

GENERAL PHILOSOPHY

WEEK 7: FREE WILL

JONNY MCINTOSH

INTRODUCTION

The following three claims constitute an inconsistent set:

1. Determinism is true
2. If determinism is true, it is not the case that we sometimes act freely
3. We sometimes act freely

How should we respond?

Some terminology:

- **INCOMPATIBILISTS** endorse (2): if determinism is true, it is not the case that we sometimes act freely
- **COMPATIBILISTS** deny (2): even if determinism is true, it may nevertheless be that we sometimes act freely

Incompatibilists will deny either (1) or (3):

- **LIBERTARIANS** are incompatibilists who claim that we do sometimes act freely, and so reject (1)
- **HARD DETERMINISTS** are incompatibilists who claim that determinism is true, and so reject (3)

Note that compatibilists may **ALSO** reject (1) or (3), as they may think there are other reasons for rejecting determinism or free will.

- Compatibilists who accept (1) are sometimes called **SOFT DETERMINISTS**.

Rejecting (3) is particularly problematic given the plausible connection between free will and moral responsibility:

- S is **MORALLY RESPONSIBLE** for her action X only if S acted freely in performing action X

The plan for today:

- Classical Views
- Peter van Inwagen
- Harry Frankfurt
- Peter Strawson

CLASSICAL VIEWS

Compatibilists need to provide an account of free will on which, even if determinism is true, we may nevertheless sometimes act freely. How might they do that?

Here's a first attempt:

In performing action X, S acts freely IFF:

- S does what she **WANTS** or **CHOOSES**, and
- S is not compelled by any external force

(Compare a bank manager opening a safe of her own free will and doing so at gunpoint.)

Determinism does not entail that noone ever does what they want in the absence of compelling external force. So determinism is compatible with free will, so defined.

CLASSICAL INCOMPATIBILISM

But this doesn't work. Both a **KLEPTOMANIAC** and an ordinary **THIEF** may, in stealing something, do what they want in the absence of any compelling external force.

The obvious difference between the kleptomaniac and the ordinary thief is that only the thief **COULD HAVE DONE OTHERWISE**. So why not simply add this as a further condition to the account?

In performing action X, S acts freely IFF:

- S does what she **WANTS** or **CHOOSES**,
- S is not compelled by any external force, and
- S could have done otherwise

But doesn't determinism imply that noone could **EVER** have done otherwise?

Thus the **CLASSICAL INCOMPATIBILIST** argument:

1. S acts freely in performing X only if S could have done otherwise
2. If determinism is true, noone could ever have done otherwise
3. So if determinism is true, it is not the case that we sometimes act freely

Note the dialectic here:

- First, the compatibilist offers an account of free will on which it is compatible with determinism.
- The incompatibilist then responds by arguing that the compatibilist account doesn't give us everything that we want.
- Moreover, the incompatibilist suggests that what's missing is incompatible with determinism.

CLASSICAL COMPATIBILISM

We'll see in later that Frankfurt responds to the classical incompatibilist argument by denying the first premise. Classical compatibilists (like Ayer, and on many readings Hume) responded by denying the second.

They did this by offering an analysis of "S could have done otherwise" on which the second premise is false:

- S could have done otherwise IFF S would have done otherwise had S wanted or chosen to

Thus, the **CLASSICAL COMPATIBILISTS** account of free will ...

In performing action X, S acts freely IFF:

- S does what she **WANTS** or **CHOOSES**
- S is not compelled by any external force
- S would have done otherwise had S wanted or chosen to

This plausibly distinguishes the kleptomaniac from the ordinary thief. Even if kleptomaniacs want to **NOT** to steal something, they often steal it anyway. So it is often false to say of a kleptomaniac that, had she wanted or chosen not to steal something, she would have refrained from doing so. The thief by contrast, would have refrained if she had wanted or chosen not to.

And the account also seems to be consistent with determinism. Determinism does not entail that S could not have wanted or chosen to perform Y instead; it only entails that she could not have wanted or chosen to perform Y **GIVEN THE ACTUAL PAST**. So even if determinism is true, it may be that S would have performed Y had she wanted or chosen to do so.

