Seyed Mohammad Hossein Mir-Mohammadi

PhD

LOOKING THROUGH SELF-DECEPTION FROM ETHICAL STANDPOINT

"Self-Deception" is a phenomenon which has a close relation with ethical issues. So, ethical perspective can explain self-deception different from those mentioned in other contemporary discussions like philosophy of mind and philosophical psychology. Such a standpoint invites new ethical questions and problems such as "is the phenomenon morally harmful?", "is the agent aware of the phenomenon?", if so, " how is it possible for the agent to deceive (her)himself intentionally? Having a look on the contemporary readings of the phenomenon, it seems impossible to solve the problems by remaining faithful to philosophical literature. However, this paper, which is based on the ethical, philosophical resources, tries to present the phenomenon in a new way. This suggestion is able to solve the problems on one hand, and remain faithful to philosophical heritage, on the other.

Key words: *self-Deception, the paradox of self-deception, the empirical and pure will.*

Introduction

Despite the fact that much of the contemporary philosophical discussion of self-deception has focused on epistemology, philosophical psychology and philosophy of mind,the relation between ethics and self-deception can open doors to new horizons. This new viewpoint can invite couple of questions such as "is the agent responsible for being self-deceived?", "isself-deception morally harmful?", "in what ways and to what extent is it morally harmful?" and so on. Additionally, it enables us to think about self-deception not only as an experienced phenomenon but also as an immoral one. If so, self-deception can be related to religious perspective as well since the religious-reliable sources convey the responsibility for being deceived. Accordingly, one, such as divine-command-theory holders, may claim that the self is responsible for being deceived religiously while it is an immoral phenomenon[1]

This paper tries to look at this phenomenon through the ethical viewpoint in a new way. Although this viewpoint takes the philosophical debates mentioned in the traditional and contemporary philosophical literature seriously, it enables us to solve the problems differently.

In the interest of finding such a scope, the paper has been charted as follows: firstly, different definitions of self-deception and its paradox in both traditional and contemporary literature will be investigated. Then, it will be explained why there is a relation between "the will", as a fundamental constituent of person, and "self-deception". Because of neglecting the role of "the will", philosophical problems of self-deception still exist. Finally, based on the mentioned relation between *the phenomenon* and *the will and its types*the new way of understanding self-deception will be suggested.

Definitions of self-deception

Here is an example about my friend. He told me that he is "very smart" and in fact "smarter than others in his age" while he does not read a lot, and does not get any signs for being smart. Knowing that my friend never lies, and knowing him to be no fool, I see he is deceiving himself.

Self-deception is a phenomenon which can be experienced by humanbeings. We are, notes William Hirstein, often deceived about how good we are at something [Hirstein, 2005, p. 213]. Anice example of this is the following fact: "94 percent of educators describe themselves as better at their job than their average colleague, which would yield a self-deceptionrate of 44 percent" [Gilovich,1991, pp. 77–78].

If so, what is self-deception? what does it refer to? Can it be modeled in the way the agent be aware of the occurrence? Is it a possible phenomenon? In the interest of finding the answers, Let us return to philosophical literature.

According to philosophical literature, other-deception has a clear meaning. "In the ordinary life", as Martin W. Mike notes,"*A* deceives *B* when *A*, who knows a truth, purposefully gets *B* to believe the opposite – a falsehood. This process usually happens by lying to *B*"[2].

Accordingly, "in a wide sense, self-deception refers to (a) activities of evading significant truths or topics, and (b) resulting states of ignorance, false, or un-

---

[1] There is a sharp distinction between lying and deception, however. The distinction should be taken into account while we try to explain self-deception. Deception has a success condition involving the target of the deception that lying lacks. I can lie to you even if you fail to believe me. But I can only deceive you if I succeed in getting you to believe something I am aware is false. SO, it implies that deception is related to a new belief which does not already exist. It is not possible to deceive you into not changing a belief you already held.]Hirstein, 2005,p. 219.

