Philosophical Psychology
Vol. 18, No. 5, October 2005, pp. 561–584

Routledge
Taylor & Francis Group

Surveying Freedom: Folk Intuitions about Free Will and Moral Responsibility

Eddy Nahmias, Stephen Morris, Thomas Nadelhoffer, and Jason Turner[1]

Philosophers working in the nascent field of 'experimental philosophy' have begun using methods borrowed from psychology to collect data about folk intuitions concerning debates ranging from action theory to ethics to epistemology. In this paper we present the results of our attempts to apply this approach to the free will debate, in which philosophers on opposing sides claim that their view best accounts for and accords with folk intuitions. After discussing the motivation for such research, we describe our methodology of surveying people's prephilosophical judgments about the freedom and responsibility of agents in deterministic scenarios. In two studies, we found that a majority of participants judged that such agents act of their own free will and are morally responsible for their actions. We then discuss the philosophical implications of our results as well as various difficulties inherent in such research.

Keywords: Free Will; Moral Responsibility; Folk Psychology; Experimental Philosophy

1. Introduction: Motivating the Project of Testing Folk Intuitions

> I am sometimes asked—in a tone that suggests the question is a major objection—why, if conceptual analysis is concerned to elucidate what governs our classificatory practice, don't I advocate doing serious opinion polls on people's responses to various cases? My answer is that I do—when it is necessary. (Jackson, 1998, p. 37)

Philosophers disagree about the proper role that folk concepts, common sense, and prephilosophical intuitions should play in shaping our philosophical ideas

Correspondence to: Eddy Nahmias, Department of Philosophy, Georgia State University, Atlanta, GA 30302-4089, USA. Email: enahmias@gsu.edu

ISSN 0951-5089 (print)/ISSN 1465-394X (online)/05/050561-24 © 2005 Taylor & Francis
DOI: 10.1080/09515080500264180

and theories. Whereas some have dismissively viewed the beliefs and intuitions of non-philosophers with either suspicion or scorn, others have attempted to refute speculative philosophical theories—such as skepticism or idealism—by appealing to what they take to be our commonsense intuitions.[2] Such intuitions play a conspicuous role in the free will debate, where philosophers often motivate their position by claiming that it is commonsensical, fits with ordinary intuitions, accounts for our practices of attributing moral responsibility, and captures a conception of freedom we value. However, few philosophers have tried to ascertain what these commonsense intuitions actually are. More often than not, philosophers are content to place their *own intuitions* into the mouths of the folk in a way that supports their own position—neglecting to verify whether their intuitions agree with what the majority of non-philosophers *actually* think.

This paper represents a preliminary attempt to correct for this oversight. In this section we describe our project and the motivation for it. We then discuss two studies we performed to test folk intuitions about free will, moral responsibility, and the ability to do otherwise (section 2). We then consider some of the difficulties facing this kind of empirical research and discuss how such research might impact the philosophical debates about free will (section 3). We hope our research will spark further interest in the examination of ordinary people's intuitions about free will and related concepts and in developing methods to study such intuitions.

We suggest that the best way to arrive at a sound—and empirically accurate, rather than merely speculative—understanding of folk intuitions is to conduct surveys of non-philosophers in an effort to generate the much needed data—to do what Jackson suggests above: ''serious opinion polls on people's responses to various cases''. Our project can thus be viewed as part of a gathering storm of 'experimental philosophy'—recent attempts to empirically test the claims philosophers make about folk intuitions in epistemology, ethics, and action theory.[3] In this paper, we expand the scope of this novel, and we believe illuminating, approach to philosophy by applying it to the free will debate. Jackson (1998) suggested that this sort of research will not be necessary when philosophers' claims about ordinary intuitions are uncontroversial—when there is agreement about what people's intuitions are and philosophers ''know that our own case is typical and so can generalize from it to others'' (p. 37).[4] In the free will debate these criteria are not met. There is substantial disagreement among philosophers about what ordinary people's intuitions are, as we'll illustrate shortly. This disagreement likely stems from the philosophers' own divergent intuitions. We suspect that this conflict indicates that philosophers' intuitions have been corrupted by their theories, making it uncertain whether their own case is ''typical'' and all the more necessary to survey the intuitions of pre-theoretical folk.

While *compatibilists* believe we can have free will and be morally responsible even if causal determinism is true, *incompatibilists* maintain that the existence of free will entails the falsity of causal determinism. *Libertarians* are incompatibilists who think we have free will; *skeptics* are incompatibilists who think we don't. To say that these parties find themselves in a stalemate would be an understatement. One indication of

how intractable the debate has become—and how entrenched the respective parties are—is the fact that so many philosophers claim that their own position has the most intuitive appeal and best fits our ordinary conception of free will and our practices of responsibility attribution.

We find, on the one hand, incompatibilists who suggest that their view is commonsensical and that compatibilism is counter-intuitive. Robert Kane (1999), for example, writes, "most ordinary people start out as natural incompatibilists.... Ordinary persons have to be talked out of this natural incompatibilism by the clever arguments of philosophers" (p. 218). Laura Ekstrom agrees that "we come to the table, nearly all of us, as pretheoretic incompatibilists" (2002, p. 310) such that it is the compatibilist who "needs a positive argument in favor of the compatibility thesis" (2000, p. 57). Galen Strawson (1986) contends that the incompatibilist's libertarian conception of free will, though impossible to satisfy, is precisely "the kind of freedom that most people ordinarily and unreflectively suppose themselves to possess" (p. 30), adding that it is "in our nature to take determinism to pose a serious problem for our notions of responsibility and freedom" (p. 89). And Thomas Pink (2004) tells us that "most of us start off by making an important assumption about freedom. Our freedom of action, we naturally tend to assume, must be incompatible with our actions being determined" (p. 12).[5] By making these sorts of claims, incompatibilists are presumably trying to situate the burden of proof on compatibilists—demanding that they explain how compatibilist theories of free will could satisfy our ordinary notions—while at the same time motivating the metaphysically demanding libertarian theories that some incompatibilists defend (e.g., agent causation) and that other incompatibilists (skeptics) attack as impossible or implausible.

On the other hand, compatibilists also appeal to commonsense intuitions, suggesting that the folk do *not* demand the libertarian conception of free will, which requires an ability to do otherwise purportedly incompatible with determinism. For instance, Daniel Dennett (1984b) claims that when ordinary people assign moral responsibility, "it simply does not matter at all... whether the agent in question could have done otherwise in the circumstances" (p. 558). Frankfurt-style cases (Frankfurt, 1969) are designed to bring to light the intuition that the freedom necessary for moral responsibility does not require the ability to do otherwise and hence is compatible with determinism.[6] Susan Wolf (1990) motivates her compatibilist "Reason View" in part by claiming that it "seems to accord with and account for the whole set of our intuitions about responsibility" (p. 89). And William Lycan (2003) rejects the incompatibilists' attempts to situate the burden of proof, arguing instead that compatibilism is "the default position... not only true, but the only position rationally available to impartial observers" (p. 107).[7]

Obviously, to the extent that incompatibilists *and* compatibilists claim that their own respective positions best accord with and account for our prephilosophical intuitions, they cannot *both* be right. Moreover, they can't really claim to know if they are right, since no one has systematically tested to see what people's prephilosophical intuitions are.[8] It is therefore important to determine the content of these intuitions.

However, finding an adequate method for testing such intuitions is no easy task. Here we discuss our own attempts to collect data on folk intuitions about free will and moral responsibility, some difficulties facing this sort of research, and suggestions for further research. We also examine the question of what role such data about folk intuitions should play in the philosophical debate.[9]

In our attempts to get at folk intuitions, we employed thought-experiment scenarios to see how ordinary people respond to different compatibilist and incompatibilist hypotheses. This data is important because any adequate philosophical analysis of free will should be, as Alfred Mele (2001) suggests, "anchored by common-sense judgments" about particular cases—after all, if a philosophical analysis of free will runs entirely afoul of what the majority of non-philosophers say, this analysis "runs the risk of having nothing more than a philosophical fiction as its subject matter" (p. 27). In this respect we also agree with Jackson (1998) that the fundamental issue in the free will debate should be "whether free action according to our ordinary conception, or something suitably close to our ordinary conception, exists and is compatible with determinism" and that to identify our ordinary conception we must "appeal to what seems to us most obvious and central about free action ... as revealed by our intuitions about possible cases" (p. 31).[10]

Consider, for instance, the aforementioned claims that the majority of non-philosophers have incompatibilist intuitions. On the surface, this is a straightforward empirical claim that entails certain predictions about how people would respond to various thought experiments. So, for example, if it were true that most non-philosophers do share the incompatibilists' intuitions, then we should expect that if they were given a thought experiment involving an (otherwise ordinary) agent in a deterministic scenario, a majority of them would *not* attribute free will and moral responsibility to the agent. Conversely, those compatibilists who claim that the folk share their intuitions should predict that a majority of people *would* attribute free will and moral responsibility to such an agent. Yet, until a concerted effort is made to probe folk intuitions and judgments via systematic psychological experiments, the truth of these sorts of empirical claims goes unchecked.