A RESPONSE

But classical compatibilism won't work! As Roderick Chisholm, an incompatibilist, pointed out, it can be true that S would have done otherwise if she had wanted or chosen to even though she is **UNABLE** to want or chose to.

Suppose John is offered a choice between eating ice cream and eating biscuits and cheese. He eats the ice cream. But he has a rare psychological condition which means he is incapable of forming a desire to eat biscuits and cheese.

John does what he wants (eat ice cream), is not compelled by any external force, and presumably would have eaten biscuits and cheese **IF HE HAD WANTED OR CHOSEN TO DO SO**.

By the classical compatibilist's lights, then, John acted freely. But John was incapable of wanting or choosing to eat biscuits and cheese. So it seems that he does **NOT** act freely!

VAN INWAGEN

THE CONSEQUENCE ARGUMENT

If classical compatibilism doesn't give us everything we want, why should we think that what's missing is incompatible with determinism? Why think that, if determinism is true, noone could ever have done otherwise?

One attempt to answer this is Peter van Inwagen's **CONSEQUENCE ARGUMENT**.

Let P_0 be the proposition expressed by a statement describing the complete state of the world at a time t_0 , let P_1 be the proposition expressed by a statement describing how things are in some respect at any later time t_1 , and let L be the proposition expressed by a statement specifying the laws of nature.

According to van Inwagen, if determinism is true then it is necessarily the case that if P_0 and L then P_1 .

(This is a relatively weak claim. It does not attempt to say what determinism is, only what it entails, and makes no mention of causation.)

Now take an action that it seems to have been freely performed: your coming to the lecture today, for example. And let P be the proposition that the agent performed that action.

Then the argument runs as follows:

1. If you could have not come to the lecture today, you could have rendered P false
2. If you could have rendered P false, and determinism is true, then you could have rendered either P_0 or L false.
3. But you couldn't have rendered P_0 false.
4. And you couldn't have rendered L false either.
5. So if determinism is true, you couldn't have **NOT** come to the lecture today.

The argument thus seems to show that, if determinism is true, you could not have done otherwise. To the extent that there's nothing special about your action, the argument generalises to show that, if determinism is true, noone could ever have done otherwise.

A RESPONSE

If you didn't come to the lecture today, and determinism is true, all that follows is that either P_0 or L would have been false - not that you would have **RENDERED** them false.

Distinguish two claims:

- **WEAK:** There is a possible situation, W , in which you do something that is contrary to the laws of nature that actually hold.
- **STRONG:** There is a possible situation, W , in which you do something that is contrary to the laws of nature that hold in that situation, i.e. in W .

Does the consequence argument rests on an equivocation between these different ways of understanding the claim that you "could have rendered L false"?

In the **STRONG** sense, premise 4 is true: there is no possible situation in which you do something that is contrary to the laws of nature of **THAT** situation.

But premise 2 is false. If determinism is true, and there is a possible situation in which you don't come to today's lecture, all that follows is that there is a possible situation in which you do something that is contrary to the laws of nature that **ACTUALLY** hold.

In the **WEAK** sense, premise 2 is true: if determinism is true, any possible situation in which you don't come to today's lecture is one in which you do something that is contrary to the laws of nature that **ACTUALLY** hold.

But isn't premise 4 false? Are there not possible situations in which you do something that is contrary to the laws of nature that actually hold?

If this is right, there is no way of reading the Consequence Argument so that both premises 2 and 4 come out as true together.

- For more on this response, see David Lewis (1981) 'Are We Free to Break the Laws', reprinted in Gary Watson, ed., *Free Will* (OUP, 2nd edition).

FRANKFURT CASES

We saw that the classical incompatibilist argument rested in part on the assumption that S acts freely in performing X only if S could have done otherwise. Given the assumption that S is morally responsible for performing X only if S acts freely in performing X, it follows that:

- S is morally responsible for performing X only if S could have done otherwise

Harry Frankfurt calls this **THE PRINCIPLE OF ALTERNATE POSSIBILITIES** (or **P.A.P.**), and argues that it is false: sometimes an agent is morally responsible for performing an action even when she could **NOT** have done otherwise.

ASSASSINATION

Suppose Jones is poised to assassinate Smith. Black, who wants Smith dead, is monitoring the situation, looking out for a certain sign that Jones is going to decide to shoot - perhaps a kind of blush on Jones's face. If there is no sign by 2pm, Black will intervene and force Jones to shoot. Just before 2pm, Black sees the sign, and Jones shoots Smith.

