

*Copula*  
**Jahangirnagar University Studies in Philosophy**  
Vol. XXXV June 2018



**Department of Philosophy**  
**Jahangirnagar University**  
Savar Dhaka Bangladesh

## **An Evaluation of Derk Pereboom's Four-Case Argument**

**Mostofa Nazmul Mansur\***

**Abstract:** Hard incompatibilism is a view which asserts that determinism and free will are inconsistent and given the facts of our best sciences determinism is true; and hence, free will does not exist. Not only that, it also claims that if the world were indeterministic and our actions were caused by states or events, still we would lack free will. In this way, it denies the truth of any libertarian account of free will based on event causation. In that sense, this is a hard position. Regarding moral responsibility, this hard incompatibilism claims that human agents are not morally responsible for their actions, unless they are the ultimate originators (agent-causes) of the actions in question. Derk Pereboom has offered an argument in favor of his hard incompatibilism, aka hard determinism, in which he shows four different hypothetical situations (thought experiments, indeed) each of which is aimed to prove that we are not morally responsible for what we do. The argument is, thus, known as the **four-case argument**. In the present paper, Pereboom's four-case argument is examined and defended. It has been shown here that Pereboom is quite correct, if we consider the 'true sense' of the term "moral responsibility." This 'true sense' of the term "moral responsibility" is considered as the **strong sense** of moral responsibility in this paper. However, a **weak sense** of the term "moral responsibility" has also been proposed in this paper. This proposed weak sense of "moral responsibility" can accommodate most of the socially-approved ways of ascriptions of responsibility. And finally, it has been claimed that our general intuition that every event has a cause and we are not the ultimate sources of our actions is true from the strong sense of moral responsibility; and our commonsense intuition that as human beings we are inherently free and responsible for our actions is true from the weak sense of moral responsibility.

The debate between compatibilism and incompatibilism is one of the most interesting topics in ethics. Compatibilism is a thesis that

---

\* **Dr. Mostofa Nazmul Mansur**, Professor, Department of Philosophy, Jahangirnagar University.

asserts that determinism is compatible with free will. On the other hand, incompatibilism is a thesis that asserts that determinism and free will are incompatible to each other. Of course, most of the incompatibilists, such as libertarian incompatibilists (agent-causal and event-causal), believe that determinism is false, and free will exists. Since free will exists, there is a scope for moral responsibility. There is another view, hard incompatibilism, recently introduced and defended by Derk Pereboom. Hard incompatibilism is a view which asserts that determinism and free will are inconsistent and given the facts of our best sciences determinism is true; and hence, free will does not exist. Moreover, it also claims that if the world were indeterministic and our actions were caused by states or events, still we would lack free will. In this way, it denies the truth of any libertarian account of free will based on event causation. In that sense, this is a hard position.<sup>1</sup> Regarding moral responsibility, this hard incompatibilism claims that human agents are not morally responsible for their actions, unless they are the ultimate originators (agent-causes) of the actions in question.

To defend his view Pereboom explains why compatibilism and event-causal libertarian incompatibilism are not adequate accounts of moral responsibility. He devised his well-known **Four-Case Argument** to undermine any prominent compatibilist account. The method of his attack is simple and interesting. Pereboom takes all the prominent compatibilist accounts into consideration, then, diagnoses what the sufficient conditions for moral responsibility are, according to these accounts respectively. First, he considers Hume-Ayer's account of compatibilism. He finds that an agent, according to Hume-Ayer's account, is free and hence morally responsible for her actions despite the truth of determinism, if the agent acts not under constraint, but voluntarily.<sup>2</sup> Then, Pereboom considers Frankfurt's view. According to this view an action of an agent is free and, hence, worthy of moral assessment, if the agent's first order desire that results in the action in question conforms to that agent's second order desires.<sup>3</sup> Pereboom, next, considers Fischer and Ravizza's compatibilist account. This account holds that an agent's action, in a deterministic world, is free and hence worthy of moral assessment, if the agent acts in a 'moderately reasons-responsive' way. In addition,

the agent has to have taken responsibility for the springs of his action.<sup>4</sup> And, finally, Pereboom considers Wallace's account. According to this view, an agent is morally responsible for her action if the agent has the powers of reflective self-control. Even if the agent loses this power in a particular circumstance, still that agent is morally responsible because he may retain the power.<sup>5</sup> Now, taking the conjunction of these conditions, Pereboom constructs a series of hypothetical cases, which is known as Four-Case Argument, where an agent fulfills all the compatibilist conditions, but still is not morally responsible for her relevant action. In this way, Pereboom tries to prove that these compatibilist accounts are false. We can now consider the cases:

