

Daniel C. Dennett
University Professor at Tufts University
Austin B. Fletcher Professor of Philosophy
External Professor, Santa Fe Institute

January 2015

Brief statement for Rome

What Neuroscience will tell us about Moral Responsibility

There has been a lot of speculation recently about how advances in the neurosciences are going to oblige society in general, and lawmakers in particular, to reform or even overthrow our current understanding of the law, citizenship and, particularly, punishment. The theme that unites these speculations is the suggestion, sometimes explicitly endorsed, that science has shown that we human beings do not have free will after all, and hence are not morally responsible agents. Punishment is therefore unjustifiable, and should be replaced by a non-punitive system of treatment, with restraint only to the extent that it protects the public from dangerous individuals. More moderate proposals urge that we reform our policies to minimize punishment, restricting it to circumstances where we have well-grounded expectations of deterrent effect, to uphold respect for the law—since no miscreant is ever *really* morally responsible.

The demand for dramatic reforms in our inhumane systems of punishment (especially in the United States) is welcome, but the reasoning behind this particular informal campaign is dubious indeed. It depends on the assumption that the kind of “free will” that is prerequisite for moral responsibility is incompatible with determinism. Science has shown that all human actions, however deliberated, are the outcomes of causal chains extending back ultimately before our birth. Some thinkers deny this well-attested empirical claim, but with more hope than evidence. The hope is motivated by the belief that *if* our choices are thus caused, they cannot be “free”—and this would be a calamity.

It seems obvious to many that we must be capable of this kind of choosing for us to be morally competent agents, but this has never been demonstrated, and has been strenuously denied by *compatibilists*, who argue that such indeterminism is not at all a prerequisite for moral responsibility. The point of contention can be focused on the claim that when a person makes a morally responsible choice, it must be the case that she “could have done otherwise”—and this is never the case in a deterministic world. But this ignores an alternative, and much more plausible, interpretation of the key phrase, which we can bring out by looking at a usefully

simple parallel in sports: who—if anybody—deserves to receive a red card in a football (soccer, to us Americans) match?

When a red card is issued, there is often heated discussion about whether it was *deserved*, and the distinction between deserved and undeserved penalties, while contentious in close calls, is obvious to all. One will seek far and wide for a football player or fan who thinks that the whole practice of issuing yellow cards and red cards and calling fouls should be abandoned, because it is too “punitive”; because it deals with human beings who could never really deserve anything-- because of the truth of determinism. It is quite clear—so clear that even young children accept it with minimal explanation or justification—that strict rules don’t just improve a game; they make it possible. If you want to play football, you have to play by the rules, and there are penalties—punishments—for violating the rules. This is fair. Life itself, as a whole, is not fair; some people are stronger, faster, more beautiful, richer, happier, more talented than others. Some are just luckier. But rules can be designed to “even the playing field” for all, and the measure of good rules is not that they never result in punishment, but that they strike a mutually acceptable balance between dangerous anarchy and over-enforcement. And one of the chief questions raised about any particular candidate for a foul is *could the player have done otherwise?* Players are held accountable for anticipating their trajectories and those of their opponents. They cannot plead “I could not have done otherwise because at the last moment I was already airborne on a collision course” if they should have foreseen this as the most likely outcome of a lunge. This is the sense of “could have done otherwise” that matters for fair rules and fair punishment, and it has nothing at all to do with whether or not determinism reigns in the physical world, or in the brains of individual people. (In fact, if causation were capricious on the football pitch so that players could not, in general, predict the outcomes of their actions, the “could have done otherwise” provision would have no application. Responsibility depends on predictability.) This is the sense of “could have done otherwise” that imposes an obligation on all participants (players of the game, or citizens of the state) to think ahead and give due consideration to likely outcomes. Nothing in neuroscience has shown that this capacity for responsible self-control is lacking in normal people.

There are those who are demonstrably not normal in this regard, and we already deem them as having diminished moral or legal responsibility, or none at all. They may have to be institutionalized against their will if they are dangerous, and they are not granted the right to sign contracts, or make legally binding promises. They, through no fault of their own, lack the requisite competence for being allowed to pass freely in the world. It is important to recognize that neuroscience does not in any way demonstrate that the difference between these unfortunate people and the rest of us is illusory.

What neuroscience has shown, and will continue to show in the coming years, is that some people whom we had thought to be normal in this regard are in fact subtly impaired in morally significant ways, and we will have to adjust our legal

systems (through legislation or legal precedent) to take account of this new knowledge, but we can be confident in advance that this will be a self-limiting process—for a quite obvious political reason: people *want* to be held responsible because it is their ticket to *social* freedom, the right to act and move as they choose, making promises, and controlling their projects. We can concentrate the forces and considerations that are at play into a simple thought experiment.

Suppose you were to learn, from well-grounded neuroscientific examination, that you are at risk of developing an impairment of judgment or self-control that will destroy your moral competence. You now have two choices:

submit to treatment that will (probably) protect you from this incapacitation, leaving you free to act in the world *at risk of being justly punished for any misdeed you commit*, or

let nature take its course, in which case you can expect to commit some destructive act sooner or later that will lead to your institutionalization.

If the treatment is easy—taking a single pill, let's imagine—the choice is also easy. If the treatment is drastic, the choice is more difficult, and in many instances, it may well prove that neuroscience can offer a terrible diagnosis with no cure in sight. Life is not fair. We are already being faced by these decisions. It has been shown that young children who fail a simple test of self-control (the famous “marshmallow test”) are much more likely to get in trouble with the law in adulthood than children who exhibit early self-control. Fortunately, there are non-invasive routines of education and practice that can repair this deficit, just as eyeglasses can restore normal vision. You wouldn't deny these routines to your own children, would you?

Who has the responsibility and the right to make such decisions? *These* are the questions we will have to address as neuroscience advances our ability to anticipate and explain deficits in human cognition and self-control, and notice that they presuppose that we—we fortunate ones—are morally responsible, and can be held accountable for our decisions.