We are not, however, suggesting that discovering what folk intuitions really are would *resolve* the free will problem. There are various responses either side could make if it turned out their view did not fit with ordinary intuitions, some of which we'll discuss below. Nonetheless, if a philosophical theory *does* turn out to be privileged by the endorsement of the folk, that would seem to position the burden of proof on the shoulders of those who argue *contrary* to folk intuitions. If it turns out that a significant majority of people make judgments that support either compatibilist or incompatibilist views, that would at least give "squatters' rights" to whichever position has such support. At least, this seems to be the idea expressed by those incompatibilists and compatibilists quoted above who claim to have commonsense intuition on their side.

Data about folk intuitions would also help situate other positions in the free will debate. For instance, some philosophers argue that ordinary intuitions are mistaken and folk concepts need to be revised or even eliminated.[11] Some offer an explanation

for why people have the intuitions they do about free will but why those intuitions do not in fact commit them to certain conceptual or theoretical views.[12] And some think that we have a variety of conflicting paradigms of free and responsible action, which has led philosophers to develop conflicting theories of free will (see Double, 1991, 1996). But these interesting positions still require that we first determine, rather than merely speculate about, what the folk concepts and intuitions actually are—that is, what it is that needs to be explained or revised, or whether we in fact have conflicting views about free will. Hence, we must first make an earnest attempt to probe the intuitions in question.

In our view, the best way to identify whether people have the intuition that determinism conflicts with free will is to survey laypersons who have not studied—and hence have not been influenced by—the relevant philosophical arguments about free will. Such people should be presented with scenarios portraying determinism and then asked whether they believe an agent acts in such a scenario of his or her own free will and is morally responsible. Negative responses would indicate the intuition that determinism conflicts with free will and responsibility; positive responses would indicate that people do not have this incompatibilist intuition. By collecting such data we hope to shed light on ordinary concepts and intuitions in a way that will be useful to all of the parties in the free will debate. However, as we will discuss in section 3, surveying intuitions about free will and determinism poses some particularly difficult problems. And while we have attempted to address some of these problems, we nevertheless view our studies as preliminary and exploratory and hope they will generate interest in further research. We also hope to prompt philosophers to be more explicit about the relationship between their conceptual analyses of free will and folk concepts and intuitions—e.g., whether they mean to be describing, explaining, revising, or eliminating the folk's concept of free will.

2. Studies on Folk Intuitions

To find out what the freedom-relevant intuitions of the folk really are we performed a number of empirical studies. These studies test the incompatibilist prediction that most people will judge that agents in a deterministic scenario do not act of their own free will and are not morally responsible for their actions.[13]

2.1. Study 1: Jeremy Cases

In some initial surveys we found that people do not understand the concept 'determinism' in the technical way philosophers use it. Rather, they tend to define 'determinism' in contrast with free will. This result alone, we suggest, does not bolster the incompatibilist position. It does not suggest that people consider 'determinism', *as defined in (one of) the technical ways philosophers define it*, to be incompatible with free will or moral responsibility. Rather, it seems that many people think 'determinism' *means* the opposite of free will, as suggested by the phrase 'the problem of free will and determinism'.[14]

So, in order to test whether folk judgments support incompatibilist intuitions, we developed thought experiments describing deterministic scenarios, roughly in the philosophical sense of the concept, without begging any questions by using the term 'determinism' or describing determinism as involving constraint, fatalism, or reductionism. The first survey we ran uses a Laplacean notion of determinism.[15] Participants read the following scenario in Figure 1 (including the timeline) and answered two questions about it (then, on the back of the questionnaire, they responded to a manipulation check, were invited to explain their answer, and offered some demographic information).

> Scenario: Imagine that in the next century we discover all the laws of nature, and we build a supercomputer which can deduce from these laws of nature and from the current state of everything in the world exactly what will be happening in the world at any future time. It can look at everything about the way the world is and predict everything about how it will be with 100% accuracy. Suppose that such a supercomputer existed, and it looks at the state of the universe at a certain time on March 25, 2150 AD, 20 years before Jeremy Hall is born. The computer then deduces from this information and the laws of nature that Jeremy will definitely rob Fidelity Bank at 6:00 pm on January 26, 2195. As always, the supercomputer's prediction is correct; Jeremy robs Fidelity Bank at 6:00 pm on January 26, 2195.

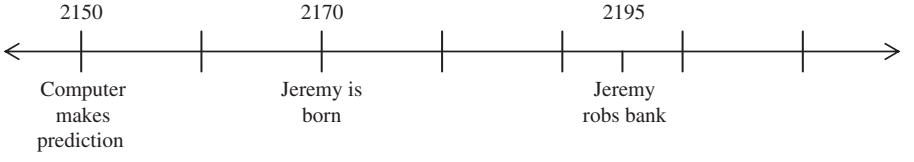

Figure 1 Jeremy Case 1: Bank Robbing Scenario.

For this study (and those discussed below), participants were undergraduates who had not studied the free will problem.[16] In pilot studies we found that some participants seemed to fail to reason conditionally (e.g., given their explanations on the back of the survey, some seemed to assume that the scenario is impossible *because* Jeremy has free will, rather than making judgments about Jeremy's freedom on the assumption that the scenario is actual). To correct for this problem, Question 1 asked participants whether they think the scenario is possible (the majority responded "no", offering various reasons on the back of the survey).[17] Then they were then asked to "suspend disbelief" for the experimental question:

> Regardless of how you answered question 1, *imagine* such a supercomputer actually did exist and actually could predict the future, including Jeremy's robbing the bank (and assume Jeremy does not know about the prediction):
>
> Do you think that, when Jeremy robs the bank, he acts of his own free will?

The results indicate that a significant majority of participants (76%) judged that Jeremy robs the bank of his own free will.[18]

We wondered whether some people were inclined to judge that Jeremy acts freely because he performs a blameworthy action. Perhaps their emotional response primed them to hold him responsible for robbing the bank, and this—combined with a tacit belief that responsible actions must be performed of one's own free will—led some participants to answer in the way they did, masking the effect of the determinism in this case (i.e., case 1, 'negative action'). To test for this possibility, we ran two more sets of surveys with identical wording except that in one case (case 2, 'positive action') Jeremy performs the praiseworthy act of saving a child from a burning building, and in another (case 3, 'neutral action') he goes jogging.[19]

However, in both these cases participants' responses closely tracked those from the blameworthy case: 68% said Jeremy saves the child of his own free will, and 79% said he goes jogging of his own free will.[20] Hence, the moral status of an action, and any emotional responses it evokes, appeared to have no significant effect on judgments of free will. The majority of participants judged that Jeremy acts of his own free will in this deterministic scenario regardless of the type of action involved (see Figure 2).