**Case 1:** Professor Plum was created by neuroscientists, who can manipulate him directly through the use of radio-like technology, but he is as much like an ordinary human being as is possible, given this history. Suppose these neurologists "locally" manipulate him to undertake the process of reasoning by which his desires are brought about and modified—directly producing his every state from moment to moment. The neuroscientists manipulate him by, among other things pushing a series of buttons just before he begins to reason about his situation, thereby causing his reasoning process to be rationally egoistic. Plum is not constrained to act in the sense that he does not act because of irresistible desire—the neuroscientists do not provide him with an irresistible desire—and he does not think and act contrary to character since he is often manipulated to be rationally egoistic. His effective first-order desire to kill Ms. White conforms to his second-order desires. Plum's reasoning process exemplifies the various components of moderate reason-sensitiveness. He is receptive to the relevant pattern of reasons, and his reasoning process would have resulted in different choices in some situations in which the egoistic reasons were otherwise. At the same time, he is not exclusively rationally egoistic since he will typically regulate his behavior by moral reasons when the egoistic reasons are relatively weak—weaker than they are in the current situation.<sup>6</sup>

Now, we see, in Case 1, Professor Plum satisfies Hume-Ayer's condition of moral responsibility since he 'is not constrained to act in the sense that he does not act because of an irresistible desire.' Professor Plum also satisfies Frankfurt's condition of responsibility since his 'first-order desire to kill Ms. White conforms to his second-

order desires.' Plum satisfies Fischer and Ravizza's condition too since his 'reasoning process exemplifies the various components of moderate reason-responsiveness.' And, finally, Plum satisfies Wallace's condition since he, Plum, 'at the same time is not exclusively egoistic' and he typically regulates his behavior by moral reasons when the egoistic reasons are weaker than they are in the situation in question. That is, Plum has the powers of reflective self-control, even if he did not have it at the time of his relevant action, he could retain that power when his egoistic reasons got weaker than the situation in question. Thus Professor Plum satisfies all the conditions of prominent compatibilist theories. Nevertheless, he is, intuitively, not morally responsible for his act of killing Ms. White because his action is determined by the neurologists' activities, which are beyond his control. Pereboom believes that even compatibilists share the intuition that Professor Plum is not responsible for his action in this case.

Now Case 2 comes in. Here, also, Professor Plum is covertly manipulated to kill Ms. White, but unlike Case 1, the manipulation occurs as a result of programming:

**Case 2:** Plum is like an ordinary human being, except that he was created by neuroscientists, who, although they cannot control him directly, have programmed him to weigh reasons for action so that he is often but not exclusively rationally egoistic, with the result that in the circumstances in which he now finds himself, he is causally determined to undertake the moderately reason-responsive process and to process the set of first—and second-order desires that results in his killing Ms. White. He has the general ability to regulate his behavior by moral reasons, but in these circumstances, the egoistic reasons are very powerful, and accordingly he is causally determined to kill for these reasons. Nevertheless, he does not act because of an irresistible desire.<sup>7</sup>

It can be shown, again, in Case 2, that Plum satisfies all the compatibilist conditions. Nevertheless, Plum, intuitively, is not responsible for his germane action since his action is determined by the neuroscientists' programming which is beyond his control. In fact, if Plum is not responsible in the Case 1, he is not responsible in the Case 2 either, because there is no significant difference, regarding responsibility, between these two cases.

Case 3 presents a scenario that is comparatively more similar to the ordinary situation. Pereboom calls it a 'near-normal' case. Here the role of neuroscientists is replaced by the rigorous training practices of the agent's early age:

**Case 3:** Plum is ordinary human being, except that he was determined by the rigorous practices of his home and community so that he is often but not exclusively rationally egoistic (exactly as in Case 1 and 2). His training took place at too early an age for him to have had the ability to prevent or alter the practices that determined his character. In these current circumstances, Plum is thereby caused to undertake the moderately reason-responsive process and to process the first—and second-order desires that result in his killing White. He has the general ability to grasp, apply, and regulate his behavior by moral reasons, but in these circumstances, the egoistic reasons are very powerful, and hence the rigorous training practices of his upbringing deterministically result in his act of murder. Nevertheless, he does not act because of an irresistible desire.<sup>8</sup>