[2] There is a sharp distinction between lying and deception, however. The distinction should be taken into account while we try to explain self-deception. Deception has a success condition involving the target of the deception that lying lacks. I can lie to you even if you fail to believe me. But I can only deceive you if I succeed in getting you to believe something I am aware is false. SO, it implies that deception is related to a new belief which does not already exist. It is not possible to deceive you into not changing a belief you already held [Hirstein, 2005, p. 219].

warranted beliefs or lack of clear consciousness. In a narrow sense, self-deception refers to a special case of such activities and states that is describable as a form of lying- lying to oneself" [Mike, 2001, p. 1551].

Self-deception, then, can be modeled as follows:

"Where *A* intentionally gets *B*, [i. e. him/herself]to believe some proposition *p*, all the while knowing or believing truly not-*p*" [Deweese-Boyd, 2010].

With an eye to philosophical debates, I shall call the mentioned model, "straight" self-deception. In "straight" self-deception people are self-deceived in believing something they want to be true.

However, not all cases of self-deception are straight. There is another type of self-deception termed "twisted" self-deception by Alfred Mele [Mele, 2001, pp. 94–118]. In "twisted" self-deception, people who are self-deceived in believing that *p* do not desire that *p*. "A typical example of such self-deception is the jealous husband who believes on weak evidence that his wife is having an affair, something he doesn't want to be the case. In this case, the husband apparently comes to have this false belief in the face of strong evidence to the contrary in ways similar to those ordinary self-deceivers come to believe something they want to be true" [Deweese-Boyd, 2010; Pears, 1982, pp. 42–44].

According to this model,a philosophical problem arises. The problem is the fact that is it really possible for A to deceive A (i. e. him/herself); and persuade him/herself into believing what he knows is false?

Many philosophers tried to answer to the question. For example, according to Jean-Paul Sartre (1905–1980) [Sartre, 1966, 86–116] and Raphael Demos (1892–1968) [Demos, 1960, 588–95] people can deceive themselves by directing their attention to a false belief while ignoring both contrary sign and the activity of disregarding sign. Sartre writes: "we say indifferently of a person that he shows signs of bad faith (i. e. falsehood) or that he lies to himself. We shall willingly grant that bad faith is a lie to oneself. (…) The essence of the lie implies in fact that the liar actually is in complete possession of the truth which he is hiding" [Sartre, 1966, p. 48]. Other perspectives, such as Sigmund Freud's (1856–1939), recognize self-deception as an unconscious phenomenon, and if so, it is an unintentional process rather than an intentional activity.

Trying to solve the problem, some philosopher-shad an attempt to present another definition for self-deception. Self-deception, in this view, is a phenomenon which has the following characters: (a) it has the production of the ill-grounded thought; and (b) then, the ill-grounded thought is protected and allowed to remain. The metaphor of *protection* applied here refers to special cases: the cases in which "a person intentionally does not allow himself to think about, remember, or perceive certain facts; or the cases in which the self-deceived person intentionally does not use certain checking procedures" [Hirstein, 2005, p. 216].

Examples of this account can be traced in different fields. For example, according to some scientists, there is no spiritual experience in people's ordinary life. So, evidences expressed by people or philosophers, like R. Otto[1] (1869–1937) and others, should be somehow wrapped. As another example, atheists who reject all philosophical books which defend the existence of God do the same. Indeed, their account has the production of the ill-grounded thought, firstly. This ill-grounded thought will be protected and allowed to remain, secondly.

This understanding of self-deception explains the mechanism behind self-deception by focusing on the brain's functions: "As … Schnider (2001) and others noted, much of the work of the brain goes into inhibiting representation from interfering with the dominant thought processes. This necessity has given the brain the power to deactivate representations or direct attention away from them in various ways. These functions might be part of the mechanism behind certain types of self-deception" [Hirstein, 2005, p. 216].

Indeed, self-deceived people tend to dwell on neither the feared thought nor the self-deceptive thought[2]. The self-deceived person need a process to improve his or her state but does not employ them. According to this account of understanding, "self-deceived people protect certain beliefs by failing to think them through clearly and vividly. If the self-deceived person thinks the ill-grounded thought clearly, it will produce a strong negative emotion and it may well start procedure to stop doing what he is doing according to that thought" [Hirstein, 2005, p. 216].