In order to test directly whether people judge that an agent is morally responsible for his actions in a deterministic scenario, we asked participants (in case 4): "Do you think that, when Jeremy robs the bank, he is morally blameworthy for it?" and (in case 5): "Do you think that, when Jeremy saves the child, he is morally

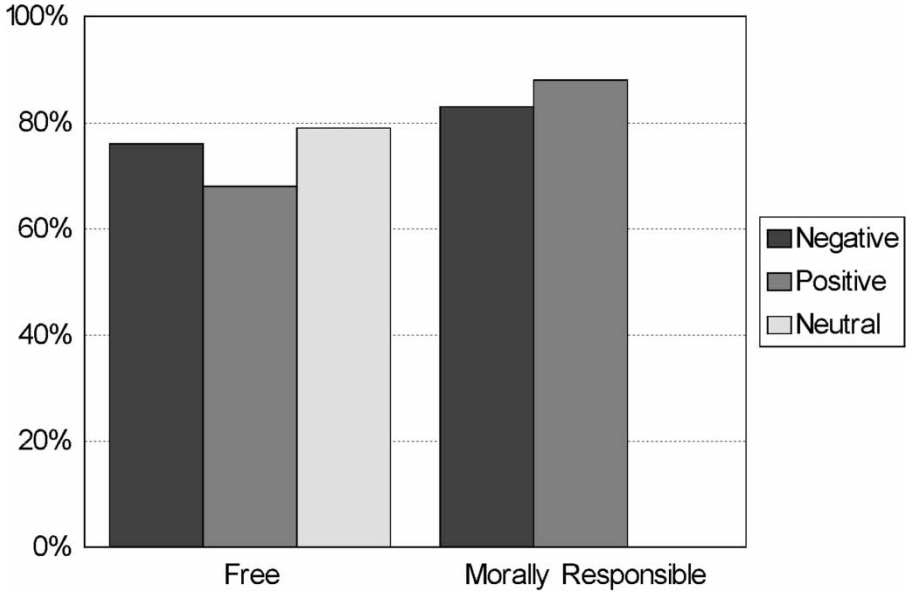

Figure 2 Judgments of Free Will and Moral Responsibility. Percentage of participants who judged Jeremy acted of his own free will when he robbed a bank (negative), saved a child (positive), and went jogging (neutral), as compared with percentage of participants who judged Jeremy morally responsible when he robbed a bank (negative) and saved a child (positive).

praiseworthy for it?'' Results from these participants closely tracked those offered by participants who had responded to questions about free will (though a slightly higher proportion of participants judged that Jeremy is morally responsible than had those who judged that he acts of his own free will): 83% judged he was blameworthy in the negative case and 88% judged he was praiseworthy in the positive case (see Figure 2).[21]

Since we surveyed one set of participants to make judgments about Jeremy's free will and another set to judge his moral responsibility, we can be confident that the consistency in the response patterns is not being driven by individuals trying to maintain consistency in their answers (e.g., first answering that he acted freely, then inferring that he must be morally responsible, or vice versa). This consistency accords well with philosophers' claims that judgments about free will are closely related to judgments about moral responsibility.

Finally, we examined whether people's judgments of an agent's *ability to choose otherwise* in a deterministic scenario would track judgments about his free will and moral responsibility, since philosophers disagree both about whether determinism conflicts with people's ordinary conception of the ability to do (or choose) otherwise and also about whether free will and moral responsibility require the ability to do otherwise. In these cases, participants were asked—again, imagining the scenario were actual—whether or not Jeremy could have chosen *not* to rob the bank (case 6), whether he could have chosen *not* to save the child (case 7), or whether he could have chosen *not* to go jogging (case 8).

In the blameworthy variation, participants' judgments of Jeremy's ability to choose otherwise (ACO) did in fact track the judgments of free will and responsibility we collected, with 67% responding that Jeremy could have chosen *not* to rob the bank.[22] However, in the praiseworthy case, judgments of ACO *were* significantly different from judgments of his free will and responsibility: Whereas a large majority of participants had judged that Jeremy *is* free and responsible for saving the child, a majority (62%) answered ''no'' to the question: ''Do you think he could have chosen *not* to save the child?''[23] Finally, in the morally neutral case, judgments of ACO were also significantly different from judgments of free will—again, whereas a large majority had judged that Jeremy goes jogging of his own free will, a majority (57%) answered ''no'' to the question: ''Do you think he could have chosen *not* to go jogging?''[24] (See Figure 3.)

We offer two related interpretations of these interesting results. First, the results may suggest that some folk have Frankfurtian intuitions (see note 6). That is, they think, as suggested by Frankfurt (1969), that an agent's action may be free and responsible without the agent's having the ability to do otherwise—though this trend did not appear with the blameworthy actions.[25] So perhaps these intuitions are more pronounced regarding agents who perform praiseworthy (or morally neutral) actions. This would support Wolf's (1980) asymmetry thesis—roughly, that we tend to judge an agent to be blameworthy only if we believe he could do otherwise, but we are willing to judge an agent to be praiseworthy even if we believe he could *not*

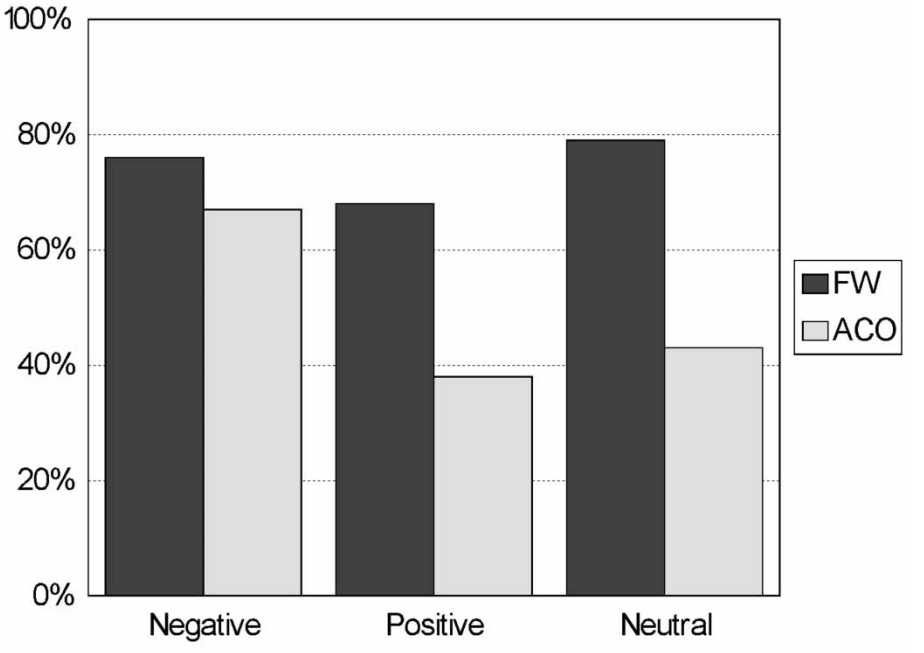

Figure 3 Comparison of Free Will and ACO Judgments. Comparison of percentage of participants who judged that Jeremy acted of his own free will (FW) to those who judged that he could have chosen otherwise (ACO) in robbing a bank (negative), saving a child (positive), and going jogging (neutral).

do otherwise.[26] In any case, it appears that judgments of free will and responsibility can diverge from judgments about the ability to do otherwise.

The main finding in Study 1 is that most people do *not* judge determinism—at least as described in this scenario—to be incompatible with an agent's acting of his own free will or with his being morally responsible for his actions.[27] Repeatedly, we found that the number of participants making judgments *inconsistent* with incompatibilist intuitions is two to three times greater than the number making judgments indicative of incompatibilist intuitions.[28]

2.2. Study 2: Fred and Barney Cases

One might object that in the Jeremy cases, by trying to avoid a question-begging description of determinism, we did not make the deterministic nature of the scenario salient enough to the participants.[29] Perhaps they were more focused on the fact that Jeremy's actions were predicted by the supercomputer than the fact that the prediction was made based on deterministic laws. (If so, it would still be an important result that most people do not judge such *predictability* to conflict with free will and responsibility.[30]) To explore this possibility, we developed another scenario that, we believe, presents determinism in a different and more salient way in

that it points out that the agents' behavior is sufficiently caused by factors beyond their control (i.e., genes and upbringing):

> Scenario. Imagine there is a world where the beliefs and values of every person are caused completely by the combination of one's genes and one's environment. For instance, one day in this world, two identical twins, named Fred and Barney, are born to a mother who puts them up for adoption. Fred is adopted by the Jerksons and Barney is adopted by the Kindersons. In Fred's case, his genes and his upbringing by the selfish Jerkson family have caused him to value money above all else and to believe it is OK to acquire money however you can. In Barney's case, his (identical) genes and his upbringing by the kindly Kinderson family have caused him to value honesty above all else and to believe one should always respect others' property. Both Fred and Barney are intelligent individuals who are capable of deliberating about what they do.
>
> One day Fred and Barney each happen to find a wallet containing $1000 and the identification of the owner (neither man knows the owner). Each man is sure there is nobody else around. After deliberation, Fred Jerkson, because of his beliefs and values, keeps the money. After deliberation, Barney Kinderson, because of his beliefs and values, returns the wallet to its owner.
>
> Given that, in this world, one's genes and environment completely cause one's beliefs and values, it is true that if Fred had been adopted by the Kindersons, he would have had the beliefs and values that would have caused him to return the wallet; and if Barney had been adopted by the Jerksons, he would have had the beliefs and values that would have caused him to keep the wallet.