Again, Plum fulfills all the prominent compatibilist conditions of responsibility. In this case, Pereboom does not claim that it is intuitively obvious that Plum is morally responsible. Instead, he argues that the replacement of 'programming' by 'rigorous training practices of the agent's early age' does not introduce any responsibility-bearing content in the agent's action. So, regarding responsibility, there is no significant difference between Case 2 and Case 3. Causal determinism—Pereboom says—that explains Plum's lack of moral responsibility forces us to say that Plum is not morally responsible in Case 3 too for the same reason.<sup>9</sup> Pereboom invites compatibilists to show any significant responsibility-introducing feature to explain why Plum is responsible in Case 3, but not in Case 2, if they are unwilling to accept that Plum is not morally responsible in Case 3. And Pereboom believes that compatibilists cannot show it because there is no such feature at all.<sup>10</sup>

From this near-normal scenario described in Case 3, Pereboom generalizes his idea to the normal case:

**Case 4:** Physicalist determinism is true, and Plum is an ordinary human being generated and raised under normal circumstances who is often but not exclusively rationally egoistic (exactly as egoistic as in Case 1-3). Plum's killing of White comes about as a result of his undertaking the moderately reason-responsiveness

process of deliberation, he exhibits the specified organization of first—and second-order desires and he does not act because of an irresistible desire. He has the general ability to grasp, apply, and regulate his behavior by moral reasons, but in these circumstances the egoistic reasons are very powerful, and together with background circumstances they deterministically result in his act of murder.<sup>11</sup>

Plum, like in Case 1-3, satisfies all the prominent compatibilist conditions. Like Case 3, Pereboom does not claim here that Plum's non-responsibility in Case 4 is intuitively obvious. Instead, he, again, challenges compatibilists to show any responsibility-introducing feature in Case 4 to explain why Plum is morally responsible in this case while he is not responsible in case 3. Pereboom believes that there is no such feature. So, he holds the view that 'the best explanation for our intuition is that Plum's non-responsibility in Case 1 generalizes all the way down to Case 4.' But Case 4 represents the normal situation given that the world is deterministic. Thus, no agent is ever morally responsible for her actions in a deterministic world. Hence Pereboom claims that compatibilism is untenable. The structure of this Four-Case Argument can be re-expressed in the following way:

**The Manipulation Argument**

1. Manipulated *S* does not freely *A* and is not morally responsible for *A*-ing.
2. Regarding free action and moral responsibility, there is no relevant difference between manipulated *S*'s *A*-ing and any action deemed to be true and for which its agent is morally responsible on any compatibilist account of free action and moral responsibility.
3. If (1) and (2), then no compatibilist account of free action and moral responsibility is true.
4. Therefore, no compatibilist account of free action and moral responsibility is true.<sup>12</sup>

Now, the question, regarding the **Four-Case Argument**, is: what makes Plum, especially in Case 1 and 2, intuitively, non-responsible? It is clear that in the first two cases Plum is not responsible for his germane action because his action is caused by a deterministic causal process arranged by manipulators which is beyond his control. This

lack of control makes Plum non-free and, hence, non-responsible for his germane action. Premise-1 of the manipulation argument is founded on this intuition. Cases 3 and 4 show that there is no significant difference, regarding the control over the action in question and hence regarding free action and moral responsibility, between cases 1-2 (manipulated situations) and cases 3-4 (near-normal and normal situations). Hence, regarding free action and responsibility, what is true of cases 1-2, is true of cases 3-4. Premise-2 of the manipulation argument rests on this discovery. Premise-3 is an assumption about the truth of the premise 1-2. And from premises 1-3 the conclusion is deduced validly. Though the argument is about compatibilist accounts of free action and responsibility, the whole argument displays a more general feature regarding control, free action and responsibility in general. This general feature is: if an agent's action is caused by a factor that is beyond the control of the agent, then the agent is not morally responsible for the action. This idea gives birth to a principle—**Principle O**. Pereboom expresses this principle in the following way:

(O) If an agent is morally responsible for her deciding to perform an action, then the production of this decision must be something over which the agent has control, and an agent is not morally responsible for the decision if it is produced by a source over which she has no control.<sup>13</sup>

