

Draft, February 6, 2012

Phenomenal Abilities: Incompatibilism and the Experience of Agency*

Oisín Deery (University of British Columbia)

Matt Bedke (University of British Columbia)

Shaun Nichols (University of Arizona)

1 Background

1.1 Introduction

Agents act. They buy detergent at the store, they go to work, they celebrate holidays, they cheat on their taxes. Sometimes we hold agents morally responsible for what they do, or what they fail to do, meting out credit and blame as the occasion merits. In typical cases, when agents act they are thought to have an *ability to do otherwise*. This is a point on which most parties to the free-will debates agree. When it comes to characterizing the ability to do otherwise and asking whether this ability is compatible with determinism, however, there is no consensus.

* We thank the following people for providing commentary on earlier drafts: Eddy Nahmias, Dana Nelkin, Paulo Sousa, David Shoemaker, and Uriah Kriegel. We are also grateful to participants at the New Orleans Workshop on Agency and Responsibility (November 3-5, 2011) for suggestions and comments.

In the ensuing debates, the *experience* as of having an ability to do otherwise occupies a central role.¹ Many libertarians, for instance, maintain that the ability experienced is incompatible with determinism (Campbell 1951; O'Connor 1995). Of course, some compatibilists have challenged this idea (Mill 1865; Grünbaum 1971; Nahmias et al. 2004). Despite the centrality of the *phenomenology of agency* in all this, there has been strikingly little work on its characteristics. Of particular significance, there is almost no empirical work on whether the experience of agency involves a phenomenology of being able to choose among alternative possibilities or whether people take their agentic experiences to have incompatibilist elements.²

This paper reports a series of experiments that investigates the phenomenology of agency. To anticipate, we found remarkably consistent results across three sets of studies: participants regarded their experience of the ability to do otherwise as incompatible with

¹ For some of the literature on the ability to do otherwise, see Moore (1912), Berofsky (2002), J. Campbell (2005), Perry (2004), Vihvelin (2004), Smith (2004), Fara (2008), John M. Fischer (2008), Randolph Clarke (2009). The claim that moral responsibility requires alternative possibilities has been disputed since Frankfurt (1969). However, it is still widely contended that free will requires being able to do otherwise.

² The one exception in the recent literature is a paper by Nahmias and colleagues (2004), which we discuss in section 1.3. Although we challenge their experimental results, we are indebted to them for pioneering the investigation. We also draw on their scholarship in setting out some of the historical statements below.

determinism. Now that we've spoiled any suspense, we will locate the issue in the broader literature.

1.2. The experience of the ability to do otherwise

Let us characterize determinism as follows: a statement of the facts of the world at an instant together with a statement of the laws of nature entail all truths about the world, including those about future human actions.³ Granting that we often feel that we have an ability to act other than we do, for present purposes incompatibilists think that the experience as of an ability to do otherwise is incompatible with determinism, while compatibilists think the opposite.

There are a number of influential introspectors on both sides of this issue. John Searle is a representative incompatibilist:

[R]eflect very carefully on the character of the experiences you have as you engage in normal, everyday human actions. You will sense the possibility of alternative courses of action built into these experiences... that we could be doing something else right here and now, that is, all other conditions remaining the same. This, I submit, is the source of our own unshakeable conviction of our own free will. (1984: 95)

Similarly, Keith Lehrer has claimed that the incompatibilist “accurately describes what I find by introspecting, and I cannot believe that others do not find the same” (1960: 150). Even such a paradigmatic compatibilist as David Hume (1960/1739) agrees with this sentiment when he writes, “There is a false... experience... of the liberty of indifference” (Bk. II, Part III, §II).

³ In the experiments below, we will get at this basic idea in a more intuitive, accessible way by presenting it in terms of causation.

Draft, February 6, 2012

The appeal to an incompatibilist phenomenology plays a particularly important role in libertarianism. Many libertarians maintain both that we experience our agency as incompatible with determinism, and that this experience provides reason to think that our agency defies determinism. C. A. Campbell writes:

Everyone must make the introspective experiment for himself: but I may perhaps venture to report... that I cannot help believing that it lies with me here and now, quite absolutely, which of two genuinely open possibilities I adopt. (1951: 463)

Campbell goes on to argue that, unless we have good reason to doubt the impression that “it lies with me” which of two possibilities I adopt, we should accept the impression to reflect the truth. Timothy O’Connor makes this move as well. First, O’Connor describes the character of the experience of decision-making:

