

---

Responsibility, History and Manipulation

Author(s): John Martin Fischer

Source: *The Journal of Ethics*, Vol. 4, No. 4, Free Will and Moral Responsibility: Three Recent Views (Dec., 2000), pp. 385-391

Published by: [Springer](#)

Stable URL: <http://www.jstor.org/stable/25115664>

Accessed: 24/02/2011 13:14

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=springer>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).



Springer is collaborating with JSTOR to digitize, preserve and extend access to *The Journal of Ethics*.

## RESPONSIBILITY, HISTORY AND MANIPULATION

I'd like to thank Michael [McKenna] for all the hard work he did in putting this conference together. And thanks to the department here in supporting him. I think it was a great idea and I have had a lot of fun. I have learned a lot.

Before I start my presentation, let me say a few things about what Professor Kane said in his session earlier. There was no time to respond at that point. We were talking about the issue of whether causal determinism in the actual sequence rules out moral responsibility directly and not in virtue of alternative possibilities. One thing he said is that encoded in some of our common sense practices is the demand for the lack of determination. He used as an example a hypothetical case, though obviously it is very similar to actual cases. The case involved a young man who brutally raped a young woman. Something like that. People are very angry. But then they attend the trial and they find out that the man was abused as a young child, sexually and otherwise. It then becomes unclear what the appropriate response is. What Professor Kane concludes from such a situation is that feeling confident in holding this individual responsible requires that we find some sort of lack of determination, as he put it.

First of all it is interesting that Gary Watson writes about just such a case of Robert Alton Harris in California. Harris brutally murdered a couple of young boys. Later he laughed about it and ate their sandwiches after he murdered them. But then there were long articles in the *L.A. Times* talking about his horrible upbringing. Gary Watson in his interesting piece talks about what an appropriate reply would be. But for the purposes of our discussion today, what I would say is that the example does not show that encoded within our common practices is the demand for the lack of determination. It shows that we demand the lack of a *certain sort of determination*. What was problematic was certain kinds of determination, such as sexual abuse. But the example doesn't show that determination *per se* rules out responsibility according to common sense.

You can see that by contrasting that sort of case with the example of Leopold and Loeb. In this sort of case someone very affluent and wealthy goes to a judge and says "Well no, I was not sexually abused. I was not physically abused. I killed because I wanted to commit the perfect crime.



But by the way, I am not morally responsible because causal determinism is true.” He would be laughed out of court. No one would take that seriously . . . well maybe a couple of you in this room would. As Bob Kane points out, that is exactly how Darrow did argue in the case of Leopold and Loeb. But remember, he did not win that case (my own view is that if Leopold and Loeb had had Johnny Cochran as their lawyer maybe Darrow’s strategy would have worked!). In summary, I think that we do demand that causation not be of certain sorts. But it is not clear that we demand the lack of causation.

What I want to talk about today is my claim that moral responsibility is essentially an historical notion. I want to say a little bit about that, and about my account of a specific way in which it is an historical notion and how that might be promising in dealing with certain cases involving direct manipulation.

Certain phenomena or notions are historical, and, in contrast, certain phenomena are current time-slice notions. A current time-slice notion is something that does not depend in any crucial way on its history being a certain way rather than another. A current time-slice notion supervenes on snapshot properties – properties that one could in principle take a snapshot of at a given time. Size, weight, height, something’s being shiny or metallic, or smooth, round, bright, colorful, symmetric: these are all examples of current time-slice properties. For example, something’s being symmetric is a matter of its snapshot properties. It does not depend on its past being a certain way. However, I think that there are other notions that are essentially historical. To begin, I am going to use two examples from Al Mele. Something’s being a sunburn is an historical notion. You can look at someone’s face and see that it is red and burned, etc. But it can only be a sunburn if it was caused by the sun. Something being genuine currency rather than counterfeit is an historical notion. It is not just a matter of how it looks at a certain point in time but rather how it came into being. Also, something’s being a genuine Picasso rather than a fake is an historical notion. Something that is molecule for molecule isomorphic, type-identical to a genuine Picasso is not necessarily a genuine Picasso.

I think that the notion of distributive justice is an historical notion. Many of you know that Robert Nozick developed a particular theory of distributive justice that he argued was an historical theory. He criticized Rawls for adopting what Nozick claimed was an ahistorical or current time-slice model of justice. That was an unfair criticism of Rawls, because Rawls was *not* saying that you could look at a certain distribution of goods at a certain time, take a snapshot, and decide whether it is just. But rather Rawls was talking about the basic social institutions and contending that

they should be set up with certain constraints (then one lets them operate over time). So I think both the theories of Nozick and Rawls are historical in that they need to know how a distribution came about as well as looking at a distribution at a time. One can see this in a clear way in that one can look at a particular distribution at a given time and one does not know, for instance, whether it came about by voluntary exchange or someone stealing from someone else. And that, presumably is crucial to knowing whether the distribution is just.