The upshot of Principle O is that it shows that, like determinism, any event-causal libertarian account cannot accommodate free will and hence fails to prove that human agents are morally responsible for their actions. Pereboom claims that an indeterministic event-causal history, which does not include agent-causation, is not relevantly different from the manipulated one described in his Four-Case Argument. This is because, like the manipulation case, in case of indeterministic causal history, the ultimate source of the agent's action is something beyond the control of the agent in question. And hence, by Principle O, the agent is not morally responsible for her action—in case of an indeterministic event-causal history—because of the lack of control over the action the agent does. Principle O suggests that an agent is morally responsible for her action if she has the control over the production of the relevant action. That is, to be morally responsible for one's actions, one must be the ultimate



originator of one's decisions and actions. In other words, according to Pereboom, only agent causation can accommodate free will and responsibility. But Pereboom thinks that our best scientific theories do not give us any reason to believe that agent-causation is true. So, he claims that we should favor hard incompatibilism and accept the fact that no one is ever morally responsible for anything that one does.

Not surprisingly, Pereboom's Four-Case Argument has received criticisms from many. John Martin Fischer claims that Plum is responsible for his action of killing Ms. White, though he is not blameworthy for his germane act. He distinguishes moral responsibility from praiseworthiness and blameworthiness. In his consideration, moral responsibility is a more abstract idea than praiseworthiness and blameworthiness, which is—he says—the *gateway* to moral praiseworthiness, blameworthiness, resentment and so forth. Fischer writes:

... an agent may be morally responsible for morally neutral behavior. Further, an agent can be morally responsible, but circumstances may be such as to render praise or blame unjustified. Once the distinction between moral responsibility and (say) blameworthiness is made, it is natural to suppose that Professor Plum is morally responsible for killing Ms. White, even if is not blameworthy (or not fully blameworthy) for doing so.<sup>14</sup>

I, of course, agree with Fischer's view that moral responsibility does not necessarily entail praiseworthiness or blameworthiness. It is true that sometimes we are morally responsible for some of our actions even though we are not blameworthy or praiseworthy for those actions. But it is not clear to me how Professor Plum of the Four-Case Argument is morally responsible for killing Ms. White but not blameworthy for that action. After all 'killing a person' is not a morally neutral action. On the other hand, if in the given circumstances Plum is not blameworthy for killing Ms. White, then Fischer needs to show how in the same circumstances Plum is morally responsible for the same action. Unless and until Fischer gives us a better explanation of why Plum is not blameworthy despite being responsible for killing Ms. White, I do not see any reason to accept Fischer's verdict rejecting Pereboom's Four-Case Argument.

McKenna claims that the manipulated agent, such as Plum in cases 1 and 2 along with cases 3 and 4, is morally responsible despite the manipulation. This line of attack is known as 'hard-line reply.'<sup>15</sup> The strategy is: first we need to improve the manipulation cases in such a fashion that it becomes transparent that the manipulated agent satisfies all the compatibilist conditions of the prominent compatibilist contenders. Now, if the compatibilist conditions are independently plausible, then the manipulated agent should be regarded as responsible for the relevant action by virtue of the compatibilist conditions despite manipulation. That is, if we accept compatibilist conditions, we cannot regard the manipulation as menacing.<sup>16</sup> This strategy seems to me unacceptable. We must note that the Four-Case Argument is launched to show that Plum is not morally responsible for his relevant action despite the fact that he satisfies the compatibilist conditions. Now, if, in reply, compatibilists argue that if Plum satisfies compatibilist conditions, then he is responsible for his germane action, then the debate will end in a stalemate situation with which we cannot proceed further.

Instead of proceeding with this stalemate situation, I, rather, consider the idea of 'magical agents' that is introduced by Ishtiyaque Haji to undermine the Principle O, the outcome of the Four-Case Argument. The story runs this way:

These [magical agents] are individuals who spring into existence in an unusual fashion, but who are otherwise very much like normal, healthy, adult human being. ... Imagine that Rosa is such a magical agent: she was "born" an instant ago, and with the exception of her unconventional existence into life, she enjoys the full complement of features that morally responsible agency demands. She hears about the plight of the children in Niger. Whipping out her "magical wallet" full of large denomination bills, she makes a bountiful donation to a well-reputed, pertinent charity. Assume, in addition, that Rosa has acted in the very fashion, down to minute details, in which Roselle, a conventional, morally sensitive, adult human being with a "normal" upbringing, has just done: Roselle, just like Rosa, has contributed to this well-reputed charity. Bracketing special pressure from the direction of determinism or indeterminism, assume that Roselle is morally praiseworthy for having made this donation. If Roselle is praiseworthy for this deed, then it seems that Rosa, too, should be praiseworthy for her similar deed. This

is because Rosa is relatively just like Roselle; she is Rosa's twin in the pertinent respect.<sup>17</sup>

