

# The Negative Effects of Neurointerventions: Confusing Constitution and Causation

Thomas Douglas and Hazem Zohny

[This is a pre-publication version. The final version has been published in *AJOB Neuroscience*.]

Birks and Buyx claim that, at least in the foreseeable future, nonconsensual neurointerventions will almost certainly suppress some valuable mental states and will thereby impose an objectionable harm to mental integrity—a harm that it is *pro tanto* wrong to impose. Of course, incarceration also interferes with valuable mental states, so might seem to be objectionable in the same way. However, Birks and Buyx block this result by maintaining that the negative mental effects of incarceration are *merely foreseen*, whereas those of neurointerventions are intended. Their thought is that harms to mental integrity are more (or perhaps only) objectionable when the suppression of valuable mental states is intended.

Birks and Buyx cash out the difference between intended and merely foreseen effects by distinguishing effects that are (partly) constituted and caused by an act. They hold that the relationship between incarcerating someone (*'Incarceration'*) and the negative mental effects of incarceration is merely *causal*, whereas the relationship between biomedically suppressing, say, a person's testosterone activity (*'Testosterone Suppression'*) and that person's 'being less likely to have a valuable sexual desire' (*'Sexual Impoverishment'*) is *constitutive*. That is to say, suppressing a person's testosterone activity is *part of what it is* for the person to be less likely to have a valuable sexual desire. In contrast, incarceration is merely a *cause* of the negative mental effects that it produces. From this, they infer that the negative mental effects of incarceration may be unintended, whereas some of the negative mental effects of neurointerventions are intended. And, as we have seen, this, in turn, is supposed to explain why *Testosterone Suppression*—and all foreseeable neurointerventions—are objectionable in one respect that *Incarceration* is not.

We dispute Birks and Buyx's characterisation of the descriptive difference between these effects. In both cases, the negative effects are caused, not constituted, by the act in question.<sup>1</sup> One way to see this is to contrast both of these cases with a case in which the relationship between an act and its effect is clearly constitutive. Promising to spend the rest of one's life with someone is plausibly a constituent in getting married; making this promise is just part of what getting married is. We discover the relationship between making this promise and marriage by looking 'inside' the act of marriage—by examining its intrinsic properties—not by examining the extrinsic events and processes that precede it.

In contrast, the relationship between *Testosterone Suppression* and *Sexual Impoverishment* is not like this. *Testosterone Suppression* is not part of what being less likely to have a valuable

---

<sup>1</sup> The relationship is also causal, we believe, in Birks and Buyx's case of shrinking, and thereby killing, a person. The shrinking causes the person's death rather than (partly) constituting it.

sexual desire is. We discover the relationship between *Testosterone Suppression* and *Sexual Impoverishment* not by examining the intrinsic properties of the latter, but by examining the extrinsic events and processes leading up to it. We discover that, among those events and processes is the suppression of testosterone levels in the blood, which, over time, produces changes to the neurochemical brain states that underpin our valuable sexual desires, and thus suppresses those desires.<sup>2</sup> Indeed Fitzpatrick (2006), whose account of the intended-foreseen distinction Birks and Buyx endorse, himself seems to acknowledge that, when one event or state of affairs produces another through a temporally extended process extrinsic to the effect, we should classify the relationship between them as causal, not constitutive.<sup>3</sup>

There is also a further problem. In order to show that *Sexual Impoverishment* is intended while the negative mental effects of incarceration are not, it is not enough to merely appeal to a descriptive difference between these two types of negative consequences. The descriptive difference must provide a *satisfying explanation* for why one consequence should be classified as intended while the other should not. Otherwise, the appeal is *ad hoc*. Yet Fitzpatrick's (2006) account of constitution which Birks and Buyx apply does not clearly provide such an explanation. As they concede in a footnote, Fitzpatrick does not provide a theory or even definition of constitution. Instead, he only provides various paradigmatic examples to illustrate what he has in mind. This lack of theoretical backing limits the explanatory value of Fitzpatrick's distinction. In explaining the intended-foreseen distinction by reference to Fitzpatrick's constitutive-causal distinction, we are seeking to explain the intended-foreseen distinction by reference to something that is itself unexplained.