According to Frankfurt, Jones is morally responsible for shooting Smith. But he could not have done otherwise: had he not decided to shoot, Black would have intervened, and ensured that he shot Smith anyway.

FRANKFURT CASES like this seem to be counter-examples to **P.A.P.**. To the extent that Jones acted freely in shooting Smith, they are also potential counter-examples to:

- S acts freely in performing X only if S could have done otherwise.

Debate over these cases is ongoing. Proponents of the so-called **DILEMMA DEFENCE** of P.A.P. argue:

- Either the sign that Jones is going to decide to shoot Smith is an infallible sign or it is not.
- If it is not, then Jones could have done otherwise; if it is, he is not morally responsible!

SIGNIFICANCE

Supposing it's successful, what is the **SIGNIFICANCE** of Frankfurt's challenge to P.A.P. for debates over compatibilism and incompatibilism?

It's not best understood as a response to van Inwagen's Consequence Argument.

At best, Frankfurt shows that free will/moral responsibility doesn't **ALWAYS** require an ability to do otherwise. But we might still think that an ability to do otherwise is required **MOST** of the time - whenever there's none of the odd, standby stuff going on.

It's better understood as simply showing that what's missing from standard compatibilist accounts of free will is **NOT** the ability to do otherwise, and so clearing the ground for an alternative positive account - such as the account Frankfurt himself offers in later work.

- For Frankfurt's positive account, see Harry Frankfurt (1971) 'Freedom of the Will and the Concept of a Person', reprinted in Gary Watson, ed., *Free Will* (OUP, 2nd edition).

STRAWSON

According to Peter Strawson, what's missing from standard compatibilist accounts of free will and moral responsibility is an account of the role that the practice of **HOLDING** someone morally responsible for her actions plays in the wider network of human attitudes and feelings.

Suppose I stand on your toe. You might feel **RESENTMENT** towards me. In doing so, you express the significance to you of the ill will or indifference towards you that my action appears to manifest.

You might also hold me **MORALLY RESPONSIBLE**. This is a generalized or vicarious analogue of resentment. It is not just that standing on your toe manifests ill will or indifference towards **YOU**; standing on **ANYONE'S** toe in such circumstances would be to manifest ill will or indifference towards **THEM**.

Both of these attitudes are what Strawson calls **REACTIVE** attitudes. Others include love, anger, gratitude.

- Your resentment is a species of **PERSONAL** reactive attitude: roughly, an attitude expressing the significance to you of the good or ill will, or indifference, towards you that my action seems to manifest.

- Your holding me morally responsible is a species of **MORAL** reactive attitude: roughly, an attitude expressing the significance to anyone suitably placed of the good or ill will, or indifference, towards such a person that my action seems to manifest.

What has all this got to do with the compatibilism/incompatibilism debate? Strawson identifies two kinds of **EXCUSING** circumstances: circumstances in which reactive attitudes are withdrawn or suspended:

1. The agent's action did not in fact manifest any ill will or disregard. (She didn't see your toe; she lost her balance; she has a compulsive disorder which makes her step on toes.)
2. The agent is psychologically or otherwise morally undeveloped or immature. (She is a child, or a rudimentary but cunningly disguised robot.)

- In the first sort of excusing circumstance, we continue to adopt a **PARTICIPANT** stance towards the agent, holding reactive attitudes towards her, but treating her action objectively.
- In the second, we take up an **OBJECTIVE** stance to the agent herself, no longer holding the full range of reactive attitudes towards her, and instead treating her as to be managed or treated.

Given all this:
FIRST: If determinism is true, it does not follow that every circumstance is one or other of these kinds of excusing circumstance. So even if determinism is true, it will often be appropriate to hold others morally responsible for their actions.

SECOND: But even if we had some theoretical reason to abandon the practice of holding others morally responsible, (a) we couldn't and (b) we wouldn't want to! Reactive attitudes are too central to our lives for us to be able to give them up.

- For further reading on this, connecting up with Hume on inductive and external world scepticism, see Strawson's (1985) *Skepticism and Naturalism: Some Varieties*.