Haji suggests that there is no responsibility relevant difference between manipulated Plum and magically created Rosa. But, intuitively, Rosa's making decision is a free action for which she bears responsibility. Principle O, though, generates that Rosa is not morally responsible for her act of charity since she has no control over the source of her spring of action. This does not go with Haji's intuition. So, Haji thinks that there is something wrong in the Principle O, in his words, "perhaps Principle O is the culprit."<sup>18</sup>

I have two observations concerning Haji's idea of 'magical agents' and his reading of Principle O. First, it seems to me that Haji considers Principle O as the background assumption of Pereboom's Four-Case (Manipulation) Argument. Haji explicitly says that the line-1 (premise-1) of the manipulation argument rests on the Principle O.<sup>19</sup> My reading of the Four-Case Argument and the Principle O is slightly different than Haji's. My observation is that Pereboom never appeals to the Principle O in constructing the Four-Case Argument. He, rather, appeals to our general intuition in the first two cases. And, in the last two cases he shows that there is no responsibility-relevant difference between first two and last two cases. Hence, the agent is not morally responsible for his germane action in the last two cases since the agent is not morally responsible in the first two cases. Principle O comes up on the basis of the success of the Four-Case Argument. The following section from Pereboom's text supports this observation:

The best explanation for the intuition that Plum is not morally responsible in the first three cases is that his action results from a deterministic causal process that traces back to factors beyond his control. Because Plum is also causally determined in this way in case 4, we should conclude that here too Plum is not morally responsible for the same reason. More generally, if an action results from a deterministic causal process that traces back to the factors beyond the agent's control, then he is not morally responsible for it.<sup>20</sup>

Thus it seems that Pereboom's overall discussion of the Four-Case Argument gets him to the Principle O. In other words, Principle O is the logical outcome of the Four-Case Argument, not the background

assumption of the argument. Since the Principle O is the logical outcome of the Four-Case Argument, one has to accept it, if one buys the Four-Case Argument. So, What Haji needs is to attack the Four-Case Argument, not to condemn the Principle O.

Second, Haji, like Pereboom, appeals towards general intuition to show that Roselle is responsible for her action. And, since there is no responsibility-relevant difference between Roselle and the magical agent Rosa, Rosa is also morally responsible for her germane action. This argument is almost similar to Pereboom's argument. Haji just turns it upside down. Pereboom goes from the non-responsibility of the manipulated agent (Plum of Case 1) to non-responsibility of the agent of the normal situation (Plum of Case 4), while Haji goes from responsibility of the agent of normal situation (Roselle) to the responsibility of the manipulated agent (Rosa). But the problem is that while Pereboom begins with a relatively 'uncontroversial' intuitive position, Haji begins with an intuition which is controversial to many. Recall the first two cases of the Four-Case Argument: incompatibilists and compatibilists including Haji agrees that Plum, intuitively, is not responsible for his killing of Ms. White.<sup>21</sup> Even McKenna needs to 'improve' the first two cases to show that Plum is responsible.<sup>22</sup> But a supporter of the hard determinism may not share Haji's intuition. About this sort of intuition Pereboom says:

... it is specified that determinism is true, but ordinary intuitions are likely to persist regardless of this stipulation, especially if the implication of determinism are not thoroughly internalized. If we did assume determinism and internalize its implications, our intuitions might well be different.<sup>23</sup>

Now, if Haji still sticks on this sort of intuition, then we will fall in a stalemate situation again. This stalemate situation, however, does not undermine Pereboom's Four-Case Argument and Principle O as well.