In this explanation which takes into account the process of human's brain, there is still a question whether or not the self-deception is voluntary. Since some checking processes are damaged in the case of self-deception,one can then frame the question whether the agent voluntary fails to control over these processes. "The voluntary control might be operated either by keeping away the thought from checking processes, or the checking processes from the thought. Both of these can be done intentionally or *non*intentionally" [Hirstein, 2005, 214–215]. Accordingly, both intentional and nonintentional process of self-deception can be found.

This account distinguishes two types of deceptions: shallow and deep deceptions. "In shallow deception, the deceiver treats his victim the way he would treat any other victim". Deep deception occurs when the deceiver takes into account differences between his own beliefand those of his victim. So, "the deceiver who assumes that his victim has basically the same beliefs as he does is still engag-

---

[1] R. Otto in *The idea of Holy*, not only explains characteristics of the moment in which the agent experience or feel the idea of holy, but presents it in a philosophical way also [Otto, 1936].

[2] Recalling the example of twisted self deception, it is possible to give an example for the feared thought and self-deceptive thought. The husband's thought ("She is having an affair") is an example of the feared thought while ("She's faithful") is the husband's self-deceived thought.

ing in shallow deception. Maxims of shallow deception can be presented when people have a natural inclination to believe positive claims about themselves, one should appear confident, and so on" [Hirstein, 2005, p. 219].

Contrary to those who hold that self-deception involves a (false) belief, there is a quite different approach to the analysis of self-deception does not center on the understanding of self-deception in terms of belief at all. This approach connects the phenomenon with personal identity. Herbert Fingarette (1921), as an example, holds that the self-deceiver "is one who is in some way engaged in the world but who disavows the engagement, who will not acknowledge it even to himself as his. That is, self-deception turns upon the personal identity one accepts rather than the beliefs one has" [Fingarette, 1982, p. 213].

Does Self-Deception contain any paradoxes?

There is a famous paradox in self-deception which prevents philosophers to accept the regular understanding of the phenomenon. In the interest of finding any paradox, let us return again to the example about my friend mentioned at the beginning.

You may think at one level that my friend *must* know that he is unintelligent because of the evidence. But he keeps telling me with a sincere voice that he is intelligent, so he must believe that he is intelligent. If knowledge entails belief, as we often say it does, then my friend must both believe that he *is* intelligent and *is not* intelligent when he is deceiving himself. So we have arrived at the famous paradox of self-deception. The core of the paradox is the question "how is it possible for a person to believe p and not-p consciously?"

Some hold that self-deception is not possible; for it is impossible to believe p and not-p simultaneously. Other philosophers hold that there is no paradox in self-deception at all [Mele, 1987, p. 121]. The third group of thinkers hold that it only *seems* that there is a paradox. Because the philosophical efforts often make the phenomenon much more difficult to understand that we have reason to believe it to be. So, the paradox can be solved.

In General, the paradox of self-deception was formulated by many philosophers[1]. Among all possible alternatives, it can be formulated as follows:

"For any A and B, when A deceives B into believing that p, A knows or truly believes that not-p while causing B to believe that p. So when A deceives A (i. e. himself) into believing that p, he knows or truly believes that not-p while causing himself to believe that p. Thus, A must simultaneously believe that not-p and believe that p" [Mele, 1987, p. 121].

The paradox of self-deception can be summarized in three main questions: "(1) how does the person both know and not know p? (2) what good does it do not to know p consciously? (3) why is the faster, more accurate, system unconscious?" These three

questions, as A. Greenwald states, can condense the paradox of self-deception [Greenwald, 1997, p. 54].

There are philosophers who solve the problem by believing that there is no paradox in self-deception. They try to explain the phenomenon differently. Herbert Fingarette's (1921) analysis, which has been presented earlier, attempts to avoid paradox although it was judged unsuccessful by other philosophers [Fingarette, 1982, p. 53]. What follows is the summary of the way in which he defines self-deception without any paradoxes.

Fingarette rejects philosophers' attempts which analyze self-deception and its paradox. He holds that their attempts "either had not addressed themselves to the proper phenomena of self-deception or rather than resolving the paradox, had portrayed it in a "variant form"" [Fingarette, 1982, p. 54]. He explains the self-deception in the form of "an unnamed mechanism that analyzes the true import of circumstances and purposefully prevents the emergence into consciousness of both the true information and the defense against it" [Greenwald, 1997, p. 54]. He assumed that his unnamed mechanism operated outside of *conscious cognition*. So, there is no paradox in self-deception at all.

