PHILOSOPHY AND THE FOLK:
ON SOME IMPLICATIONS OF EXPERIMENTAL WORK FOR PHILOSOPHICAL DEBATES ON FREE WILL

(Forthcoming in *The Journal of Cognition and Culture*, Vol. 6, No. 1-2, 2006.)

Manuel Vargas
University of San Francisco

1. Philosophy and the Folk

Consider two relatively innocuous data points.

First, unlike a good many philosophical puzzles that absorb the efforts of professional philosophers, the web of problems surrounding free will does not take philosophical training to appreciate. It is a ubiquitously accessible problem discussed at length by novelists, poets, musicians, scientists, religious believers, atheists, and more than a few undergraduates in late-night discussions. At least in the Western philosophical tradition it is also a very old problem: versions of it can be found at least as far back as the Stoics and the Epicureans, and arguably in Aristotle. Taken as a whole, these considerations suggest that at least a significant source of puzzles about free will can be found in aspects of our thinking that are available to us at easily accessible levels of reflection.

Second, over the past 30 years or so, the philosophical arsenal of incompatibilists—those who think that free will and/or moral responsibility is incompatible with determinism, etc.— has grown considerably. So have the resources of compatibilists, those who hold that free will (and/or moral responsibility) is compatible with determinism (or mechanism, or whatever the purported threat is taken to be). The spiraling increase in sophistication of both sides of what might appear from the outside to be a multi-millenial stand-off does not inspire confidence that philosophy alone has the resources to resolve the problem.

In light of these data points about the history of the problem(s) of free will and its persistence, I am inclined to think that perhaps the most interesting recent development in this debate has to do with the introduction of experimental methods to it, and along with it, the introduction of the opinions of non-philosophers to philosophical debates. A number of philosophers and psychologists have taken to studying what "the folk"—people not trained in the philosophical debates on these subjects—have to say about free will and the conditions under which they are willing to ascribe freedom and responsibility.

The limits and importance of experimentally-informed philosophy remains a subject of contention between philosophers. The *promise* of such work, though, is relatively clear: minimally, experimental methods might provide us with principled and measurable ways to assess some aspects of thought that are rooted in comparatively easily accessible levels of reflection. One interesting result of recent work in this vein, especially that of Shaun Nichols and his sometimes collaborator Joshua Knobe, suggests that, contra compatibilists, incompatibilism really is rooted in aspects of ordinary thinking, but contra incompatibilists and some of the folk, compatibilism too is rooted in common sense.[1]

In what follows, I consider the philosophical implications of this research in two parts. First, I will raise some minor questions about the methodology. Second, and primarily, I focus on what the truth of this data and its interpretation would mean for philosophical debates about free will.

2. The experimental work

---

[1] See Shaun Nichols, "Folk Intuitions on Free Will," *Journal of Cognition and Culture* 6, no. 1 & 2 (Forthcoming in 2006). See also Shaun Nichols and Joshua Knobe, "Moral Responsibility and Determinism: Empirical Investigations of Folk Concepts," (unpublished manuscript).

Nichols maintains that there are two clear results in the experimental work he discusses: (1) we have incompatibilist intuitions under some conditions and (2) we have compatibilist intuitions under others.

In one experiment, subjects are given descriptions of two different universes, one in which everything is completely caused by whatever happened before it, and the other a universe in which almost everything is determined by whatever happened before it, *except* human decision-making. Close to 95% of the respondents describe the second universe (the one where human decision-making is not completely determined by prior events) as the one most like ours. This result seems to strongly favor the long-standing claim by incompatibilists that their characterization of human agency as incompatibilist is natural, intuitive, and widespread. It is difficult to see what could be driving those responses if there were not incompatibilist (specifically, libertarian) self-conceptions at work on *some* level. How deep this aspect of our self-conception runs is open to philosophical dispute (see section 4, below).

Compatibilist intuitions are elicited in different contexts, however, especially when it comes to making predictions about human behavior. When asked to predict what other agents do, the evidence suggests that we tend to straightaway assume the truth of determinism, and to go on and make responsibility ascriptions anyway. For example, in one experiment Nichols discusses, subjects are asked to make predictions about what exact psychological duplicates will do. In particular, subjects are asked if a duplicate "will definitely decide the same thing" as the other duplicate. The majority of respondents gave determinist answers and signaled a willingness to ascribe responsibility.[2]

---

[2] As Nichols notes, this is consistent with other research on responsibility attributions. See, for example, Eddy Nahmias et al., "Is Incompatibilism Intuitive?" *Philosophy and Phenomenological Research* (Forthcoming).