One way to get rid of this stalemate is to appeal to our best sciences and find necessary conditions for moral responsibility. Our best sciences favor determinism. Everything that happens, according to our best sciences, has a cause behind it. That cause, again, is caused by another cause. In this way the whole universe is bonded with a deterministic causal chain of states or events. There is no scope of agent causation. Agents may have desires, beliefs, attitudes

and so on which cause the agents' actions. But the causes of their desires, beliefs, attitudes and so on rest on some other states or events that are beyond the agents' control. Even if we believe in the sort of indeterminism which is posited by some current interpretations of quantum mechanics, we will find no feature that goes in favor of agent-causation. If there is any uncertainty in quantum level or micro level that does not entail that human agents have control over the relevant causes of their actions. If I flip a coin, then there is always uncertainty whether heads or tails will come up. But that does not mean that I have any control over the event of heads' or tails' coming up. If I want heads to come up, and it happens that heads comes up, still it does not mean that my wanting of heads to come up has any control on the event of heads' coming up. Analogously, there might be some uncertainty just before I make some decision; but that does not mean that I have control over that decision, because my own desires, beliefs, attitudes and so on, by which I finally make the decision after the claimed uncertainty, are event-caused by the distant past and laws over which I do not have any control. So, there is no scope of agent-causation, even if indeterminism is true. Agents are never the ultimate sources of their actions. On the other hand, an agent is morally responsible for her action only if she has ultimate control over her actions. It is unjustified to hold someone responsible for an action over which one has no control. So, this control or ultimate origination is a necessary condition for moral responsibility.<sup>24</sup> Now, the conjunction of the facts of our best sciences (deterministic or indeterministic) and the ultimate origination as the necessary condition for responsibility shows that human agents are not ever responsible for their actions. That means that the conclusion drawn by Pereboom's Four-Case Argument along with his Principle O, is supported by our best sciences. This support from our best sciences may be of some help to break the stalemate situation and to get the victory over compatibilist and event-causal libertarian accounts regarding moral responsibility.

But most of our ordinary people are unwilling to accept this sort of hard incompatibilism. They think that if we accept hard incompatibilism, then we are required to reject all the morality, values, laws, concepts of good and bad, right and wrong etc. This is

not a right place to discuss it in detail. But it must be mentioned here that these ordinary ideas are not true. If there is no moral responsibility, still there will be good actions and bad actions, right and wrong, and other sorts of moralities. In a moral responsibility free world, there will be social values and even laws; we won't need to empty our jails. In short, everything will remain almost the same in a responsibility free world. So, a moral responsibility free world won't be an unlivable world.

I could conclude here claiming the victory of the Four-Case Argument as well as hard incompatibilism and a responsibility free world. But one true story illustrated in Lynn Baker's paper "Moral Responsibility without Libertarianism" forces me to think twice before I conclude saying that moral responsibility is a mere illusion. Here is the story:

In January 2001, two small-town teenage boys wantonly and brutally murdered two well-respected Dartmouth University Professors. The motive was unclear. There was some talk of Robbery; there were surmises about the thrill of killing. The boys did not personally know the victims. In court, the older of the two, who, by all accounts was the leader of the operation, showed no remorse at all. He sat stone-faced and sullen. The slightly younger boy wept and managed to address the children of the victims with an apology.<sup>25</sup>

What is the basic difference, regarding morality, between these two boys? Well, the first boy does not accept moral responsibility while the second one does. Is it not the case that by all morally relevant notions the second boy is morally superior to the first boy? If so, then what does make him morally superior? It is his acceptance of moral responsibility which makes him morally superior to the first boy. If it is so, then we should not say that moral responsibility is a 'mere' illusion. By saying that moral responsibility should not be considered as 'mere' illusion, I am not changing my view regarding Pereboom's Four-Case Argument, Principle O or hard incompatibilism in general. I do agree with Pereboom's account regarding moral responsibility. He is quite correct, if we consider the true sense of the term "moral responsibility." I, now, call it strong sense of moral responsibility. In addition to the strong sense of moral responsibility, here, I propose a weak sense of moral responsibility,

which can accommodate most of the socially-approved ways of ascription of responsibility, e.g. it can accommodate the acceptance of the moral responsibility of the second boy of the true story illustrated above. I formulate the two senses of moral responsibility in the following way:

**Moral Responsibility (Strong Sense): Principle O**

**Moral Responsibility (Weak Sense):** *A* is responsible for her action *x* at *t*, if *x* is not necessary at *t* and *A* believes that she could have done otherwise at *t*.