[T]he agency theory is appealing because it captures the way we experience our own activity. It does not seem to me (at least ordinarily) that I am caused to act by the reasons which favor doing so; it seems to be the case, rather, that *I* produce my decision *in view of* those reasons, and could have, in an unconditional sense, decided differently... (1995: 196)

Next, O’Connor says that we should take these experiences to reflect something important about the *nature* of decision-making:

Such experiences could, of course, be wholly illusory, but do we not properly assume, in the absence of strong countervailing reasons, that things are pretty much the way they appear to us? ... Skepticism about the veridicality of such experiences has numerous isomorphs that, if accepted, appear to lead to a greatly diminished assessment

Draft, February 6, 2012

of our knowledge of the world, an assessment that most philosophers would resist.
(1995: 196–197)

A number of compatibilists have challenged the basic phenomenological claim. These compatibilists deny that we experience our agency as incompatible with determinism. John Stuart Mill, for instance, writes,

Take any alternative: say to murder or not to murder... If I elect to abstain: in what sense am I conscious that I could have elected to commit the crime? Only if I had desired to commit it with a desire stronger than my horror of murder; not with one less strong. When we think of ourselves hypothetically as having acted otherwise than we did, we always suppose a difference in the antecedents: we picture ourselves as having known something that we did not know, or not known something that we did know; which is a difference in the external motives; or as having desired something, or disliked something, more or less than we did; which is a difference in the internal motives. (1865: 285)

On Mill's view, the feeling of the ability to do otherwise is always contingent on our supposing that the situation prior to the decision was somehow different. Adolf Grünbaum repudiates any incompatibilist element with equal vigor:

Let us carefully examine the content of the feeling that on a certain occasion we could have acted other than the way we did... Does the feeling we have inform us that we could have acted otherwise under exactly the same external and internal motivational conditions? No, ... this feeling simply discloses that we were able to act in accord with

our strongest desire at that time, and that we could indeed have acted otherwise if a different motive had prevailed at the time. (1952: 672)

Grünbaum's last sentence here gestures at the payoff of denying the phenomenological claim of incompatibilist agency: if Mill and Grünbaum are right, then the feeling of being able to do otherwise is consistent with determinism, and this would undercut a crucial motivation for libertarianism.

This situation might seem to be a dialectical stalemate (cf. Fischer 1994: 84). However, these philosophers are making general claims about the nature of our experience of agency. These are empirical claims, and they can be illuminated by taking up empirical methods.

1.3. Previous Work on the Phenomenology of Free Will

We are not the first to recommend a more systematic investigation that is partly empirical. Nahmias and colleagues (2004) suggest that we find out how people actually tend to describe their agentic experience (what they call the phenomenology of free will), including their experience as of being able to do otherwise:

Taking a cue from recent empirical work on “folk intuitions”, we think the best way to understand the phenomenology of free will—if there is one—is to find out what ordinary people's experiences are like. If this is not possible, philosophers' competing introspective descriptions will remain in yet another free-will stalemate. (164)

Nahmias and colleagues undertook this task in survey studies. Their studies appear to lend some support to the idea that the phenomenology of agency is compatibilist. However, we

think the studies have significant shortcomings, so let us briefly describe one of those studies, and then identify what we find lacking.

In one study, Nahmias and colleagues pitted compatibilism and incompatibilism against each other directly. The study was based on “competing libertarian and compatibilist accounts of our experience of the ability to choose otherwise” (174). Their survey asked participants to imagine (or recall) an experience of making a difficult choice:

Imagine you’ve made a tough decision between two alternatives. You’ve chosen one of them and you think to yourself, “I could have chosen otherwise” (it may help if you can remember a particular example of such a decision you’ve recently made).

Which of these statements best describes what you have in mind when you think, “I could have chosen otherwise”?

- A. “I could have chosen to do otherwise even if everything at the moment of choice had been exactly the same.”
- B. “I could have chosen to do otherwise only if something had been different (for instance, different considerations had come to mind as I deliberated or I had experienced different desires at the time).”
- C. Neither of the above describes what I mean. (2004: 175–76)

The majority of the participants gave the response that fits with compatibilism (i.e., B).

While this study is clearly focused on an issue that divides compatibilists and incompatibilists, there are a number of limitations to the study. First, participants are told to think of a decision and then told to think something else about the decision: that they *could have done otherwise*. It is thus unclear whether their initial recollection *actually* carried with it a sense of an ability to do otherwise. So if people make compatibilist judgments about these decisions, it might be because they are considering cases in which the phenomenology of the ability to do otherwise is absent.