Despite this seemingly potent description of complete causation by genes and environment, a significant majority of participants (76%) judged both that Fred kept the wallet of his own free will and Barney returned it of his own free will (case 9).[31] This response pattern was very similar to the pattern of participants' judgments about free will in the Jeremy cases, suggesting that this scenario probed similar intuitions about the relationship between deterministic causation and free will.[32]

We also tested whether participants judge that Fred is "morally blameworthy for keeping the wallet" and that Barney is "morally praiseworthy for returning the wallet" (case 10). For most participants (94%) these judgments were consistent, and the response patterns were not significantly different from the judgments we collected about free will: 60% judged that Fred is blameworthy and 64% judged that Barney is praiseworthy.[33] Finally, we tested whether participants judge that Fred and Barney could do otherwise than they did (case 11). Again, results closely tracked judgments of freedom and responsibility, with 76% of participants responding that both Fred and Barney could have done otherwise.[34] (See Figure 4.)

2.3. Discussion

We suggest that—in the absence of further studies contradicting our results or alternative explanations of them—these studies suggest that ordinary people's pre-theoretical intuitions about free will and responsibility do *not* support incompatibilism. It appears to be false—or certainly too hasty—to claim that

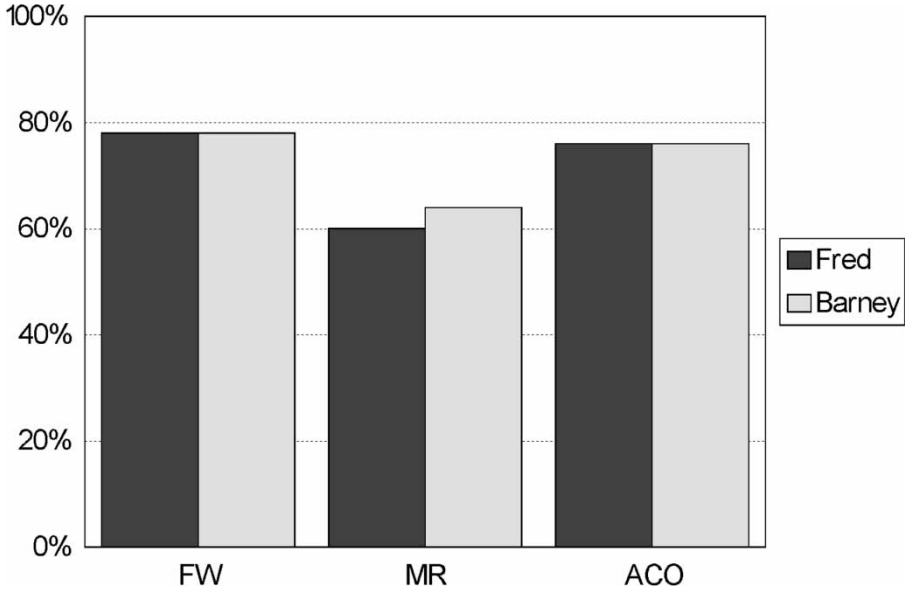

Figure 4 Judgments of Free Will, Moral Responsibility and ACO. Percentage of subjects who judged that Fred and Barney acted of their own free will (FW), were morally responsible for their actions (MR), and could have chosen otherwise (ACO).

"most ordinary persons . . . believe there is some kind of conflict between freedom and determinism" (Kane, 1999, p. 218), and that "We come to the table, nearly all of us, as pretheoretic incompatibilists" (Ekstrom, 2002, p. 310). Rather, when pretheoretic participants considered an agent (Jeremy) whose action is unerringly predicted based on the state of the universe before his birth and the laws of nature, a significant majority judged that the agent acts of his own free will and is morally responsible for the action, and this tendency shows up when the action in question is morally negative, positive, or neutral. Hence, most participants did *not* recognize an incompatibility between determinism described in these terms and free will or moral responsibility. Furthermore, when participants considered agents (Fred and Barney) whose beliefs, values, and actions are completely caused by their genes and upbringing, a significant majority judged that these agents act of their own free will and are morally responsible for their actions. Hence, most participants did *not* recognize an incompatibility between determinism described in these terms and free will or moral responsibility. Judgments about the ability of agents in a deterministic scenario to do or choose otherwise were more complex and, given their important role in the philosophical debates, clearly demand more research.[35]

Notice that, by claiming that most people do not express incompatibilist intuitions in these cases, we are *not* endorsing the stronger claim that the folk *do* have compatibilist intuitions. To be sure, compatibilists should take comfort from results suggesting that their opponents do not have the support of pre-theoretical intuitions on their side. Our results, however, were not unanimous—we consistently found

a non-negligible minority of participants offering incompatibilist judgments. Such results may support the idea that individuals have conflicting intuitions about free will or moral responsibility.[36] Or they may indicate that intuitions vary significantly between different individuals who themselves have consistent intuitions.[37]

Nonetheless, we think that our results place the burden of proof on the shoulders of incompatibilists. Incompatibilists are especially apt to cite folk intuitions in support of their view, in part because their conception of free will is more metaphysically demanding and therefore requires extra intuitive support to offset the strength of their claims.[38] Put simply: if our ordinary intuitions do not demand indeterminism, then why should our theories? If incompatibilists claim that compatibilism is a ''wretched subterfuge,'' a radical revision of commonsense beliefs, then we recommend that some empirical evidence should be offered to back up this claim.

3. Objections, Replies, and Philosophical Implications

Our studies falsified the prediction that most lay persons will express incompatibilist intuitions by judging that agents in a deterministic scenario lack free will and moral responsibility. Again, however, we view these results as preliminary, not conclusive, and hence as motivation for further research on folk intuitions about freedom and responsibility and for further consideration of the role such intuitions should play in the free will debate. One shortcoming of our studies is that they were limited to a college student population. They should be replicated using participants with more diverse educational backgrounds and socio-economic statuses. Ideally, they would also be replicated with participants from other cultures to test whether there are important differences among various cultures' conceptions of free will and moral responsibility.[39] Furthermore, the method of surveying judgments in response to scenarios, though useful for our purposes, should be combined with other methods, including coding participants' responses in more open-ended interviews about the relevant concepts (see Nahmias et al., 2004), and experimentally investigating relevant behavioral responses to situations that involve people's ascriptions of moral responsibility (e.g., praising, blaming, excusing).[40]

A potential problem more specific to our studies is that the presence of determinism might not have been salient enough in the scenarios. We have already explained why we think it would be problematic to use the word 'determinism' to test whether ordinary people have incompatibilist intuitions. But if the description of determinism is overly watered down, then people may simply fail to notice its presence and therefore botch the (supposed) inference to the conclusion that agents in the scenario are not free and responsible. We agree that the more salient determinism is in the scenarios, the more significant the results would be (see Black & Tweedale, 2002). Amping up determinism, however, is not as easy as it might seem. First of all, using a technical definition of 'determinism' would likely be impractical. For instance, using van Inwagen's (1983, pp. 58–65) definition of determinism and telling participants that, in Jeremy's world, the past and the laws of nature entail any

true proposition, would require providing participants with enough education to ensure that they comprehend the technical notions of 'proposition' and 'entail', along with understanding necessity and indicative conditionals. It is unlikely that participants could be sufficiently educated on these issues without significant training. (We discuss below problems with using participants trained in philosophy for research like ours).

Attempts to make determinism more salient *without* teaching participants technical notions run the risk of adding new factors that may themselves mask whatever effects determinism has on people's judgments with more obvious threats to free will. For instance, determinism would be very salient if we changed the Jeremy scenario so that his action was caused by a brilliant neuroscientist who manipulated the initial conditions of Jeremy's life in a way that she knew would inevitably lead to his specific action. Here, the deterministic causation is clear—but so is the presence of a 'covert controller'. If participants then judged that Jeremy did not act freely or was not morally responsible for his act, we would not be able to determine whether it was because their intuitions about free will and moral responsibility are responding to the *determinism* in the scenario or to the presence of an active manipulator.[41] It would also beg the question to describe determinism as entailing that the agent could not do otherwise, since some compatibilists disagree with incompatibilists about whether the ordinary conception of 'can' relevant to freedom and responsibility is inconsistent with determinism. Indeed, our diverse results from participants' judgments about the ability to do or choose otherwise suggest that further research will be required to understand this concept.[42]

With that being said, if one is able to find a way to increase the salience of determinism without masking it with a different free-will threat, we welcome the attempt. If turning up the volume on the ''determinism knob'' of the scenarios *does* cause participants to withdraw their judgments of free will and moral responsibility (and if this is clearly not a result of masking determinism), then we would withdraw our current interpretation of the data. If a more clearly deterministic case does *not* result in more judgments of unfree and unresponsible, though, our interpretation is strengthened. In any case, the claim that people will withdraw judgments of freedom when determinism is made more salient is an empirical claim. Philosophers are well positioned to develop the relevant thought experiments to test this claim. But how people will respond to such scenarios cannot be settled from our philosophical La-Z-Boys. This response generalizes to any empirical or methodological objections to our studies. We are not closing the door on the question of whether the folk have incompatibilist intuitions; on the contrary, we are opening the door to further research exploring such questions.

