

Incompatibilism without agent causation

PHIL 43503

Jeff Speaks

March 26, 2009

1 Three desiderata for a theory of free will

O'Connor lists three aspects of free will which a theory of free will should explain:

1. Agent-control: the fact that free actions 'flow from', or are under the control of, the agent.
2. Alternate possibilities: in free actions, alternate actions were genuinely open to the agent.
3. The confluence of the agent-control and alternate possibilities requirement: it must be explained how the agent controls which of the alternate possibilities in question is realized.

Compatibilist theories (like Frankfurt's) clearly satisfy the first condition, but have more trouble with the second. Incompatibilist theories

2 Two sorts of incompatibilist theories of free will

An incompatibilist theory of free will is a theory which says that actions are free if and only if they satisfy some condition C , where this is a condition which no action could satisfy in an indeterministic world.

We can divide incompatibilist theories of free will thus defined into two categories: those which do, and those which do not, make use of the notion of agent-causation.

O'Connor's concern is with incompatibilist theories of free will which do not make use of this notion.

These theories are further divisible into two categories.

Some of these theories explain free will without making use of any causal relations at all. This is what O'Connor calls 'simple indeterminism.'

Others make use of causal relations – just not agent-causal ones. O'Connor calls these varieties of 'causal indeterminism.'

3 Simple indeterminism

The idea that basic free actions are undetermined mental acts which have a certain ‘actish’ phenomenal quality. (To get at this phenomenal quality, think of the difference in ‘feel’ between saying a word to yourself and having the word pop into mind.)

The problem of the possibility of mental actions with this phenomenal quality being caused by an evil neurosurgeon.

4 Causal indeterminism

Causal indeterministic theories try to explain free action in terms of causal (but not agent-causal) relations involving those actions. These will typically be causal relations involving the mental states, like the beliefs and desires, of the agent.

The basic problem for causal indeterminism comes from the following plausible principle:

If an agent’s act was caused but not determined by his prior inner state, and if nothing besides that inner state was causally relevant to the act, then the agent had no choice about whether the inner state was followed by the act.

van Inwagen’s example of the thief and the poor box.

The problem is that if an agent has no choice about whether his inner states are followed by the action, and no choice about his inner states, it follows (from the no choice principle) that he has no choice about the action.

It seems that in making room for alternate possibilities, causal indeterminist theories sacrifice the agent-control which deterministic theories of free will provide. Or, more precisely, they fail the third, ‘confluence’ desideratum listed above.

A natural reply is that the relationship between an agent’s inner states and action is deterministic, so long as those inner states are taken to include mental events like an agent’s decision to give certain weights to some possible reasons for action. On this sort of view, the indeterminism comes not in the relation between these inner states and the action, but rather in the causal history of these inner states. One might take the ‘decision to assign weights’ to be undetermined, for example.

A worry is that this just pushes the problem back a step. Is the agent free with respect to the decision to assign weights to the various factors? If so, then the agent must have control over this decision (the second requirement above). But how is this to be explained?