

In Defense of P.F. Strawson's Account against a Metaphysical Challenge

1. A Big Picture

- (1) The purpose of this presentation is to address a traditional metaphysical worry regarding the relationship between determinism and freedom in Strawson's theory of moral responsibility.
- (2) Strawson's theory of moral responsibility: The separation of moral responsibility from metaphysical issues about freedom.
 - moral responsibility requires actual moral emotion (=reactive attitudes¹)
- (3) Problem: Strawson's notion of **excuses**² seems to involve a kind of freedom: **freedom to do otherwise**³.
 - It makes Strawson's theory of moral responsibility vulnerable to the **metaphysical challenge**⁴ pertaining to freedom.
- (4) Solution: I propose to revise excusing conditions in terms of a different kind of freedom based on **sourcehood**⁵, as conceived by **Harry Frankfurt**⁶.
 - i. The Revision of Excusing Pleas on the basis of Sourcehood (henceforth **REPS**⁷)
 - ii. Working under REPS, I claim that moral responsibility requires sourcehood. By arguing that sourcehood is compatible with determinism, I show that Strawsonian moral responsibility can be safeguarded against the incompatibilists' challenge.
- (5) Conclusion: There is a significant advantage to REPS over alternative possibilities: it can counter the metaphysical issue stemming from the relationship between determinism and freedom.

2. My Argument: A Defense of Strawson's Account against the Metaphysical Challenge

(P1) For any agent S and action A, S is not the source of A only if S's circumstances satisfy the excusing conditions.

¹ Reactive attitudes: We adopt reactive attitudes, such as resentment or gratitude, in response to the quality of an agent's will—a good or ill will with which they act; based on the reactive attitudes adopted, we hold the agent responsible.

² Excusing conditions: Excusing conditions are conditions in which the reactive attitudes are suspended and an agent is exculpated from moral responsibility for their action.

³ Freedom to do otherwise (=alternative possibilities): For any given agent S and action A, S is free to perform A only if S could have acted differently than A at the moment when S decided to do A

⁴ The metaphysical challenge as advanced by incompatibilists:

(M1) If determinism is true, then no one could have done otherwise.

(M2) If determinism is true, then all reactive attitudes are suspended or withheld.

⁵ Sourcehood: Sourcehood is commonly understood as an agent's being the source of their action. Here 'the source' refers to the expression of the agent's fundamental practical self (i.e., their deep self) or what they really want to do.

⁶ Frankfurt's sourcehood: Frankfurt defines sourcehood as an identification model (henceforth sourcehood), i.e. sourcehood is governed by motives internal to the self when an agent identifies with said motives.

For any agent S and action A, S is free to perform A only if S is the source of A—i.e., A is brought about by motives with which S is identified.

⁷ (1) Strawson's original excusing conditions: "He couldn't help it", "He was pushed", "He had to do it", "It was the only way", "They left him no alternative", etc. (Strawson 1962/2008: 7).

(2) REPS: "He couldn't help it because he is moved to it by forces that he is not identified with", "He acted against his own wishes", "He did what he did not really want to do because he lost his body balance through being pushed", "He hadn't realized that he was performing the action, although he knew the action could harm us", and so on.

(P2) For any agent S and action A, if S's circumstances satisfy the excusing conditions, then reactive attitudes in response to A could not serve as the basis of moral responsibility.

Putting P1 and P2 together, we arrive at the following claim: if S is not the source of A, then the reactive attitudes in response to A could not function as the basis for moral responsibility. The contrapositive of this claim brings us to the conclusion:

(C1) For any agent S and action A, reactive attitudes in response to A can serve as the basis of moral responsibility only if S is the source of A.

Let us consider whether Strawson's theory of moral responsibility based on REPS can address the metaphysical challenge.

The reconstruction of the incompatibilists' claim:

(M1') If determinism is true, then no one can be the source of their actions.

