation," Sartre moves on to examine the way in which this permanent possibility for negation and the secretion of nothingness affect the way in which the For-itself deals with itself. For it is not only possible for the For-itself to deny things in the external world: the being of the For-itself also allows it to deny itself. It is in response to this capability that Sartre formulates the question, "What are we to say is the being of man who has the possibility of denying himself?" (*Ibid.*, 87).

Sartre realizes the impossibility of fully answering such a question, and therefore limits his questioning to what he calls "one determined attitude which is

essential to human reality and which is such that consciousness instead of directing its negation outward turns it toward itself." Sartre calls this attitude "bad faith (*mauvaise foi*)" (*Ibid.*). Essentially, Sartre views bad faith as the process of lying to oneself. However, he realizes the incongruous nature of such self-deception, and openly wonders how such a thing is even possible.

"The essence of a lie implies that the liar is in complete possession of the truth which he is hiding." This is indubitable, but it does raise the fundamental question of how a person can both be aware of a certain fact (in her capacity as liar) and yet remain unaware of that same fact (in her capacity as being lied to). Sartre adds, "Better yet, I must know the truth very exactly *in order* to conceal it more carefully. . ." (*Ibid.*, 89). This would seem to imply a logical contradiction, one that would destroy the possibility of bad faith.

The answer, to Sartre, lies in the very being of the For-itself. As alluded to previously, "we have to deal with human reality as a being which is what it is not and which is not what it is" (*Ibid.*, 100). Consider the example of Egbert, a repressed homosexual. Egbert has since his childhood consistently found himself being attracted in a distinctly sexual way towards certain male friends and colleagues, and has largely failed to understand the attraction of his male friends and colleagues towards women. He has even gone so far as to pursue sexual encounters with other men. He recognizes the greater pattern that such desires represent, namely, his homosexuality.

However, Egbert was also raised according to strict, American fundamentalist Christian morality, and therefore has also been raised to understand homosexuality as evil, anti-Christian, and hateful towards God and family. Having heard of the evils of homosexuality since he was old enough to understand English, he is firmly convinced that homosexuality is sin, that it removes one from the presence of God, and that above all else, *it must not be done*.

Egbert does not wish to remain abstinent for life, however, and therefore seeks to find a way of reconciling these two parts of his being. He finds this reconciliation through bad faith. Egbert decides to believe that he is a heterosexual. Each time he has a sexual encounter with another man, he tells himself that it was an isolated incident, a mistake, a slip of some kind and not indicative of a greater pattern of behavior. He then asks God for forgiveness and moves on, still to some degree believing himself to be a heterosexual.

In Sartian terms, he is accepting both his *facticity* and his *transcendence*, but is refusing to connect the two. To Sartre, the term *facticity* is meant to designate the For-itself's "necessary connection with the In-itself, hence with the world and its own past. It is what allows us to say that the For-itself *is or exists*" (*Ibid.*, 802). The body of the For-itself is an example of its facticity. It is what is given about a particular person's situation, before that person makes the effort of transcendence.

By *transcendence*, Sartre means the "process whereby the For-itself goes beyond the given in a further project of itself" (*Ibid.*, 804). One's place of birth, one's parents, one's body, all of these are the givens that make up one's facticity. If one were an In-itself, these characteristics would be able to completely define one in the same way that the facticity of a stone (e.g., its hardness or its color) is able to completely define it. Because of the nature of the For-itself as a free, choosing being, however, none of the things that make up its facticity can completely encompass the whole of its being. All of the For-itself's actions represent, to a greater or lesser degree, its transcendence, since they fundamentally consist of the For-itself choosing to become what it is not already.

Normally, these two aspects of human reality "are and ought to be capable of a valid coordination." However, Egbert, in so far as he has allowed himself to slip into a pattern of bad faith, "seeks to affirm their identity while preserving their differences" (*Ibid.*, 94). When Egbert engages in sexual relations with another man, he understands the facticity of his situation: what it is that he is doing, the person with whom he is doing it, etc. He also understands his transcendence in relation to both this action and the rest of his life. He (at least subconsciously) realizes that this one encounter is but one part of a greater whole, that of his homosexuality, and believes this to be an immoral enterprise, one that must be stopped.

However, in bad faith he is able to understand both of these facets of his life without making the connections that are necessitated by logic. On the one hand, he can accept his facticity without making the connection to his transcendence. As long as he can pretend with every sexual encounter that he has merely temporarily slipped from the "righteous" path, and not accept his true sexuality, he can continue to have sexual relationship after sexual relationship without ever confronting the disparity between his moral beliefs and his actions. However, he can also, within the structure of his bad faith, accept and understand his transcendence, enjoying the sexual act as a distinctly homosexual act, thus simultaneously accepting and rejecting his transcendence.

It is at this point that the inadequacy of the strong-motive, internalist position becomes clear. According to such a position, Egbert's situation is impossible. He fully believes that the continued practice of homosexual acts is the height of immorality. He has never even heard it suggested that anything to the contrary is true. According to the strong-motive internalist, this inherently implies that he *must* do everything within his power to see that he never has sex with another man. However, the reason that Egbert set up his web of bad faith in the first place was so that he would be able to continue to have homosexual encounters without changing his moral beliefs. Therefore, unless one is prepared to deny the existence of sexually active homosexuals who believe that homosexuality is evil, it would seem inconsistent to believe anything stronger than the idea that the

holding of a moral belief inherently provides some motivation to act based upon said moral belief.

Bibliography

- Brink, David O., *Moral Realism and the Foundations of Ethics*. New York: Cambridge University Press, 1989.
- Harman, Gilbert, *The Nature of Morality: An Introduction to Ethics*. New York: Oxford University Press, 1977.
- Mackie, J.L., *Inventing Right and Wrong*. New York: Penguin Books Ltd., 1977.
- Sartre, Jean-Paul, *Being and Nothingness: A Phenomenological Essay on Ontology*. New York: Washington Square Press, 1956.