One significant problem facing any such test of folk intuitions about the relationship between determinism and free will is that it requires participants to reason conditionally. To determine whether or not people believe free will is compatible with determinism, it is not enough to see whether (a) they believe we have free will and (b) they believe our world (or at least its human population)

behaves in a deterministic way. If people believe both (a) and (b), that would strongly suggest they have a compatibilist conception of free will. But if they believe (a) but not (b), that would *not* show they have an incompatibilist theory—after all, they may have reasons unrelated to free will for believing the world is not deterministic. Rather, incompatibilism is the belief that free will is possible *only if* determinism does not obtain. So, to test this belief directly, we must see what people believe about agents' freedom and responsibility on the assumption that determinism *does* obtain.

In our attempts to make determinism salient to participants, we described scenarios that many found implausible or impossible. Especially in the Jeremy study, almost all participants responded that they believed the scenario is impossible, for a variety of reasons (see note 17). Despite our efforts to induce participants to make their judgments based on a conditional acceptance of the scenario, it is likely that some did not do so, as evidenced by those who missed the manipulation check on the back of the surveys, which tested whether they recognized the deterministic nature of the scenario. Some of these participants may even have been expressing incompatibilist intuitions, e.g., by asserting that Jeremy acts of his own free will and then (mis)responding that the scenario could therefore *not* have been describing a fully deterministic situation. Hence, the challenge is to describe determinism in a way that participants find salient, intelligible, and somewhat believable.

Our second scenario (Fred and Barney cases) seems to have satisfied these conditions better than our first, given that fewer participants responded that it was impossible and fewer missed the manipulation check. Another scenario we have used for pilot studies involves the 'rollback' thought experiment initially advanced to probe intuitions about whether indeterminism can help secure free will (see Kane, 1999, p. 222). The scenario asks participants to imagine that the universe is re-created at a certain point in time just before a particular event (e.g., an agent's decision) with all conditions and laws being identical. Then participants can be probed on whether they believe our world is deterministic (would the same event occur every time?) and whether they believe physical events are deterministic whereas agent's decisions are not (would the same *decision* occur every time?). Finally, determinism can be stipulated (e.g., the same decision occurs every single time) to test whether participants think such a decision (their own or another's) is free and deserves praise or blame. Such scenarios can also be varied to test whether indeterminism (e.g., one decision occurs half the time and another occurs half the time) increases or decreases participants' judgments of free will and moral responsibility.[43]

Even if we are able to get participants to recognize determinism and respond based on the assumption of determinism, it may be argued that most of them are unable to recognize the *implications* of determinism for free will and moral responsibility. This problem can take a *motivational form*—participants don't *want* to see the implications of determinism—or a *cognitive form*—they cannot *comprehend* the (alleged) connection between determinism and free will.

The motivational problem suggests that people may be so attached to being free and responsible—and to holding other people morally responsible—that it is very difficult to get them to judge that agents (at least agents resembling them) are not free

and responsible.[44] This kind of "attachment" may, for example, hinder participants from recognizing the deterministic elements of scenarios like the ones we used. The idea here is that a person who harbors incompatibilist sympathies, and who views the agent(s) in a scenario as being sufficiently like himself, may fail to acknowledge the deterministic aspects of a scenario. However, if we can ensure that participants are in fact recognizing the deterministic nature of a scenario, this motivational problem should not be an issue. In this case the participants either (a) express an *incompatibilist* intuition by judging that the agent is *not* like them because he is determined (and hence respond that he, unlike them, is *not* free and responsible), or (b) express a *compatibilist* intuition by responding that the determined agent is, like themselves, free and responsible. After all, people do not need help recognizing that certain types of causal processes, such as brainwashing or direct neural manipulation, compromise free will and moral responsibility, presumably because they judge that these processes are unlike those that govern their own behavior (see note 41).

Another possibility is that people *want* incompatibilist free will, but, if they find themselves in a situation (actual or hypothetical) without it, they will settle for a less satisfying kind of compatibilist free will instead. On this reading, folk come to the table with the intuition that to be free *in the fullest sense*, determinism must be false, but a less robust (but better-than-nothing) sort of freedom is available if determinism turns out to be true.[45] If accurate, this would suggest people have a libertarian conception of free will but, in lieu of that, they would accept a compatibilist conception as a moderate revision rather than giving up entirely on the idea that we are free and responsible agents. Whether or not the folk actually *have* this complex psychology is an interesting question that would require studies that go beyond what we have done here.

In any case, there is a lesson buried in this discussion. We commonly hear that ordinary people have (or fail to have) something that might be called "the incompatibilist intuition" (Pereboom, 2001, p. 89) or "natural incompatibilist instincts" (Kane, 1999, p. 218). It is not clear what the *content* of this 'intuition', or the nature of these 'instincts' is supposed to be. Is it a propensity to call determined people unfree? Is it a desire to have a roughly incompatibilist rather than a compatibilist kind of freedom, although either deserve the name 'freedom'? Or is it something else entirely? In addition to knowing what theoretical work philosophers' appeals to folk intuitions are supposed to be doing, we also need to know what the *content* of those intuitions is supposed to be.

One suggestion we expect to hear is that "the incompatibilist intuition" is the propensity of participants to judge an agent in a deterministic scenario to be unfree but only once they recognize the *implications* of determinism. Thus, in order to see if people have the incompatibilist intuition, we need to help them see these implications—help them overcome what we called above the *cognitive form* of the problem. But getting people to "see the implications" of a case—especially a case like this, where the referent of '*the* implications' is so hotly contested—is tantamount to giving the arguments in favor of one position or another. These arguments rely on premises that are themselves controversial—and that are likely to be supported by

further appeal to intuitions. Furthermore, this does not seem a very plausible view of what the incompatibilist intuition should be: if it takes a basic incompatibilist argument to make incompatibilism the ''intuitive view'', then it seems that it is the incompatibilists who are talking the folk *into* incompatibilism. At a minimum, the compatibilists would not be talking the folk *out* of anything, in contrast to Kane's (1999) claim that ''ordinary persons have to be talked out of [their] natural incompatibilism by the clever arguments of philosophers'' (p. 218).

We have no doubt that many ordinary people, after some exposure to incompatibilist arguments, will make incompatibilist judgments. Similarly, people exposed to compatibilist arguments are likely to make compatibilist judgments. But as soon as philosophers provide such arguments, they are influencing people's views and shaping their intuitions. Even if a philosophy teacher is trying to maintain neutrality, it is unlikely her views will not influence the way she presents the material to her students.[46] And even if her own views don't influence her students, the material itself consists of conflicting claims meant to persuade people; getting a balanced view of the debate would require exposure to the philosophical literature seldom found outside of graduate school. Such an education would, however, compromise one's status as one of the ''folk'' whose intuitions we are trying to understand. Of course, it may be possible to *test* whether people's judgments change after they have studied the free will debate and whether they change in different ways depending on how they study it. Perhaps a compatibilist teacher, despite any efforts to present both sides of the debate fairly, would influence his students to respond with more compatibilist judgments, whereas an incompatibilist teacher would produce students who respond with more incompatibilist judgments.[47] If so, it would be difficult to get pre-theoretical intuitions from people who have been exposed to the relevant philosophical arguments.

Perhaps we should instead be interested in the *reflective* intuitions of well-trained philosophers rather than folk uninformed about the complex philosophical issues surrounding free will.[48] Van Inwagen (1992, p. 58) has suggested that arguments about free will should be pitched at an audience whose members have been exposed to the best arguments on both sides of the debate but have not reached a firm conclusion about the compatibility question (see also Lycan, 2003; Mele, 1995). If our project motivates philosophers to adopt this view, that would at least be a significant result: no longer could any side claim that they have common sense on their side, since the well-trained philosopher's sense is not exactly *common* sense.