The upshot seems to be that folk concepts of freedom, responsibility, and agency are fragmented or disunified in interesting ways. In cases where the descriptions are concrete and affect-provoking, we tend to opt for compatibilist conditions. In cases that are described more abstractly and that appear to involve less affect, we favor incompatibilist (and generally libertarian) analyses. This result—incompatibilist intuitions are elicited in abstract, low-affect contexts, and that compatibilist intuitions are elicited in concrete, high-affect contexts—is an important one, and I will discuss it in some detail a bit later in this paper.

3. Methodological concerns

Regarding the experiment that yields incompatibilist results, some philosophers will complain that "complete" causation, as described in the vignette given to the subjects is, on at least some interpretations, compatible with indeterministic causation. This possibility is problematic, given that the experiment is designed to test necessary or deterministic causation. I doubt this is much of an issue, however — it strikes me as implausible that many people are reading the first scenario as indeterministic.

A somewhat more serious concern is that the usage of the phrase "will definitely" might introduce some noise to the data. In colloquial American English, "definitely" seems to be consistent with thinking that something has an extremely high probability, without that probability necessarily being 1. I might think that the Kings will definitely beat the Lakers without thinking that the Kings literally can't lose. (Note that even the phrase "can't lose" is another one of those phrases susceptible to a gap between literal and colloquial uses.) If subjects are reading "definitely" in the "highly probable but not absolutely certain" sense, then the experiment may yield over-reporting of apparently determinist (i.e., compatibilist) replies.

These are minor worries that, I suspect, could be addressed without fundamentally different results. What might elicit more varied results is testing the account across various populations. For example, do differences in religious affiliation (especially when indexed to more careful measures than self-reporting) make a difference in when people are willing to attribute responsibility? Do cultural or regional differences make a difference? Does socio-economic status or political affiliation? Do gender, race, and ethnic affiliation make a difference? How robust are these results, cross-culturally? In short, how large is the "we" whose judgments are being reported in these studies?[3]

The effect of religious commitments seems especially salient given the role that various theological views have had in generating some versions of the problem of free will, and the cross-cultural stability of these judgments might cast some initial light on the extent to which the "problem of free will" is a local product of highly contingent psychological mechanisms, a cultural accident, local moral norms, or instead a deep feature of agency, its phenomenology, or the structure of widespread moral thinking. Of course, considerably more research would need to be performed before we could provide more than tentative answers, and I doubt Nichols and other empirically-oriented philosophers would disagree.

4. Philosophical Implications

To my mind, the most interesting aspect of these experimental results concerns not the results themselves but their potential implications for philosophical debates.

---

[3] There have been some surprising results with purportedly stable philosophical intuitions when they are tested cross-culturally. See E. Machery et al., "Semantics, Cross-Cultural Style," *Cognition* 92 (2004). For what it is worth—and my sense is that it is worth very little—I would expect to find that with lower SES groups you will get higher compatibilist responses in high affect contexts, some regional variation in these contexts as well (e.g., "Culture of Honor" effects), and minor but real effects depending on the intensity of religious commitments (e.g., frequently practicing non-Calvinist Christians may more frequently give incompatibilist responses).

First, these results yield something of a victory for incompatibilists. Incompatibilists need not—indeed they generally do not—deny that there are conditions under which we operate with compatibilist notions of abilities. What incompatibilists maintain is free will and moral responsibility (in some important sense(s)) are incompatible with determinism. For compatibilism to be a meaningful position, it must hold that there is *no* important sense in which free will and moral responsibility are incompatible with determinism. However, if Nichols' data are correct, contra to what traditional compatibilists have claimed, we really do imagine ourselves to be agents with genuine, metaphysically robust alternative possibilities available to us. Moreover, we really do—at least in moments of cool, abstract consideration—tend to favor an alternative possibilities requirement on moral responsibility. On the plausible assumption that these senses are important— and minimally, they are important for our self-conception—it turns out that compatibilism is dead wrong about there being no important sense in which free will and moral responsibility are incompatible with determinism.[4]