Robert Nozick indicated, actually, in his book *Anarchy, State, and Utopia* in a very suggestive passage, that love is also an historical notion. There has been an interesting discussion of that claim. You can distinguish a couple different elements of the claim that love is historical. One is the idea that love is non-fungible. What I mean by that is illustrated by the ghoulish scenarios in which you are asked to imagine that you are coming home and as you arrive your wife and your children are hit by lightning bolts – suddenly vaporized. Yet, just at that same moment, due to a cosmic accident, molecule for molecule duplicates of your wife and your children are created. They come into being quite independent of the event that destroyed your wife and your children. Obviously a very wildly implausible scenario, but seemingly conceivable. Now we can ask ourselves, assuming that you do love your wife and your children, would it be appropriate for you to continue to have these attitudes towards these new individuals? Some argue that, no, love is *historical* in the sense that it would be inappropriate for you to transfer that set of attitudes constitutive of love to these new individuals because they are different individuals and love relates to particular individuals with whom you have had interactions in the past.

There is a separate idea involved in the claim that love is historical. This is the claim that it is conceptually impossible, or necessarily false that one can have “love at first sight.” Carl (Ginet) was telling us that when he first met his wife Sally, it was love at first sight. But according to this view, you can’t have love at first sight. It is like the claim that you can’t have virtue pills that induce virtue suddenly, that you need a certain process. Well, I don’t know what I think exactly about these claims about love. But that gives you a kind of flavor of some of the notions that people have claimed are historical. You could ask the same kinds of questions about reference, the reference of terms, and claims of knowledge, and justification of beliefs. For instance, certain epistemologists argue that belief can be justified only if the belief in question was formed as the result of the right sort of process. It is not just a matter of looking at a time-slice and considering the relationship between the evidence and the belief, but it is

a matter of how the belief came into being. That is of course controversial. Others (“internalists”) would argue against these sorts of “externalist” views.

So there are current time-slice notions and historical notions. My claim is that moral responsibility is an historical notion. Harry Frankfurt has contended vigorously on a number of occasions (although I am not sure he has *argued*) that moral responsibility is a current time-slice notion – that it doesn’t matter what the history is; rather it is just a matter of how the different elements of our psychic harmony are related. On his view there are first-order desires and second-order desires. The first-order desire on which we act is called the will. For Frankfurt this is a stipulative definition. Then the second-order desire about which first-order desire should be our will is called one’s second order volition. If there is a mesh between one’s second-order volition and one’s will, that is supposed to be sufficient (for Frankfurt) for moral responsibility. Or at least in some of his earlier writings it looked like that was what he was saying. So then it would just be a matter of looking at psychic elements and their arrangements and not caring how those elements and that arrangement came into being. Basically, that is Frankfurt’s view: assessing moral responsibility is a matter of looking at the elements that exist at a certain time but not looking at the history behind them.

I actually believe that this view is false. I am not sure that I can argue for the falsity of it, but I have contended vigorously that it is false! I would try to motivate the historicity claim in two ways. One, it seems to me that there are cases in which one freely puts oneself in a position in which one is out of control and yet one is responsible for those later acts. The existence of those scenarios makes it possible to consider two kinds of cases. The one case would be like what happened to Cary Grant in *North by Northwest*, the Hitchcock film. Someone forces the agent to drink, pouring liquor down his throat. Later he is drunk and out of control. He gets in a car and runs over an innocent person because he is not in control. That is very different from a case in which someone who is not an alcoholic freely gets very drunk and gets in a car and does the same thing. My point is that if you look at the drunk drivers in the two cases, there may be no differences between them in terms of their current time-slice characteristics (at the time of the drunk driving). Their central nervous systems could be molecule for molecule isomorphic resulting in the same level of drunkenness and consequent lack of responsiveness. But one is morally responsible for what he did and the other isn’t. The difference is a matter of their histories, of how they got to be the way they were. So that

is the first kind of motivation for the claim that responsibility is essentially historical.

One can also motivate the idea by thinking about cases of hypnosis, direct electronic stimulation of the brain, and brainwashing. Again, one could look at two individuals at a certain time. Assume again that the two people have the same arrangement at that time of psychic elements, the same first-order desires and second-order desires, or in a different terminology, the same preferences and values. Yet one has got those in the normal way of human moral development and the other has, unbeknownst to him, been directly stimulated, or he has had subliminal advertising applied to him and he has never consented, or hypnosis to which he did not consent. To me, the idea is that if you look at these two individuals at a certain time, there may be no difference, and yet the moral responsibility characteristics of the two individuals differ based on their histories. Another way of putting the point is that it is possible to have two individuals who, in terms of the current time-slice properties, are exactly the same, type identical, and yet one is responsible and the other isn't. So responsibility doesn't supervene on the current time-slice properties. It is a matter of history.