Here, the first condition, '*x* is not necessary at *t*', precludes the agent to be considered as morally responsible for her relevant action when the agent is 'directly' manipulated, *i.e.* manipulated by another agent's actions, such as neuroscientists or programming arranged by any other agent. This sort of manipulation necessitates *A*'s *x*-ing. Hence, the proposed weak sense of moral responsibility makes the agent free from bearing moral responsibility when the manipulation is menacing. And if the manipulation is benign, then the agent is worthy to be considered morally responsible for his germane action, since a benign manipulation does not necessitate the agent acting in a certain pattern. The second condition, '*A* believes that she could have done otherwise at *t*' does not require the existence of an alternative. It simply requires a belief in agent that she has the ability to do otherwise. This condition makes an agent morally responsible in the normal situation. In a normal situation the agent believes that she could have done otherwise instead of doing so-and-so. Moreover, in a normal situation, the action *x* of the agent *A* at time *t* is not necessary in the sense that the agent's various beliefs, her fear to be considered as blameworthy, her eagerness to be considered as praiseworthy and so on might work as factors of her causal history that could change the pattern of the agent's relevant action at *t*. Taking these two senses of moral responsibility in mind, if we look back to the Four-Case Argument, we see that Plum is not responsible (both from strong and weak sense of responsibility) for killing Ms. White in first two cases because the manipulators' action or programming necessitate Plum's action of killing of Ms. White. But in the last two cases, if we consider the weak sense of moral responsibility, we see Plum is morally responsible for his act of killing since in these near-normal and normal situations Plum has the

belief that he could have done otherwise, and Plum's germane action is not necessary given the above interpretation of necessitation. Plum, of course, is not morally responsible in these last two cases, if we consider responsibility in its strong sense. Considering moral responsibility in this way saves this idea of responsibility from being treated as 'mere' illusion.

Finally, we, the ordinary people, have two apparently inconsistent commonsense intuitions regarding responsibility. On the one hand, we have general intuition that every event has a cause and we are not the ultimate sources of our actions (and hence we are not responsible for our action). On the other hand, we have commonsense intuition that as human beings we are inherently free and responsible for our actions. I believe that the both of these commonsense intuitions are true. The first intuition is true from the strong sense of moral responsibility favored by hard incompatibilism. And the second intuition is true from the weak sense of moral responsibility that I have proposed in this paper.

---

#### References:

1. Haji, Ishtiyaque (2008), *Incompatibilism's Allure: Principle Arguments for Incompatibilism*, Broadview Press, P. 121.
2. Pereboom, Derk (2001), *Living Without Free Will*, Cambridge: Cambridge University Press, Pp. 101-02.
3. Ibid, P. 103.
4. Ibid, Pp. 108-09.
5. Ibid, Pp. 108-09.
6. Ibid, Pp. 112-13.
7. Ibid, Pp. 113-14.
8. Ibid, P. 114
9. Ibid, P. 115.
10. Ibid, Pp. 114-15
11. Ibid, P. 115.
12. Haji (2008), P. 120 (Shortened)
13. Pereboom (2001), P. 54.
14. Fischer, John Martin (2006), *My Way: Essays on Moral Responsibility*, New York: Oxford University Press, P.233
15. Haji (2008), P. 125.
16. Ibid, P. 126
17. Ibid, Pp. 133-34 (My bracket)
18. Ibid, P. 134
19. Ibid, P. 120.



20. Pereboom (2001), P.116.
21. Haji (2008), P. 126.
22. Ibid, P. 125
23. Pereboom (2001), P. 117.
24. It might be asked that if I accept 'ultimate origination' as a necessary condition for moral responsibility, then why do I not buy any compatibilist account of ultimate origination, instead of Pereboom's incompatibilist one. In reply, I would say, I buy Pereboom's account, because it is supported by our best sciences. The compatibilist accounts of ultimate origination, such as Frankfurt's hierarchical conception of ultimate origination or Fischer-Ravizza's compatibilist account of ultimate origination, are not satisfactory, indeed. In final sense, these are event-causal accounts. The higher-order desires (more precisely, the second-order desires), in Frankfurt's case, are not really "one's own", because these must be caused by some beliefs and attitudes which are, somehow, caused by some external events or states that are beyond the agent's control. Similarly, the cluster of beliefs, attitudes that form the agent's reason-responsiveness, in Fischer-Ravizza's case, are, scientifically speaking, caused by a historical chain involving distant past and laws which are beyond the control of the agent. The account of ultimate origination proposed by Galen Strawson is an incompatibilist account like Pereboom's. It appeals to infinite regress, but Pereboom's account gives us 'almost' same result without appealing to the problematic issue of infinite regress. Hence, I am inclined to accept Pereboom's idea of ultimate origination.
25. Baker, Lynne Rudder (2006), "Moral Responsibility Without Libertarianism", *Nous* 40:2, P. 323(Shortened)