(This evidence also puts an interesting sort of pressure on semicompatibilists, at least when the view is construed as constrained by folk concepts. Semicompatibilism is the view moral responsibility is compatible with determinism, independent of whether or not we lack a freedom that requires alternative possibilities.[5] If the empirical evidence is correct, semicompatibilism is in at least semi-trouble in light of the evidence that suggests that we strongly disfavor holding people responsible when determinism is true and the case is described abstractly and with a minimum of conditions that trigger affect. Showing that we do ascribe responsibility in some cases where an agent lacks alternative possibilities would not change the fundamental trouble raised by the empirical data, for the problem would remain that there are conditions under which our responsibility ascriptions really do require alternative possibilities, even if there are conditions where they do not.)

---

[4] This evidence might be taken to give compelling support to Saul Smilansky's view that there is a "fundamental dualism" in ordinary thinking, with both compatibilist and incompatibilist strands. See Saul Smilansky, *Free Will and Illusion* (New York: Clarendon, 2000). Something like this view can also be found in Ted Honderich, *A Theory of Determinism* (New York: Oxford, 1988). As my remarks in the text suggest, I am inclined to count these views as incompatibilist, rejecting their construal of incompatibilism as requiring that commonsense is monistic in its implications. Quibbles about labels aside, I think that if the experimental results do hold up, it does vindicate a point that both Smilansky and Honderich have been at pains to emphasize: we have distinct sources of various aspects of our thinking about responsibility.

[5] The best known statement of the view can be found in John Martin Fischer and Mark Ravizza, *Responsibility and Control: A Theory of Moral Responsibility* (New York: Cambridge University Press, 1998). The view has also been embraced in Nomy Arpaly, *Merit, Meaning, and Human Bondage* (Princeton, NJ: Princeton, Forthcoming in 2006).

Although these results seem to constitute something of a victory for incompatibilism, the resources and sophistication of compatibilism should not be underestimated. Even if we grant the truth of the evidence mustered by Nichols, several routes remain open to compatibilists. First, compatibilists concerned to defend their diagnosis of folk thinking as compatibilist could argue that what is happening is that we misunderstand our own concepts, or that even if we can be said to have libertarian self-assessments, this assessment does no real or deep work in the ascription of responsibility, or at any rate, it ought not. Here is one way the compatibilist might pursue this line of response:

> While abstract characterizations of universes or of instances of blameworthiness might elicit incompatibilist answers by test subjects, what is happening is that they are misconstruing their own commitments. What the compatibilist evidence shows is that our core conception of responsibility—what is important about the concept (or even its conditions of application, if you admit of the distinction)—is compatibilist.

This route is a logical possibility and one with some historical antecedents.[6] Nevertheless, it strikes me as unpromising in this context. Compatibilists would need the incompatibilist self-conceptions that subjects give evidence of to do no work at all in our ascriptions of responsibility. However, it looks like an incompatibilist conception of agency is precisely what is doing work in those abstract description cases where determinism was judged by a majority of subjects to render not responsible the involved agents. What the compatibilist would need, then, is to appeal to some further view about our judgments of responsibility relying on, our otherwise being bound up in, a kind of affective competence. The idea would have to be something like a view that held that our proper understanding of responsibility requires the engagement of affect; the low-affect situations that trigger incompatibilist judgments are cases where our evaluation of responsibility is bypassed or otherwise prone to error. As Nichols and Knobe have argued, however, this seems implausible because the low-affect situations yield consistent answers (incompatibilist ones), which is not what you would expect if there were no comprehension at all.[7] For there to be a *systematic* error, one consistently yielding a false judgment (by hypothesis, an incompatibilist judgment), the compatibilist would need to explain the psychological mechanisms involved in the consistent production of that error, and more importantly, *why* it is an error. As Nichols notes, an alternative explanation—that there is a performance error when high affect interferes with what is an otherwise competent (incompatibilist) responsibility judgment—is at least as plausible. So, although compatibilists might yet find some way to make sense of the idea that we can acknowledge that we do make overwhelmingly incompatibilist judgments in a range of cases without those judgments representing any core aspect of our folk concepts, this looks like an uphill battle for compatibilists.

---

[6] I have in mind "conditional analysis"-style compatibilists in the vein of G.E. Moore who acknowledged that people may sometimes say and even believe that free will requires alternative possibilities in some sense that holds fixed the past and the laws, but (they go on to argue), what people *really* mean is something benignly compatibilist.