This might be of little concern if the constitutive-causal distinction were itself straightforward—or at least, more straightforward than the intended-foreseen distinction. In that case, explaining the intended-foreseen distinction by invoking the constitutive-causal distinction would at least bring *some* increase in clarity. Unfortunately, the constitutive-causal distinction, at least as understood by Fitzpatrick, is not straightforward. Indeed, some of the supposedly paradigmatic examples of constitution that Fitzpatrick puts forward are contestable. For instance, Fitzpatrick holds that a bomber obliterating an area *constitutes* its occupants being destroyed. Yet it also seems plausible to say that the constitutive relation runs in the other direction—destroying the occupants of an area partly constitutes obliterating the area—and, as Nelkin and Rickless (2013) point out, it is not clear that the relation of constitution could go both ways.

The foregoing suggests that Birks and Buyx have failed to establish a difference in *intendedness* between *Sexual Impoverishment* and the negative effects of incarceration. But of course, there may still be *some* morally significant difference between them. Perhaps, for example, there is a *closer* or *qualitatively different* causal relationship between *Testosterone Suppression* and *Sexual Impoverishment* than there is between *Incarceration* and the negative mental effects of *Incarceration*.

For instance, it may be that what is of moral significance here is the proportion of the causal processes required to produce an effect that is entailed by the act that causes it. Once we have successfully suppressed a person's testosterone, a high proportion of the causal processes required to produce *Sexual Impoverishment* have been completed; there's not much causation

---

<sup>2</sup> Note that there plausibly is some set of neurochemical states *S* such that producing those states is partly constitutive of *Sexual Impoverishment*. These are the neurochemical states on which valuable sexual desires supervene. But those who administer testosterone do not intend to eliminate those states. They intend only to eliminate the neurochemical states on which *disvaluable* sexual desires supervene.

<sup>3</sup> See Fitzpatrick [cited in Birks and Buyx] at pp. 604-6.

still left to occur. By contrast, it might seem that successfully incarcerating someone is only a small step on the way to producing the negative mental effects of incarceration. Incarcerating someone entails completion of only a small proportion of the causal processes required to bring about those negative mental effects—there’s a lot more causation that must occur.

It seems to us, however, that, if proportion of causation is what matters for the moral significance of a negative outcome, at least some of the negative consequences of incarceration will be similar in their moral significance to the negative consequences of *Testosterone Suppression*. Once one has successfully incarcerated someone, at least some of the negative mental effects of incarceration require only a small further causal leap. One negative mental effect of incarceration, for example, is sensory deprivation—the impoverishment of the sensory experiences open to the individual. It seems to us that, once one has successfully incarcerated someone, not much more needs to occur for such sensory impoverishment to result.

At this point, we could, of course, continue the search for a morally significant difference in the causal relationship between *Testosterone Suppression* and *Sexual Impoverishment*, on the one hand, and *Incarceration* and its negative mental effects, on the other. But what is the motivation for continuing this search? If the motivation is simply to account for the widespread intuition that there is something especially problematic about neurointerventions, we suggest that it may be more promising to look elsewhere. We speculate that intuitions against neurointerventions may be driven in large part by the assumption that these interventions will (probably) have *quantitatively greater* negative effects than environmental interventions like incarceration. Of course, if this is so, then the intuitions are open to empirical falsification, and the appropriate way to proceed is to consider whether they are empirically warranted.

## References

- Fitzpatrick, William J. 2006. The Intend/Foresee Distinction and the Problem of ‘Closeness’. *Philosophical Studies* 128, 3: 585-617.
- Nelkin, Dana K. and Samuel C. Rickless. 2015. So Close, Yet So Far: Why Solutions to the Closeness Problem for the Doctrine of Double Effect Fall Short. *NOUS* 49, 2: 376-409.