Although the analysis of self-deception presented by Fingarettesolves the paradox apparently, it is set beside certain of Freud's doctrine. [2]If so, there is still an important question in his theory should be answered by him. This question can be formulated as follows: how is it possible to justify Freud's doctrine philosophically?

Indeed, the question put all Freudian-like understandings of self-deception under question. Although Freud's psychological theory has been accepted by many psychologists, it still stands in need of philosophical justification. Therefore, philosophical models which remain faithful to Freud's doctrine should, in advance, explain why his theory is acceptable philosophically.

Is it possible, then, to define self-deception on one hand, and be able to solve of the phenomenon's paradox philosophically, on the other? The following part tries to investigate for the answer.

Will and Self-Deception: A Suggestion for Understanding Self-Deception

As it has been mentioned earlier, the Freudian-like approaches, i. e. non-internationalist approaches, put the contradictory belief at an unconscious level. If so, there is no room to hold that the self is responsible morally or religiously for being deceived. Indeed, according to Freudian-like approaches, there is no relation between morality and self-deception at all. Although these approaches have been supported by psychologists, they still stand in need of philosophical justification.

---

[1] See also Audi 1982, p. 133, Bach 1981, pp. 351–370, Canfield and Gustavson 1962, pp. 32–36, Pugmire 1969, p. 339, Siegler 1968, p. 147.

[2] For Freud, "the censor operated from a base within the conscious ego, and although it appeared to have ego's reasoning powers, nevertheless was assumed to operate without ego's consciousness" [Freud. 1923/1961,3–66 cited from Greenwald, 1997].

Let us return to the Mele's model once again in order to explain why it is not necessary to put the contradictory belief at the unconscious level. Melehas argued that the phenomenon "is simply a hiatus between what the evidence suggests, on one hand, and what the self is lead to believe, on the other". What happens is the fact that "the self's desire or emotion, finally, induces him(her) to believe". Indeed, the available evidence suggests the proposition *p* while the self, because of the desire, believes the false proposition (not-*p*) [Mele, 2001, p. 138]. So, the phenomenon can be experienced consciously.

Although Mele's model tries to support the view in which the phenomenon happens in the conscious level, it does not still give us a complete explanation why no paradox remains in the model. There is still a possible room for the self's belief about *p* even in a case which the desire plays a role for inducing the self to believe not-*p*. If so, is it possible to present an account which considers self-deception as a non-paradoxical account, without using the unconscious level?

"Will and Self-deception" is a suggestion for explaining self-deception not only as a conscious and non-paradoxical phenomenon but also as a phenomenon which has a close relation with morality.

Based on the statements from Kant's Second Critique, a new model of self-deception can be presented. In this model, the phenomenon will happen either in the objects related to the empirical will or in the process which the pure will requires. So there is no room for any paradoxes or challenges between two contradictory prepositions. To make the suggestion more clear, let us turn to Kant's Second Critique.

Recall Kant's statements of empirical and pure will. For empirical will, *"*all practical principles which presuppose an object (matter) of the faculty of desire as the ground of determination of the will are empirical, and can furnish no practical laws" [Kant, 1889, p. 107]. If self-deception relates to the self's desire, as Melesuggested, the empirical will plays the role in self-deception. If so, the phenomenon can be explained by focusing on the empirical will.

The Kantian-based formulation of the empirical will can be found from two deductions based on his Second Critique:

(1) Empirical principles as suchare of one and the same kind,and come under the general principle of self-love [Kant, 1889, p. 108];

(2) The principle of self-love cannot set up as a practical (universal) law [Kant, 1889, p.,113];

Therefore,

(3) Empirical principles cannot set up as a practical (universal) law.

And,

(4) Empirical principles as such are of one and the same kind, and come under the general principle of self-love [Kant, 1889, p. 108];

(5) The principle of self-love only desires [Kant, 1889, p. 126];

Therefore,

(6) Empirical principle(s) only desire(s);

Couple (3) and (6) with the fact that "all material practical rules place the determining principle of the will in the lower desires, and if there were no purely (…) laws of will adequate to determine it, then we could not admit any higher desire at all" [Kant, 1889, p. 109].