[7] See §5 of Shaun Nichols and Joshua Knobe, "Moral Responsibility and Determinism: Empirical Investigations of Folk Concepts," *Unpublished manuscript* (2004).

I am inclined to think that the Nichols and Nichols and Knobe evidence helpfully illuminates a perpetual challenge for working out the metaphysics of moral responsibility: distinguishing our genuine "theoretical beliefs" (roughly, folk conceptual metaphysics) regarding moral responsibility from the pragmatic dimensions of holding people responsible. There are plenty of good pragmatic reasons for us to hold one another responsible, and for us to assume that a person is morally responsible, especially when they have brought about some strongly undesirable consequence. Indeed, the phenomenon of apparent over-ascription of culpability is familiar to legal scholarship, and there is no reason to doubt that something similar is happening in the context of moral responsibility.[8] We tend to assume people are responsible agents unless we have reason to think otherwise, and we need pretty good reason to *not* assign responsibility to people unfamiliar to us. The difficulty of reading our metaphysics off of our practices is that we run into the difficulty of distinguishes between the pragmatic elements of our practices and the underlying theoretical beliefs that make up the folk metaphysics concerning the practice. What I suspect the data reflects is the operation of distinct psychological mechanisms involved in the metaphysics (or if you like, the connotational semantics) of responsibility, and the pragmatics of holding people morally responsible. Incompatibilist judgments reflect those mechanisms concerned with our theoretical beliefs about moral responsibility—the folk metaphysics or connotational semantics of moral responsibility—whereas compatibilist judgments reflect the concrete, pragmatics of participating in a socially-embedded practice of assigning responsibility. More would need to be said to defend the proposal I have just floated, but something like this seems to capture what is going on in the data uncovered in recent experimental work on responsibility.

---

[8] A number of interesting studies have been done on people's willingness to sentence to death even when there is good evidence that the perpetrator lacked significant normative competence and/or the ability to understand the consequences of his or her actions. In one study, taken from a sample population of former jurors in a major metropolitan area on the East Coast, death sentencing responses for 10 year olds convicted of a heinous crime ran as high as 71% of male respondents, and 52.4% of female respondents. See Catherine A. Crosby et al., "The Juvenile Death Penalty and the Eighth Amendment: An Empirical Investigation of Societal Consensus and Proportionality," *Law and Human Behavior* 19, no. 3 (1995). Although respondents did display some sensitivity to age (and thus, presumably, cognitive sophistication) as exculpating, the willingness to execute people likely lacking capacities crucial for responsibility on most philosophers' accounts is striking. For additional evidence that sentence allocation and attitudes towards punishment don't display much sensitivity to factors like cognitive and moral competence, see Simona Ghetti and Allison Redlich, "Reactions to Youth Crime: Perceptions of Accountability and Competency," *Behavioral Sciences and the Law* 19 (2001). These authors speculate that one thing that may be driving insensitivity to competency is an interest in tracking "criminal dispositions" and the perceived benefits of identifying and responding to those perceived dispositions. If so, this might explain why people are more willing to execute than would be suggested by their own estimations of competency.

If things were as I have described, would this spell the end of compatibilism? I doubt it. Compatibilists could give up the business of characterizing folk concepts of freedom and responsibility, instead treating compatibilism as a purely prescriptive project distinct from conceptual analysis-like endeavors. Prescriptive compatibilism, shorn of the project of folk-conceptual compatibilism strikes me as exactly the right project to pursue if the empirical results hold up. It would allow the sophisticated prescriptive compatibilist to grant incompatibilism the comparatively small victory of having described an important part of our pre-philosophical intuitions that make up our commonsense picture of agency, without taking on some of the more metaphysically demanding aspects of libertarianism. The concession about how to construe our folk concepts should be easy enough to grant when we acknowledge that the more important issue is what we *ought* to think about free will and responsible agency, and not what we happen to currently think about these things. Going this route, however, requires that one be willing to embrace at least some degree of revision in our self-conception, and this sort of move raises its own questions and philosophical puzzles.[9] As Nichols notes, there are plenty of considerations that should inform a prescriptive account, and any adequate account of prescriptive compatibilism (or prescriptive incompatibilism, for that matter), would need to offer some account of the considerations that appropriately govern the reconstruction of common sense. In short, while the experimental evidence Nichols cites would not mean the end of compatibilism, it would require it to revise its dominant self-conception and to undertake burdens that are rarely if ever explicitly recognized. To the extent that the experimental evidence supports these developments in the literature, I think it is clear that there is an important philosophical upshot to this work.