(M2) If determinism is true, then all reactive attitudes are suspended or withheld. (from M1' and C1)

I aim to refute the argument by showing that M1' is false. If I succeed in doing so, the incompatibilists' challenge would be neutralized. It would mean that REPS can indeed shield Strawson's theory of moral responsibility from the metaphysical worry. Thus, my focus here is on showing the plausibility of R1:

(R1) For any given agent S and an action A, even if determinism is true, S can be the source of A.

The plausibility of R1 derives from Frankfurt's hierarchical theory of desires.

i. Second order volitions⁸

ii. Freedom of the will=the Conformity between first order desires and second order volitions=Sourcehood:

When an agent's second order volitions are in accordance with their first order desires, they have freedom of the will—i.e., sourcehood. In Frankfurt's terms, Sourcehood can be defined as follows: an agent's action is brought about by the motives with which the agent identifies—i.e., by the conformity between their first order desires and their second order volitions.

iii. Freedom of action ≠ Freedom of the will

To show the plausibility of R1, I assume that even if determinism is true, the following two claims are still true:

(A1) Second order desires and volitions are independent of first order desires.

(A2) The fundamental self that is comprised of second order desires and volitions stands unthreatened by the truth of determinism.

Putting these two claims together, we arrive at the conclusion:

(A3) If the truth of determinism leads to human desires being causally necessitated, then although their first order desires are causally necessitated, their second order desires and volitions can nonetheless be free.

A possible objection:

⁸ **First order desires** refer to all the immediate and reflexive desires that arise within our minds; **second order desires** are desires for first order desires. If one translates first order desires into action, that comprises **one's will**; if one translates second order desires into action, it is called **second order volitions**.

Incompatibilists might deny A2. They believe that if determinism is true and everything which occurs in the world—including all human actions and desires—is necessitated by antecedent causal conditions, then one's fundamental self and thus second order desires and volitions must be causally necessitated as well.

The response to the objection:

I reject the incompatibilists' claim because if our second order desires and volitions must be causally necessitated, then it would lead to everyone losing their identity. I hold that second order desires and particularly volitions capture the essence of what it means to be a specific person.

- i. The identity as a person: the unwilling addict case
- ii. The lack of the identity as a person: the wanton case

If determinism is true, then all humans lose their freedom of action.

However, even if determinism is true, an agent can have freedom of the will. That is, even if determinism is true, an agent can be the source of their action when their causally necessitated action is in exact accordance with their true will—i.e., there is conformity between their first order desires and second order volitions.

Case 1. Jones and Black

Black wants Jones to vote for Biden. Jones does not know about it but Black carefully observes Jones and waits until Jones decides what to do. If Jones decides to do what Black wants him to do, then Black does nothing. However, if Jones decides to do something other than what Black wants him to do—i.e., to vote for Trump, then Black would intervene to force Jones to vote for Biden.

- i. In Case 1, Jones does what Black wants him to do in every circumstance
- ii. Consider a circumstance in which Black does not need to intervene because Jones decides to do what Black wants him to do.

For our purposes, now picture a case in which Black is replaced by the truth of determinism (Case 2).

Case 2. Jones and Determinism

Suppose that determinism is true and Jones's every action is causally necessitated by antecedent causal conditions. That is, it is determined that he will vote for Biden. But Jones is entirely unaware of this. So if Jones decides to do what is causally necessitated—i.e., to vote for Biden, then he performs it without any interference. However, if Jones decides to do what is not necessitated by the prior causal conditions and natural laws—i.e., to vote for Trump, then he could not do what he decided to do. This is because his action in voting for Biden has already been necessitated and thus he inevitably fails to resist it.

- i. All of Jones's actions and desires are causally necessitated as well.
- ii. Even if determinism holds, his first order desires are causally necessitated and yet his second order desires and volitions can nevertheless be free.
→ If this is right, then R1 seems plausible.

Since I showed R1 to be plausible, it follows that M1' and M2 are false. If we accept this line of reasoning, then we are led to the conclusion that the incompatibilists' claim does not threaten C1. It implies that REPS can defend Strawson's theory of moral responsibility against the metaphysical challenge regarding the relationship between determinism and freedom.