The worry here is that a conceptual analysis that fits the reflective intuitions of well-trained philosophers may have ''nothing more than a philosophical fiction as its subject matter'' (Mele, 2001, p. 27). We face a dilemma. If 'free will' is beholden only to the post-theoretic intuitions of philosophers, then it is a technical concept, not unlike 'determinism', 'validity', or 'skepticism' in philosophy (or 'quark' or 'momentum' in physics). If it is a technical concept, then philosophers should not be particularly interested in folk intuitions about it. But this route seems problematic since the concept of free will most philosophers are interested in is the one that is necessary for moral responsibility and attributions of praise and blame. And it

would be strange indeed if people's actual notions of moral responsibility and their actual practices of attributing praise and blame had no important relationship to philosophers' theories about them. Even if 'free will' is a technical concept, it will have to be at least constrained by the non-technical concepts and practices of responsibility attribution if it is to do the work the philosophical community has set for it. On the other hand, if 'free will' is not meant to be a technical philosophical concept—or if it is, but it is constrained by folk intuitions about moral responsibility—then philosophers should be interested in prephilosophical judgments relevant to the concept. In that case, philosophers should be interested in the *facts* about such judgments and the concepts they reveal. To understand what those facts are, we must consider surveying the folk; thus, folk intuitions are once again relevant.

We conclude with one more reason why it is unlikely philosophers will, or should, give up talking about folk concepts and intuitions regarding free will and moral responsibility. A significant recent trend in the free will debate involves philosophers on all sides of the debate analyzing the concept of free will in the light of what we value—they aim, in Dennett's (1984a) language, for a concept of free will "worth wanting." Though this often means looking for a concept that supports our practices of ascribing moral responsibility, it also involves the values we place on human dignity, uniqueness, creativity, hope, and love (see, e.g., Kane, 1996, ch. 6). And philosophers have also noted that moral responsibility itself involves a range of issues, from reactive attitudes to practices of punishment and reward to moral rehabilitation (see, e.g., Pereboom, 2001). Considering the issue in this way has led to the recognition that various conceptions of free will may secure various things we value, perhaps to different degrees, and that various putative threats to free will, such as determinism, may threaten different things we value, perhaps to different degrees. For instance, as Randolph Clarke (2003) suggests, it may be that an event-causal libertarian theory (like Kane's) can secure some aspects of dignity we value even if it does not undergird moral responsibility any better than compatibilist accounts. Similarly, as Fischer and Ravizza (1998) argue, it may be that determinism precludes the freedom to act otherwise but that such freedom is not required for moral responsibility. Fischer (forthcoming) has written, "The discovery that causal determinism is true would significantly alter our picture of ourselves. . . . But I do not believe we would need entirely to jettison . . . deliberation, moral responsibility, and judgments of deontic morality [which] are compatible with causal determinism and the lack of free will (in the sense involving alternative possibilities)" (p. 23).

These moves suggest that certain theories of free will and moral responsibility may require us to revise some, but not all, of our current concepts, beliefs, and practices about freedom and responsibility (see Vargas, forthcoming). But in order to know whether a particular theory demands revision (or even elimination) of our concepts, beliefs, or practices, we have to *know* what these are. And though armchair speculation may help here, it is unlikely to be sufficient—especially given the vigorous debates among philosophers about what is intuitive when it comes to freedom and responsibility. Rather, we should make some attempt to get systematic,

empirical data about the relevant concepts, beliefs, and practices. Though this may ultimately be work best left to psychologists and sociologists, philosophers will have to help lay out exactly which issues should be tested and how to set up scenarios to test the relevant intuitions.

Because any attempt at conceptual revision, especially in light of various values we attach to freedom concepts, will depend on our understanding of how people actually use these concepts and what values they actually attach to them, we submit that our general methodological proposal may be particularly important to the current free will debate. Beyond the particular studies we carried out here in an attempt to get at folk judgments about incompatibilism, we need to achieve a better understanding of a wide range of folk judgments, practices, and intuitions concerning freedom and responsibility. As Vargas (forthcoming) has written, ''Our ability to determine whether a theory [of free will and responsibility] is revisionist turns on the extent to which we have correctly characterized responsibility practices, attitudes, and beliefs.''

Furthermore, we hope this paper will motivate philosophers in the free will debate to do two things. First, we think philosophers should be more explicit about how they think their theories of free will and moral responsibility align with—and diverge from—people's actual intuitions and practices. And second, to the extent that they claim their theories *do* align closely with these intuitions and practices, they should be more explicit and careful about their reasons for believing such alignment exists. We are not suggesting that philosophers should make a practice of surveying the folk about every question of interest. They can at least, however, perform the following thought experiment: ''How would I determine whether the folk would agree with this particular claim about free will and moral responsibility? If it turned out they *did* agree, what evidential value would that have for the claim? If, on the other hand, it turned out they did *not* agree, what problems would that raise for the claim, and how might I respond to them? And if it simply does not matter whether or not prephilosophical people would agree with the claim, why is that?'' Considering questions like these will help ensure that, as we continue to argue about free will, our arguments will be about more than just a philosophical fiction.

Acknowledgements

For their helpful suggestions, we would like to thank Joshua Knobe, John Doris, Tamler Sommers, Mike McKenna, Brandon Schmeichel, George Graham, two anonymous referees, and, especially, Shaun Nichols, Manuel Vargas, and Al Mele.

Notes

[1] Authorship is equal.

[2] For discussions see, for example, Jackson (1998) and essays in DePaul and Ramsey (1998).

[3] See, for example, Adams and Steadman (2004), Doris and Stich (forthcoming), Knobe (2003a, 2003b, 2004), Mele (2003), Nadelhoffer (2004, 2005), Nahmias, Morris, Nadelhoffer, and Turner (forthcoming), Nichols (2004), Nichols, Weinburg, and Stich (2003), Weinburg, Nichols, and Stich (2001), Woolfolk, Doris, and Darley (forthcoming), and Turner and Nahmias (in preparation).

[4] However, Jackson's example of an uncontroversial case, Gettier examples in epistemology, has in fact been challenged by empirical results suggesting that people from different cultures or of different socio-economic statuses have different intuitions about such examples (see Nichols et al., 2003; Weinburg et al., 2001). See also Stich and Weinberg (2001).

[5] See also Smilansky (2003, p. 259), Pereboom (2001, p. xvi), O'Connor (2000, p. 4), and Campbell (1951, p. 451).

[6] Frankfurt cases aim to pump the intuition that an agent can be responsible for an action even if he could not do otherwise: for instance, if there were a neuroscientist ready to manipulate the agent to do *A* if he were not going to do it on his own, then the agent could not do otherwise than *A*, but if the agent did *A* on his own without the neuroscientist intervening, it seems the agent is nonetheless responsible for *A*-ing.

[7] Lycan (2003) argues that it is the *incompatibilist* who must provide a positive argument for her position, in part because "numerous commonsense claims of free action will always be more plausible than are the *purely philosophical* premises" of incompatibilist arguments (p. 120).

[8] Compatibilists and libertarians offer a parallel set of conflicting, but empirically untested, claims about the phenomenology of free will—for instance, our experiences of the ability to do otherwise and of the self as source of actions (see Nahmias, Morris, Nadelhoffer, & Turner, 2004).

[9] See also Nahmias et al. (forthcoming), Nichols (2004), and Nichols and Knobe (in preparation).

[10] Compare Fischer and Ravizza's (1998) conception of how to understand moral responsibility: "we shall be trying to articulate the inchoate, shared views about moral responsibility in (roughly speaking) modern, Western democratic society . . . . Here we shall be identifying and evaluating 'considered judgments' about particular cases—actual and hypothetical—in which an agent's moral responsibility is at issue" (pp. 10–11).

[11] For a helpful discussion of 'revisionism' in the free will debate see Vargas (forthcoming).

[12] For instance, Graham and Horgan (1994) offer a contextualist sort of explanation for why people might express incompatibilist intuitions in certain philosophical contexts, and Velleman (2000) developed an error theory for why our *epistemic* freedom leads us to believe we have *metaphysical* (libertarian) freedom.