---

[9] I have tried to address some of the issues in Manuel Vargas, “The Revisionist's Guide to Responsibility,” *Philosophical Studies* 125, no. 3 (2005).

The experimental work yields a second lesson important for philosophers working in this domain. What it shows is that, at minimum, we should very carefully attend to how our own reactions to various thought experiments and purported counterexamples might be a function of high-affect triggers (or low-affect triggers) of judgment.[10] Until this issue is better understood, we may be vulnerable to cynical uses of this data: if you want your thought experiment to elicit compatibilist intuitions, you would do well to frame the example in terms that trigger emotional reactions, reactions whose presence tends to strongly favor compatibilist reactions. The reverse is true for incompatibilists.

To see why this might matter, reflect on the enormous literature that has grown up around "Frankfurt-style examples." Until Frankfurt offered a purported counterexample to it, the consensus among philosophers (both compatibilist and incompatibilists) was that something like a principle of alternative possibilities was correct: roughly, that in order to count as responsible for something you had to have the ability to do otherwise.[11] The main issue between compatibilists and incompatibilists was how to understand the ability to do otherwise.

---

[10] For example, I suspect that something like high-affect framing plays a role in subjects' reaction to some especially violent examples used in a study that purports to show the intuitiveness of compatibilism in Robert L. Woolfolk, John Doris, and John Darley, "Identification, Situational Constraint, and Social Cognition: Studies in the Attribution of Moral Responsibility," *Cognition* (forthcoming).

[11] Harry Frankfurt, "Freedom of the Will and the Concept of a Person," *Journal of Philosophy* 68, no. 1 (1971).

Incompatibilists held that the power required something that would be incompatible with determinism and compatibilists sought to show that it required an ability fully compatible with determinism. What Frankfurt's example seemed to show was that an agent could be responsible without having the ability to do otherwise. Here is a stripped-down example of the sort he offered:

> Consider an agent, Ajay, who is deciding between two options, the Good Choice, and the Bad Choice. Unbeknownst to Ajay, someone nefarious has put a device in his head that will ensure he selects the Bad Choice, if it learns that he is about to make the Good Choice. However, in the actual sequence of things Ajay makes the Bad Choice on his own, so the device never activates leaving Ajay's deliberation untouched and altogether free of interference.

From this example we are to conclude that although Ajay lacked the ability to do otherwise (this is what the intervener who never has to intervene is supposed to show), he is nonetheless responsible for making the Bad Choice.

(By way of side note, it is an interesting and unstudied question whether ordinary people find these cases as compelling at anything like the rates at which professional philosophers seemed to have found these cases compelling. The *philosophical* consensus at this point seems to be that there are good reasons to be skeptical about these cases as Frankfurt first sketched them, but many remain optimistic about the viability of more recent Frankfurt-style cases.[12])

Here's a testable hypothesis: the more concrete and emotionally-laden a Frankfurt case is made to be, the higher the rate at which people will want to ascribe moral responsibility. If

---

[12] For a selection of some of the most important work in this literature, see David Widerker and Michael McKenna, eds., *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities* (Burlington, VT: Ashgate, 2003). See also John Martin Fischer, "Recent Work on Moral Responsibility," *Ethics* 110, no. 1 (1999).

Nichols and Nichols and Knobe are right that concrete, high-affect conditions trigger psychological processes of that have broadly compatibilist conditions responsibility attribution, this is what we should expect to find. Indeed, if we accept their view of different psychological mechanisms of responsibility attributions, it is natural to wonder whether concrete examples will always skew in favor of a responsibility attribution unless powerful considerations recommend against it.

In a recent paper, Monterosso, Royzman, and Schwartz have argued that experimental data shows that different kinds of explanation have significant effects on the perceived culpability of actors.[13] What they find is that physiological explanations of why an event took place tend to undermine responsibility attributions significantly more than "experiential" explanations that appeal to the agent's history or experiences. If this trend holds—and here, I am merely speculating—what we should expect to find in Frankfurt cases is that when underlying physiological features are emphasized (e.g., Ajay's brain chemistry) this will lead to exculpatory judgments more frequently than cases that cases that ignore these aspects and instead focus on what deliberation was like for the agent.