More specifically, what kind of history do we need? In my view, in order to be responsible we have to *take responsibility*. So it is a kind of subjective approach to moral responsibility. That is just one element. I think that one also has to act from appropriately reasons-responsive mechanisms. But those mechanism have to be one's own in some sense. The way that you make those mechanisms your own is by taking responsibility for them. Let me just sketch in a rough way what I mean by "taking responsibility." I don't mean that you ever say, "I take responsibility for my actions." or even explicitly think about taking responsibility. It is kind of a stipulative notion of taking responsibility. What it means is coming to have a certain set of beliefs about oneself. First of all, one must believe that one's choices and bodily movements are actually efficacious, that certain upshots in the world in fact result from one's own deliberations, choices and bodily movements. One can think, of course, in terms of the ordinary course of human moral development. Consider first, how a child, fairly early on, realizes that when he chooses to punch his sister and he moves his arm in such a way that his sister is hit, she cries as a result of his choices and bodily movements. So that seems to be a minimal condition. Then what I want to say is that one has to see oneself as an apt or fair target of the reactive attitudes, at least in certain contexts, on the basis of one's bodily movements and choices. You have to see yourself as a fair target in the social game of responsibility, which involves certain attitudes such as indignation, resentment, and grati-

tude (among others). And, finally, these beliefs about yourself can't just appear out of nowhere or by accident, or on the basis of no evidence. They have to be based on one's evidence in the appropriate way.

In the typical instance, that third condition is supposed to capture what is plausible about normal human moral development. As we grow up we learn that when we punch our sister, our parents have certain reactive attitudes on the basis of what we have done. We learn over time that on the basis of moving our bodies and creating certain upshots in the world, in certain contexts we should expect, or it is appropriate to expect, the reactive attitudes. So the condition is supposed to capture what (to some extent) is going on in human moral development. I want to add to these conditions the claim that we take responsibility in the first instance for *kinds* of mechanisms or *kinds* of processes that issue in our behavior. There is, for instance, the ordinary human mechanism of practical reasoning. There is also action from unreflective habit. Consider a case in which we are punished, and in which we first deliberated about the pertinent action. Suppose we say, "Boy, my sister's birthday presents look really intriguing!" Suppose we go ahead and we say, "She won't really mind if I open one." Then we open it. Well, that is a deliberative context. We learn that if we deliberate and we make those kinds of choices, it is reasonable to expect that there will be certain reactive attitudes taken by or held by others. But also a crucial milestone comes when we learn that we can also be held accountable for our actions that come from non-reflective habit.

So we take responsibility at a certain point for the kinds of mechanism that issue in our actions. And I think that this is part of what it is to be responsible. To be explicit: To be responsible is to act from a mechanism that is one's own appropriately reasons-sensitive mechanism. Part of what is involved in being one's own mechanism is that one takes responsibility for it, and one takes responsibility for it by having the beliefs described above.

I am just sketching things which I try to develop at greater length and in more detail in my book with Mark Ravizza. There are many details that I am just skipping over. But I want to end with this: I think that manipulation cases are compatibilism's dirty little secret. Compatibilists don't like to admit that this is a problem. It is to Bob Kane's and other incompatibilists' credit that they have pushed us to confront cases of covert non-constraining control. There can be thorough-going global kinds of manipulation. We compatibilists have to deal with this. In my view, honestly, Harry Frankfurt really has not addressed that problem. He has discussed it in different ways and in different places and it doesn't add up to anything – in my view. But I would like to suggest that at least the approach

that Mark Ravizza and I have begun to develop has some promise of usefully illuminating the manipulation cases. We would say that when your brain is being directly manipulated in one of the covert non-constraining control cases, one's behavior does not issue from *one's own mechanism*. One has presumably in the typical case taken responsibility for practical reasoning and for the kind of mechanism which is ordinary human practical reasoning, and perhaps also for unreflective action from habit. But one has not (thereby) also taken responsibility for action as a result of a scientist manipulating one's brain – a different kind of mechanism. That is not my own mechanism because I have not taken responsibility for it.

Now let me admit that there are (approximately) a million problems with this theory. One has to say something about mechanism-individuation. One has to talk about the various different kinds of manipulation cases and sort them out. But at least what I hoped to present is a way of honestly trying to come to grips with the manipulation cases. Ultimately it may fail. But at least I think it has some promise. And it is natural at some level to say that moral responsibility involves acting on one's *own* mechanism that is suitably reasons-responsive. When someone is directly manipulating my brain in a covert way, there is some natural sense in which it is not *my* own mechanism.