Stephen Morse and I agree on most things. We are both physicalists, and as such believe there is only one sort of stuff in the world, physical stuff, and humans are made of this stuff (and only this stuff). We are both compatibilists, which means that even though physical stuff operates according to the laws of physics, and thus humans do too, our actions are “free enough” (because they come from psychological states that are ours) for us to be responsible for them. We both understand that responsibility assessments depend upon folk psychological concepts, and thus any evidence relevant to criminal responsibility must be couched in folk psychological terms. We both believe folk psychology is true such that the law’s assessment of responsibility is generally accurate, even given new scientific means of understanding psychological states. Finally, we both believe mental causation is necessary for criminal responsibility, which means a person’s mental states must cause an action resulting in criminal harm for a defendant to be criminally responsible.

However, we part ways regarding the reducibility of the mental. Stephen has adopted the position of non-reductivism, meaning he believes conscious mental states cannot be reduced to their underlying physical states. At times, Stephen’s position on the irreducibility of the mental appears to ground both his distinction between mind and mechanism and the threat Stephen says neuroscience could pose to law. For example, Stephen has said scholars who “assume the validity of a complete reduction of mental states to brain states at the level of (apparently) neural networks,” such as Greene and Cohen, support the view that humans are just victims of neuronal circumstances or mechanisms, and thus not conscious agents that can be held responsible (Morse 2011: 122).

The version of non-reductivism Stephen has adopted is John Searle’s “biological naturalism.” The two major points I will make in this paper are (1) a close analysis of Searle’s non-reductive theory reveals that in order to save mental causation, and yet claim conscious properties don’t reduce, Searle must endorse the reduction of intentional mental content (but not conscious properties) to underlying brain states, and (2) this causal reduction of mental content means such content is causal solely due to its underlying physical/functional role and not via its conscious properties. Thus according to Searle’s theory, conscious properties are epiphenomenal, meaning decision-making is purely mechanistic and would operate in exactly the same way without conscious properties.

Stephen might avoid this result by claiming that Searle’s theory doesn’t demand that intentional mental content reduce even though its causal properties do. However, as J. Kim has noted, on this reading of Searle’s theory mental content is either epiphenomenal, or mental causation is over-determined. That is, either the mental is not causal via its content but instead only via its underlying physical state, or the
mental is causal in addition to, and separate from, its underlying state, resulting in over-determination. This appears to be a general problem with taking a non-reductive position regarding mental causation: whatever aspect of the mental that fails to reduce either becomes epiphenomenal, or participates in causal over-determination.

By highlighting the weaknesses of the non-reductive position, I hope to convince Stephen that the best, and possibly only, way to preserve conscious agency is to allow causal reduction to reach all the way up to conscious properties. A plausible reductive account needn’t endorse an ontological reduction of consciousness, as it may only be possible to access first-person properties if one is the holder of a conscious mental state. However, a “multifield and multilevel” (Morse 2011: 122) reductive account would claim causally effective conscious states could be understood in neuroscientific terms. According to this view, irreducible conscious states are not the hallmark of legal agency, but instead, legal agency hinges on the physical causal processes that give rise to consciousness.