What particular lessons we should draw about Frankfurt cases in light of these considerations is not clear, even if these highly speculative conclusions were vindicated.[14] If experimental evidence did show that the power of Frankfurt cases depends on various underappreciated facts about their framing, philosophers would surely adopt a range of different responses. Some would argue that the opinions of the folk are irrelevant, and that it is philosophical judgments that matter, reflecting more sophistication, nuance, accuracy, or what have you. Others would favor a kind of intuitional egalitarianism, arguing that the judgments of philosophers are not privileged, and that our assessments of Frankfurt cases should depend on folk judgments about them. Still others would favor various views in between, or argue that these issues are all tangential, irrelevant to the issue of what we ought to believe.

---

[13] John Monterosso, Edward B. Royzman, and Barry Schwartz, "Explaining Away Responsibility: Effects of Scientific Explanation on Perceived Culpability," *Ethics and Behavior* 15, no. 2 (Forthcoming).

[14] Fruitful exploration might also rely on more than surveys and intuition-testing. For example, brain-based imaging could prove to be a useful compliment and test for various hypotheses about the mechanisms involved in assessments of responsibility. We might be able to learn whether certain ways of framing examples were more and less likely to trigger affect-associated areas of the brain, and we could then further sort examples on the basis of whether they were more or less likely to trigger responsibility judgments in low-affect contexts.

We are a very long way from settling these issues. It is clear, however, that we need to begin thinking about these issues, and asking ourselves whether the power of particular examples or arguments turns less on their global epistemic credentials and more on the psychological mechanisms the particular of an example can be set up to trigger. It goes without saying that the philosophical significance of this too —whether it matters that an example derives its power from one psychological mechanism rather than another—should matter. In sum, the introduction of empirically-informed work on agency, ethics, and moral psychology raises a host of questions about whether and why philosophers have found particular examples compelling or not, and these merit careful exploration.

Before concluding, I want to briefly mention a third and final issue on which experimental work might have implications for philosophical debates about free will: Why has there been a free will debate at all, much less one that has gone on for this long? Philosophers have offered a number of explanations, and Nichols' own suggestion is this: we have distinct psychological mechanisms or modes of analysis that favor both compatibilist or incompatibilist conceptions of our agency, which in turn fuel the ongoing debate. I am agnostic about this proposal, but it is suggestive. Suppose we accept the "two mechanisms" view of responsibility assessments. Beyond explaining the persistence of the debate, it might also suggest something about how philosophers come to hold the views they do hold in the debate. On the two mechanisms view, thoroughgoing compatibilists and incompatibilists are presumably privileging the results of one kind of psychological process over the other. Which one a given philosopher takes to be paradigmatic may be contingent on a range of extra-philosophical factors. These things might include who one's teachers were, what one's religious convictions might be, how forcefully a view was initially presented, or the order in which one was exposed to the problems of one or another view. If so, we would have the beginnings of an account not only of the persistence of the debate but the outlines of an explanation for why particular philosophers hold the views they do.

These remarks are little more than speculation and should be taken with more than a couple grains of salt. That said, this possibility does illustrate a further way in which the introduction of experimental methods to philosophical topics might have both philosophical and metaphilosophical import.[15] For some, the effects of extra-philosophical factors on one's own philosophical views is a comparatively innocuous truism about philosophy. Clearly, though, this possibility would be less graciously received by others.