According to the characteristics of the empirical will, it can be concluded that the empirical will can be determined either by the lower or higher desires. In other words, the empirical will relates to different types of desires. If so, it is possible to experience conflicts among them in a particular context. So, self-deception can be explained as the result of a conflict among the different types of desires which the empirical will relates to:

The empirical will can be determined by a lower or higher desire: the lower desire determines the will while the higher desire still exists; or the higher desire determines while the lower desire still exists. Consequently, the object of the lower and higher desire may differ. It means that in the conflict, as we suppose it, the lower desire may determine the empirical will by an object, e. g. p, and the higher desire may determine the empirical will by another object, e. g. not-p. So, there is no paradox here while the will can distinguish both objects.

Based on the statement, Let us explain the example of my friend and his idea about himself mentioned earlier. My friend had two desires: (a) a desire to be proud of himself, (b) a desire to be a realist. By comparing these two desires, it will be concluded that in that situation, it is impossible for both desires to determine the will. A conflict among these two desires starts: the desire (a) will determine the will and then he would say he is "very smart" and in fact "smarter than others". Suppose that desire (b) determines the will. Consequently, he will say that he is not smart.

Not all cases of self-deception can be reduced to the conflicts among desires while the pure will has an important role to play. The pure will, has its own mechanism to be determined. If so, it can be considered as a basis of new explanation of self-deception as well.

The pure will relates to "pure reason [which] is practical of itself alone, and gives (to man) a universal law which we call the Moral Law" [Kant, 1889, p. 120]. Basically, there is a certain way of the engagement related to the pure will and pure practical reason. The phenomenon can be happened, then, while the self does not engaged in the way the pure will requires.

The certain way of engagement can be clarified by focusing once again on Kant's Second Critique. Recalling the memorably statement of Kant's ethical doctrine that there is a connection between pure practical reason and pure will. Kant explains the connection as follows:

"The voice of reason in reference to the will is so clear, so irrepressible, so audible even, to the commonest men" [Kant, 1889, p. 124]. As an example,

"the self who has lost at play may be vexed at himself and his folly, but if he conscious of having cheated at play(although he has gained thereby), he must despise himself as soon as he compares himself with the moral law". Kant concluded, then, that "there is something further in the idea of our practical reason, which accompanies the transgression of a moral law-namely, its ill desert [Kant, 1889, p. 127]. Indeed, the connection exists even if the self never takes it into account.

Self-deception can be explained by the language of the engagement which the pure will requires:

The individual should be engaged in a certain way which the pure will requires; yet (s)he does not speak of the engagement while (s)he is hiding the clear voice from him(her)self. The engagement seems to exist in a certain way even if the individual does not take it into account. Therefore, the individual is deceiving him/herself.

This new way of understanding self-deception has its own prominences. Firstly, the cases of irrationality can be changed to rational oneseither by the desires' influences or the language of engagement. Consequently, there is no need to suppose that the cases of irrationality are the result of a conflict between differentiated parts of the mind.

Secondly, as any type of desires can be related to the empirical will, it may actually differ from one person to another. Unlike all Freudian-like accounts, which presuppose that the levels of mind are universal, the suggestion enables us to take the differences among people seriously. If so, the moral responsibility will be put on the self's shoulder while he/she, him/herself, can examine whether the phenomenon happens.

Finally, the philosophical paradox of self-deception can be solved philosophically. The ordinary life which can experience challenges between two different desires never considers them as a paradox in the self. Additionally, not taking into account the certain engagement which the pure will requires cannot be considered as a paradox either.

If so, it is possible to remain faithful to ethical, philosophical heritage for defining the self-deception on one hand, and solve the paradox, on the other.