[13] We should note that two of the four authors are compatibilists and two are incompatibilists. Every effort was made not to prejudice the data through misleading experimental design, and the diversity of opinions within the team was beneficial to this end (though frustrating at times!). We also sent our surveys to a dozen or so philosophers with various views in the free will debate and received no responses suggesting that the surveys were in any significant sense misleading or problematic.

[14] Examples of participants' definitions of 'determinism' include: "Being unable to choose", "That people have a set fate", and "The lack of free will". Many others thought it meant 'determined', as in 'resolute'. Fischer (1994, p. 152) predicted that the phrase 'causal determinism' may seem threatening to people but that this alone would not indicate they have incompatibilist intuitions.

[15] Laplace defines 'determinism' in terms of an intelligent being such that if it had complete knowledge of the laws of nature and the current state of the universe, it could know all past events and predict all future events. Van Inwagen's (1983) technical definition of 'determinism' is related: a proposition describing the complete state of the universe at one time and the laws of nature logically entails a proposition describing the complete state of the universe at any other time. Laplacean determinism entails van Inwagen determinism, but the converse is not true.

[16] Participants were drawn from an Honors student colloquium and several introductory philosophy classes at Florida State University (before studying the free will problem).

Any participants who indicated that they had taken a previous college philosophy course were excluded from the results. We also excluded those participants who missed the manipulation check and the few who answered "I don't know" in response to the experimental question.

[17] The reasons participants offered for believing the scenario to be impossible were wide-ranging, including, for instance, that the computer could never acquire that much data, that people could undermine the predictions by learning about them, and that chaos theory or quantum theory makes such predictions impossible. See section 3 for a discussion of the problem of surveys that require conditional reasoning.

[18] This result (16:5) is statistically significant using a χ^2 goodness-of-fit test: $\chi^2(1, N = 21) = 5.76$, $p = 0.016$.

[19] For all of these cases we used different sets of participants to avoid any order or interference effects due to participants' attempting to keep their judgments consistent across different cases—i.e., any given participant read just one survey and answered just one experimental question.

[20] Case 2 results (15:7) approach statistical significance: $\chi^2(1, N = 22) = 2.9$, $p = 0.088$). Case 3 results (15:4) are statistically significant: $\chi^2(1, N = 19) = 6.37$, $p = 0.012$. Binomial proportion comparison tests show no significant differences between the response patterns of participants in case 1 and case 2 ($p = 0.56$), or between those in case 1 and case 3 ($p = 0.84$).

[21] Case 4 results (15:3) are significant: χ^2 $(1, N = 18) = 8$, $p = 0.005$. Case 5 results (16:2) are significant: $\chi^2(1, N = 18) = 10.9$, $p = 0.001$. Binomial proportion comparison tests show no significant differences between the response patterns of participants in case 1 and case 4 ($p = 0.58$), between those in case 1 and case 5 ($p = 0.30$), or between those in case 2 and case 5 ($p = 0.12$).

[22] Binomial proportion comparison tests between the response patterns of participants in case 1 and those in case 6 (14:7) show *no* significant difference ($p = 0.50$).

[23] Binomial proportion comparison tests between the response patterns of participants in case 2 and those in case 7 (8:13) show that there *was* a statistically significant difference ($p = 0.05$).

[24] Binomial proportion comparison tests between response patterns of participants in case 3 and those in case 8 (6:8) show a statistically significant difference ($p = 0.03$).

[25] To test this interpretation, it would be better to have the *same* participants making judgments about both the ability to do otherwise and free will (or moral responsibility). See Woolfolk et al. (forthcoming) for empirical results suggesting that ordinary people have Frankfurtian intuitions.

[26] We suspect that it is particularly difficult to probe intuitions about modal concepts such as the ability to do otherwise (see note 35). See Turner and Nahmias (in preparation).

[27] Participants' explanations of their answers suggest that some of them even had fledgling compatibilist theories in mind. For instance, some suggested that Jeremy acts freely because no outside forces compel him, because he controls his actions or consciously decides to do them, or because the prediction is based on what *he* decides to do. (Similarly, some of the minority who judged that Jeremy is *unfree* offered fledgling *in*compatibilist arguments.)

[28] We should add that in numerous pilot studies as we fine-tuned descriptions of the scenarios, experimental questions, and manipulation checks, we were surprised by the consistency of the results: in almost every set of surveys, 70–85% of participants judged that Jeremy acts of his own free will or that he is morally responsible for his action.

[29] We should note that we tried several variations of the wording of the Jeremy case, including one that said, "The computer then deduces from this information and the laws of nature that it is *physically impossible* for Jeremy to do anything other than to rob Fidelity Bank [save the child, etc.]" and results showed the same trends (as reported in note 28). For further discussion of this concern about the salience of determinism, see section 3.

[30] We wonder whether some participants, having reconciled themselves to the problem of God's foreknowledge and free will, were less inclined to see the supercomputer's foreknowledge as a threat to free will. We obtained information about participants' self-reported religiosity, but we have not yet correlated this information with the results.

[31] Methods for study 2 were the same as study 1 except that participants answered *two* experimental questions (one about Fred and one about Barney). Fewer participants judged that the scenario was impossible and fewer missed the manipulation check (three participants who answered differently for Fred and Barney were not included in the analysis of results). Results for case 9 (26:8 judging that both agents acted of their own free will) were statistically significant using a χ^2 goodness-of-fit test: χ^2 (1, $N = 34$) = 9.53, $p = 0.002$.

[32] A binomial proportion comparison test shows no significant difference between the response patterns of participants in case 1 and those in case 9 ($p = 0.98$).

[33] Comparison tests show no significant differences between the response pattern of participants in case 9 and the pattern of *blameworthy* judgments in case 10 (21:14, $p = 1.06$), or in the response patterns of participants in case 9 and the pattern of *praiseworthy* judgments in case 10 (23:13, $p = 0.25$). Interestingly, a smaller majority of participants in these cases judged that the agents were morally responsible than those judging Jeremy to be morally responsible (in cases 4 and 5).

[34] Comparisons between the response patterns of participants in case 8 and those in case 11 (20:11) were not significantly different ($p = 0.29$). Because we asked the same participants about both Fred's and Barney's ability to do otherwise, and all participants made consistent judgments, we were unable to confirm the interesting results from study 1 suggesting that such judgments vary depending on whether the action is blameworthy or praiseworthy/neutral.

[35] As we suggested above, in some cases the pattern of responses suggests to us that people may have Frankfurtian intuitions whereby they are willing to assign blame despite believing the agent could not do otherwise. In other cases, the responses suggest people may be employing a conditional conception of the ability to do otherwise, such that the agents could do otherwise despite their actions being determined.

[36] Compare Double (1991, 1996). We take very seriously the possibility that free will and moral responsibility are best analyzed as context-relative concepts (see Graham & Horgan, 1994). If so, empirical research aimed at mapping the contours of the relevant contexts would be important.

[37] This would suggest an interesting explanation (perhaps found in the work of William James and Ludwig Wittgenstein) for the long-standing free will stalemate. After decades of argument, compatibilists and incompatibilists tend to retain conflicting intuitions about crucial disputes in the debate; perhaps they *end up* where they do in part because they *start off* where they do—as ordinary folk with different intuitions about the concept of free will and its relationship to determinism.

[38] Libertarian theories require, at a minimum, indeterministic causation—in just the right place (Kane, 1996)—and often also an ontologically distinct type of causation, agent causation, that demands either substance dualism or emergent causation (O'Connor, 2000). Furthermore, in general, claims about the *in*compatibility of any two concepts require more evidence to establish than claims about their compatibility, in part because concepts that are not obviously incompatible should be assumed to be compatible barring an argument to the contrary. (See Lycan, 2003; Nahmias et al., forthcoming.)

[39] See Weinburg et al. (2001) and Nichols et al. (2003) for experimental work on epistemological intuitions suggesting that these intuitions vary among people with different educational backgrounds and between people from American and Asian cultures.

[40] Much of the work on people's "folk theories" of biology, physics, and psychology (e.g., theory of mind) involves such experiments. Some find differences between participants'

reported intuitions about the relevant situations and their actual behavior in such situations (see, for example, Keysar, Lina, & Barr, 2003).

[41] We ran pilot studies suggesting this. Scenarios describing an agent who was neurally manipulated to deliberate and act in a certain way elicited almost unanimous judgments that he was neither free nor responsible. When combined with the results from studies 1 and 2, this suggests to us that people do *not* have the intuition, sometimes advanced by incompatibilists, that determinism is no different than such covert manipulation (see Pereboom, 2001), though further tests of this question would be useful.