5. Concluding remarks

Experimental work on free will and moral responsibility is promising and overdue. Its potential philosophical implications are significant in at least three ways. First, it may provide resources to support various forms of incompatibilism, and in turn, it may give impetus to sophisticated consequentialists responses. (Or, if further experimental results disfavor incompatibilists, it would put pressure on them to explain the importance of a kind of agency with which—in the imagined scenario—we are generally not concerned.) Second, experimental work raises important questions about the methodological assumptions that go undiscussed in a good deal of the philosophical literature. Minimally, it makes it harder for philosophers to ignore the possibility that an adequate account of free will—libertarian, compatibilist, or other—may require forsaking some elements of commonsense. Thirdly, and relatedly, the work has the interesting feature of studying the very thing that makes free will such a widely-discussed philosophical conundrum: widespread, easily elicited commonsense intuitions about freedom and responsibility. While philosophical literatures sometimes tend to move gradually further away from the problems that spawned them, the work of Nichols and other empirically-minded philosophers is fruitfully returning us to the historical breeding ground of the problem. It does not strike me as a bad thing to reacquaint philosophers with the folk. I doubt that the reintroduction of the folk into the debate in a systematic way will be decisive on many issues, but it will spur philosophical thinking in important ways. And anyway, it is perhaps salubrious to remind ourselves what intuitions look like before layers of philosophical training and investment in one or another view work their magic professional philosophers' own conceptions of common sense.

---

[15] There have been some stabs at trying to understand the way psychological and social mechanisms might drive the philosophical positions people adopt, both at a time and over time. For some examples, see Randall Collins, *Sociology of Philosophies* (Cambridge, MA: Belknap, 1998), Colin McGinn, *Problems in Philosophy: The Limits of Inquiry* (Cambridge, MA: Blackwell, 1993), Shaun Nichols, "The Rise of Compatibilism: A Case Study in Quantitative History of Philosophy," (unpublished manuscript).

Works cited

Arpaly, Nomy. *Merit, Meaning, and Human Bondage*. Princeton, NJ: Princeton, Forthcoming in 2006.

Collins, Randall. *Sociology of Philosophies*. Cambridge, MA: Belknap, 1998.

Crosby, Catherine A., Preston A. Britner, Kathleen M. Jodl, and Sharon G. Portwood. "The Juvenile Death Penalty and the Eighth Amendment: An Empirical Investigation of Societal Consensus and Proportionality." *Law and Human Behavior* 19, no. 3 (1995): 245-61.

Fischer, John Martin. "Recent Work on Moral Responsibility." *Ethics* 110, no. 1 (1999): 93-139.

Fischer, John Martin, and Mark Ravizza. *Responsibility and Control: A Theory of Moral Responsibility*. New York: Cambridge University Press, 1998.

Frankfurt, Harry. "Freedom of the Will and the Concept of a Person." *Journal of Philosophy* 68, no. 1 (1971): 5-20.

Ghetti, Simona, and Allison Redlich. "Reactions to Youth Crime: Perceptions of Accountability and Competency." *Behavioral Sciences and the Law* 19 (2001): 33-52.

Honderich, Ted. *A Theory of Determinism*. New York: Oxford, 1988.

Machery, E., R. Mallon, Shaun Nichols, and Stephen Stich. "Semantics, Cross-Cultural Style." *Cognition* 92 (2004): B1-B12.

McGinn, Colin. *Problems in Philosophy: The Limits of Inquiry*. Cambridge, MA: Blackwell, 1993.

Monterosso, John, Edward B. Royzman, and Barry Schwartz. "Explaining Away Responsibility: Effects of Scientific Explanation on Perceived Culpability." *Ethics and Behavior* 15, no. 2 (Forthcoming).

Nahmias, Eddy, Stephen Morris, Thomas Nadelhoffer, and Jason Turner. "Is Incompatibilism Intuitive?" *Philosophy and Phenomenological Research* (Forthcoming).

Nichols, Shaun. "Folk Intuitions on Free Will." *Journal of Cognition and Culture* 6, no. 1 & 2 (Forthcoming in 2006).

———. "The Rise of Compatibilism: A Case Study in Quantitative History of Philosophy." (unpublished manuscript).

Nichols, Shaun, and Joshua Knobe. "Moral Responsibility and Determinism: Empirical Investigations of Folk Concepts." *Unpublished manuscript* (2004).

———. "Moral Responsibility and Determinism: Empirical Investigations of Folk Concepts." (unpublished manuscript).

Smilansky, Saul. *Free Will and Illusion*. New York: Clarendon, 2000.

Vargas, Manuel. "The Revisionist's Guide to Responsibility." *Philosophical Studies* 125, no. 3 (2005): 399-429.

Widerker, David, and Michael McKenna, eds. *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities*. Burlington, VT: Ashgate, 2003.

Woolfolk, Robert L., John Doris, and John Darley. "Identification, Situational Constraint, and Social Cognition: Studies in the Attribution of Moral Responsibility." *Cognition* (forthcoming).