BIBLIOGRAPHY

1. Aristotle (1934): Aristotle, Nicomachean Ethics in Aristotle in 23 Volumes, Vol. 19, translated by H. Rackham. Cambridge, MA, London: Harvard University Press William Heinemann Ltd. 1934, 420 p.
2. Audi (1982): Audi, R., Self-Deception, Action, and will. *Erkenntnis* 18, 1982.
3. Bach (1981): Bach, K. (1981) An Analysis of Self-Deception Philosophy and Phenomenological Research 41, 1981, pp. 351–370.
4. Canfield (1962): Canfield, J. &Gustavson D. Self-Deception. Anlyis 23, 1962, pp. 32–36.
5. Davidson (1982): Davidson, Donald, 1982- paradoxes of Irrationality in R. Wollheim& J. Hopkins, (eds.) Philoso-phical Essays on Freud, Cambridge: Cambridge University Press, 1982, pp. 289–305.
6. Demos (1960): Demos, Raphael, Lying to Oneself. Journal of Philosophy 75, 1960, pp. 588–95.
7. Deweese–Boyd (2012): Deweese–Boyd Ian. Self–Deception. The Stanford Encyclopedia of Philosophy (Fall 2010 Edition), Edward N. Zalta (ed.) URL = <http://plato. stanford. edu/archives/spr2012/entries/self–deception/>.
8. Fingarette (1982): Fingarette, Herbert, "Self-deception and the "splitting of the ego"" in Philosophical essays on Freud, edited by Richard Wollhein and James Hopkins, Cambridge University Press, 1982, pp. 212–227.
9. Freud (1923): Freud, S. "The ego and the id. In J. Strachey (Ed. & Trans.), The standard edition of the complete psycho-logical works of Sigmund Freud (vol. 19, pp. 3–66) London: Hogarth. (original work published in 1923), pp. 3–66.
10. Gilovich (1991): Gilovich, T. How We Know What Isn't So. New York: Macmillan, 1991, 216 p.
11. Greenwald (1997): Greenwald, Anthony G. Self-Knowledge and self-deception: further consideration, in The Mythomanias: The Nature of Deception and Self-Deception, Michael S. Myslobodsky, Lawrence Erlabum Associates, publishers Mahwah, New Jersey, 1997, pp. 51–72.
12. Hirstein (2005): Hirstein, William, Brain Fiction: Self-Deception and Riddle of confabulation, The MIT Press, 2005, 242 p.
13. Johnston (1988): Johnston M., Self-deception and the nature of mind, Perspectives on Self-deception, eds. B. P. McLaughin and A. O. Rorty. Berkeley: University of California Press, 1988, 345 p.
14. Kant (1889): Kant, Immanuel, The Critique of Practical Reason in "Kant's Critique of Practical Reason, etc." translated by Thomas Kingsmill Abbot, Longmans, Green,& Co., Paternoster-Row, forth edition,1889, pp. 101–244.
15. McNaughton (1991): McNaughton, David, Moral Vision; An Introduction to Ethics, Blackwell, 1991.
16. Mele (1987): Mele, Alfred, Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control, Oxford: Oxford University Press, 1987.
17. Mele (2001): Mele, Alfred R., Self-Deception Unmasked, Princeton University Press, 2001.
18. Mike (2001): Mike W. Martin, Self-deception in Lawrence C. Becker and Charlotte B. Becker, Encyclopedia of Ethics, Rutledge, 2nd. ed., vol. 3, 2001.
19. Otto (1936): Otto, R, The Idea of the Holy: an Inquiry into the Non-rational Factor in the Idea of the Divine and its Relation to the Rational, translated by John W. Harvey, Oxford University Press, 1936, 239 p.
20. Pears (1982): Pears, David, Motivated irrationality, Freudian theory and cognitive dissonance in R. Wollheim& J. Hopkins, (eds.) Philosophical Essays on Freud, Cambridge: Cambridge University Press, 1982, pp. 279–288.
21. Plato (1997): Plato, Phaedrus, in Plato, complete Works, edited by John Cooper, Hackett Publishing Company, Indianapolis, Cambridge, 1997.
22. Pugmire (1969): Pugmire, D. Strong' Self-Deception. Inquiry 12, 1969, pp. 339–349.
23. Rorty (1994): Rorty, Amelie O. User-Friendly Self-Deception, Philosophy, Vol. 69. 1994, 25 p.
24. Sartre (1966): Sartre, Jean-Paul. Being and Nothingness. Translated by Hazel E. Barnes. New York: Washington Square, Classic discussion in chapter on "Bad Faith". 1966, 202 p.
25. Siegler (1968): Siegler, F. (1968). An Analysis of Self-Deception. 1968, Nous 2, pp. 147–164.