

## Appendix

### Frankfurt-Case

The following is an example of a so-called “Frankfurt-case”, taken from HK. An enormous literature in philosophy is devoted to dealing with these kinds of examples, attempting to reconcile intuitions about responsibility with the counterfactual and causal features that these examples contain. Surprisingly, almost none of it uses causal models, and yet doing so reveals the causal structure to be entirely unproblematic.

**Example 7 (Frankfurt).** *Imagine Jones poisons Smith, who dies. Unbeknownst to Jones, Black was observing his behavior: if Jones had not poisoned Smith, Black would have given Jones a gun and manipulated him in some way or other so that Jones would shoot Smith. Black is both a perfect observer and manipulator of Jones’s behavior, and is thus guaranteed to succeed in his plans. Intuitively it is clear that Jones is responsible for Smith’s death, despite the fact that he could not have prevented it. (Typical Frankfurt cases focus on responsibility for an action, as opposed to responsibility for the consequence of an action, and therefore scenarios are normally formulated such that Black manipulates Jones to perform the same action. Except for the shift from the action to the consequence though, those cases are structurally isomorphic.)*<sup>8</sup>

The **Epistemic Condition** of both BvH and HK is obviously fulfilled, for Jones believes that Smith’s death is completely dependent on his poisoning. We consider the following equations to assess the causal condition:  $SD = JP \vee JS$  to capture the fact that Smith dies ( $SD$ ) if either Jones shoots ( $JS$ ) or poisons ( $JP$ ) him,  $JS = BM$  to capture that Jones shoots only when Black hands him a gun and manipulates him ( $BM$ ), and finally  $BM = \neg JP$  to capture that Black’s action depends on Jones’s failure to poison.

Regardless of whether we apply the NESS definition, the CNESS definition, or the HP definition,  $JP = 1$  comes out as a cause of  $SD = 1$ , and thus the **Causal Condition** is satisfied. (This is easy to see by observing that the structure of this example is a standard case of Early Preemption.)

BvH claim that their account can handle Frankfurt-cases like this, but that is a mistake. Recall that their variables represent the agents’ strategies rather than their actions, and that we are limited to using a single equation. The outcome function they use when discussing a Frankfurt-case is equivalent to the equation  $SD = JP \vee B$ , where  $B$  represents Black adopting his preferred strategy. Therefore on their account both Jones and Black come out as causes of Smith’s death, which is not a sensible result. BvH admit that their NESS definition is unable to handle conditional strategies like that of Black, but contend that since we are here focussing on Jones this is not a problem. Obviously simply stating that one should only focus on the sensible results of one’s theory is not a satisfactory way of defending it... (This example also highlights a more philosophical problem with their approach: it is not at all clear what it means for a strategy to be a cause. The broad consensus is that causal relata are either events/omissions or properties of events, whereas conditional strategies are neither.)

### Counterexample to the HP-definition

We here consider a second counterexample to the HP definition that was suggested in [24]. The example is of particular interest as it was presented precisely within the context of the relation between causation and moral responsibility.

**Example 8.** *We have equations  $Y = X \vee D$  and  $X = D$ , and we consider a context such that  $D = 1$ . This looks very much like a standard case of overdetermination in which  $X = 1$  and  $D = 1$  are both overdetermining causes. Yet  $X = 1$  is not an HP-cause of  $Y = 1$  (and it is a CNESS-cause). The reason for this is that  $Y = 1$  depends counterfactually on  $D = 1$  by itself, whereas it does not depend on  $X = 1$  by itself and nor does it when we take  $D = 1$  as our witness  $\vec{W} = \vec{w}$ . Rosenberg & Glymour [24] argue that this result shows the HP definition cannot offer a basis for moral responsibility, by offering the following scenario to go along with these equations:*

---

<sup>8</sup>This analysis can just as easily be applied to these more typical Frankfurt cases. Still, for those who are sceptical that proponents of Frankfurt cases are equally comfortable as I am with moving from actions to consequences, I point out that Fischer & Ravizza apply this shift in exactly the same manner as I do when discussing responsibility for consequences [12, ch. 4].