[42] Similar worries would apply to any description of determinism that suggests that the laws of nature control or constrain us, that our actions are fated by the past, that our conscious decisions are causally inefficacious, etc. Such question-begging descriptions of determinism would not show that the folk are incompatibilists about free will and *determinism* but only that they are incompatibilists about free will and constraint, fate, epiphenomenalism, etc. Notice that introductions to the problem of free will often use such misleading images to portray the threat of determinism.

[43] Nichols (2004) used a roughly similar set-up and found that adults and children are more likely to judge that a *person* could behave otherwise than that a physical *object* could. Nichols suggested that these results indicate that people have a theory of agent causation, but for a variety of reasons we think such a conclusion is too hasty. (See Turner & Nahmias, in preparation.)

[44] Compare Peter Strawson's (1962) claim that the reactive attitudes are "natural, original . . . in no way something we could choose or could give up" (note 7).

[45] This view is similar to van Inwagen's (1983) view, dubbed "metaphysical flip-flopping" by Fischer & Ravizza (1998), that if he were convinced determinism were true, he would become a compatibilist rather than a skeptic about free will.

[46] Kane (1999) suggests that students "subjected to [a compatibilist] argument may have the uneasy feeling they are being had" (p. 218). O'Connor (2000) writes, "Does freedom of choice have this implication [that causal determinism is false]? It seems so to the typical undergraduate on first encountering the question" (p. 4). We wonder whether these philosophers' students have these reactions in part due to the incompatibilist sympathies of their professors.

[47] In fact, we ran one such test. One of the authors (E.N.), who happens to be a compatibilist, ran the Fred and Barney survey on his Intro students soon after a two-week section on the free will debate. The results, it turned out, were not significantly different from the results garnered from 'untrained' participants: 83% of the 'trained' participants judged that Fred and Barney acted of their own free will, where 76% of untrained participants had made such judgments). Participants were asked whether studying the free will debate had influenced their judgments, and roughly half said it had.

[48] Though we need to be wary of selling the folk short. My (E.N.) furniture delivery men, on discovering I was a philosopher, told me they had solved the old riddle of whether the cup was half full or half empty: "It's half full," they explained, "if it's been filled halfway up and it's half empty if it's been emptied halfway." An excellent example of contextualist semantics!

References

Adams, F., & Steadman, A. (2004). Intentional action in ordinary language: Core concept or pragmatic understanding? *Analysis*, *64*, 173–181.

Black, S., & Tweedale, J. (2002). Responsibility and alternative possibilities: The use and abuse of examples. *The Journal of Ethics*, *6*, 281–303.

Campbell, C. A. (1951). Is "free will" a pseudo-problem? *Journal of Philosophy*, *60*, 441–465.

Clarke, R. (2003). *Libertarian accounts of free will*. New York: Oxford University Press.

Dennett, D. (1984a). *Elbow room: The varieties of free will worth wanting*. Cambridge, MA: MIT Press.

Dennett, D. (1984b). I could not have done otherwise: So what? *Journal of Philosophy*, *81*, 553–565.

Doris, J., & Stich, S. (forthcoming). As a matter of fact: Empirical perspectives on ethics. In F. Jackson & M. Smith (Eds.), *The Oxford handbook of contemporary analytic philosophy*. Oxford, England: Oxford University Press.

Double, R. (1991). *The non-reality of free will*. New York: Oxford University Press.

Double, R. (1996). *Metaphilosophy and free will*. New York: Oxford University Press.

Ekstrom, L. W. (2000). *Free will*. Boulder, CO: Westview Press.

Ekstrom, L. W. (2002). Libertarianism and Frankfurt-style cases. In R. Kane (Ed.), *The Oxford handbook of free will* (pp. 309–322). New York: Oxford University Press.

Fischer, J. M. (1994). *The metaphysics of free will*. Cambridge, MA: Blackwell.

Fischer, J. M. (forthcoming). Responsibility and alternative possibilities.

Fischer, J. M., & Ravizza, M. (1998). *Responsibility and control*. New York: Cambridge University Press.

Frankfurt, H. (1969). Alternate possibilities and moral responsibility. *Journal of Philosophy*, *66*, 5–20.

Graham, G., & Horgan, T. (1994). Southern fundamentalism and the end of philosophy. In M. R. Depaul & W. Ramsey (Eds.), *Rethinking intuition* (pp. 271–292). Lanham, MD: Rowman & Littlefield.

Jackson, F. (1998). *From metaphysics to ethics: A defense of conceptual analysis*. New York: Oxford University Press.

Kane, R. (1999). Responsibility, luck, and chance: Reflections on free will and indeterminism. *Journal of Philosophy*, *96*, 217–240.

Keysar, B., Lina, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, *89*, 25–41.

Knobe, J. (2003a). Intentional action and side-effects in ordinary language. *Analysis*, *63*, 190–193.

Knobe, J. (2003b). Intentional action in folk psychology: An experimental investigation. *Philosophical Psychology*, *16*, 309–324.

Knobe, J. (2004). Intention, intentional action and moral considerations. *Analysis*, *64*, 181–187.

Lycan, W. (2003). Free will and the burden of proof. In A. O'Hear (Ed.), *Minds and persons: Royal Institute of Philosophy supplement* (pp. 107–122). Cambridge, England: Cambridge University Press.

Mele, A. (1995). *Autonomous agents*. New York: Oxford University Press.

Mele, A. (2001). Acting intentionally: Probing folk notions. In B. Malle, L. Moses & D. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 27–43). Cambridge, MA: MIT Press.

Mele, A. (2003). Intentional actions: Controversies, data, and core hypotheses. *Philosophical Psychology*, *16*, 325–340.

Nadelhoffer, T. (2004). The butler problem revisited. *Analysis*, *64*, 277–284.

Nadelhoffer, T. (2005). Skill, luck, and folk ascriptions of intentional action. *Philosophical Psychology*, *18*, 343–354.

Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (2004). The phenomenology of free will. *Journal of Consciousness Studies*, *11*, 162–179.

Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (forthcoming). Is incompatibilism intuitive? *Philosophy and Phenomenological Research*.

Nichols, S. (2004). The folk psychology of free will: Fits and starts. *Mind and Language*, *19*, 473–502.

Nichols, S., & Knobe, J. (in preparation). Moral responsibility and determinism: Empirical investigations of folk intuitions.

Nichols, S., Stich, S., & Weinburg, J. (2003). Metaskepticism: Meditations in ethno-epistemology. In S. Luper (Ed.), *The skeptics* (pp. 227–247). Burlington, VT: Ashgate.

O'Connor, T. (2000). *Persons and causes: The metaphysics of free will*. New York: Oxford University Press.

Pereboom, D. (2001). *Living without free will*. New York: Cambridge University Press.

Pink, T. (2004). *Free will: A very short introduction*. New York: Oxford University Press.

Smilansky, S. (2003). Compatibilism: The argument from shallowness. *Philosophical Studies*, *115*, 257–282.

Stich, S., & Weinberg, J. (2001). Jackson's empirical assumptions. *Philosophy and Phenomenological Research*, *62*, 637–643.

Strawson, G. (1986). *Freedom and belief*. Oxford, England: Oxford University Press.

Strawson, P. (1962). Freedom and resentment. *Proceedings of the British Academy*, *48*, 1–25.

Turner, J., & Nahmias, E. (in preparation). Are the folk agent-causationists? A response to Shaun Nichols' ''The folk psychology of free will.''

van Inwagen, P. (1992). Reply to Christopher Hill. *Analysis*, *52*, 56–61.

van Inwagen, P. (1983). *An essay on free will.* Oxford, England: Oxford University Press.

Vargas, M. (forthcoming). The revisionist's guide to moral responsibility. *Philosophical Studies*.

Velleman, D. (2000). Epistemic freedom. In D. Velleman (Ed.), *The possibility of practical reason*. New York: Oxford University Press.

Weinburg, J., Nichols, S., & Stich, S. (2001). Normativity and epistemic intuitions. *Philosophical Topics*, *29*, 429–460.

Wolf, S. (1980). Asymmetrical freedom. *Journal of Philosophy*, *77*, 157–166.

Wolf, S. (1990). *Freedom within reason*. New York: Oxford University Press.

Woolfolk, R., Doris, J., & Darley, J. (forthcoming). The intuitive scientist as moral cognizer: Constraint, identification, and responsibility.