Second, the key question is about experiences sometime in the past, rather than present-focused experiences where the phenomenology of agency is actually present and thus presumably more accessible.

Third, Nahmias and colleagues asked participants about *difficult* decisions, and this presents the opportunity to interpret “could have done otherwise” in confounding ways. Consider Martin Luther’s decision to renounce his writings or be declared an outlaw and heretic. Legend has it that, after praying and consulting with advisors for a day, he said, “Here I stand. I can do no other.”, thereby reaffirming his writings. Luther might have chosen B in Nahmias’s survey. But if he did, we should not conclude that there is no sense of “could have done otherwise” that captures some aspect of Luther’s phenomenology and that is incompatible with determinism. For Luther could have responded as he did to express his commitment to his cause, a commitment that would only change if the considerations before him and his reasons for breaking with the Roman Catholic Church presented themselves differently. This commitment-expressive meaning of “could not have done otherwise” is consistent with other senses of “could have done otherwise”—consider whether Luther thought it was up to him

whether to renounce his views—that might or might not be incompatible with determinism. Difficult decisions are subject to confounds like this, so the above survey does not cleanly address the question whether there is some aspect of the phenomenology of agency that is in tension with determinism.

Fourth and last, it is not clear whether the participants really understand the intended meaning of “even if everything at the moment of choice had been exactly the same” or “only if something had been different”.

We wanted to run more comprehensive studies that fix these shortcomings. The result was the following three studies, which share a common structure. First, participants were asked whether they had an experience as of the ability to do otherwise when faced with a simple decision. Next, they were given a description of determinism. Of course, we did not use the term “determinism”, since that might have conjured up unwanted associations in participants. Rather, we used a technical term—“causal completeness”. To address concerns about comprehension of the materials, we used the familiar psychological technique of *training to criterion*, where we asked a series of questions that tested and, if necessary, corrected, the participant’s understanding of determinism. Participants who passed the training were asked about the compatibility of their experience with determinism. In study 1, this question focused on both a first-person, present-focused experience in a hypothetical deliberative context and a past-focused judgment about such a situation. In study 2, we explored the phenomenology of actual rather than imagined choices. In study 3, we tested whether epistemic phenomenology—the phenomenology of uncertainty—feels incompatible with determinism.

2 Study 1

2.1 Overview

In our first study, we had participants imagine a decision about whether to go left or right on a sled. In one condition, the sledding scenario was set in the future; in the other condition, the scenario was set in the past. After reading the scenario, participants in condition 1 were asked whether they had a feeling of an ability to do otherwise; participants in condition 2 were asked for a retrospective judgment about whether they could have done otherwise. Participants who affirmed feeling (or having) an ability to do otherwise were directed to the training section in which “causal completeness” (i.e., determinism) was explained to them. Participants who passed the training were reminded of their affirmation regarding the ability to do otherwise and asked about consistency with causal completeness.

Our prediction was that when asked about the phenomenology of imagined decision making, participants would tend to affirm a feeling of an ability to do otherwise and also regard this feeling as incompatible with determinism.; but when asked for a retrospective judgment about the ability to do otherwise, we predicted that participants would be less likely to treat the ability to do otherwise in an incompatibilist way.⁴

⁴ See e.g., van Inwagen 1983 (8–13) for an overview of other uses of “can”.

Method:

Participants:

84 participants were initially recruited online through the Mechanical Turk (MTurk) website.⁵ The survey itself was conducted using SurveyMonkey. 2 participants did not complete the survey. They were excluded from the analysis.

Materials:

Each condition had three parts.

Part 1: The ability to do otherwise

Participants were presented with a vignette and a question about the ability to do otherwise. For condition 1, this went as follows:

Please read the following passage, and answer the questions that follow as best you can:

Imagine that you are sledding down a snowy path on a mountainside. Your sled has a steering mechanism that allows you to control the direction of the sled. Below you is a fork in the path with snow built up in the middle, and you can tell that, if you don't

⁵ MTurk is a website supported by Amazon.com (<https://requester.mturk.com/mturk/welcome>) that provides users the opportunity to fill out surveys for modest compensation. Recent work indicates that the data gathered through MTurk is at least as reliable as that gathered through standard psychology pools composed of undergraduates (see Buhrmester et al forthcoming).

direct your sled one way or the other, the contours of the mountain will channel you and your sled either to the left or to the right.

Ability Question

Consider how things seem to you as you approach the fork in the path. In particular, consider what it's like to decide which way the sled will go.

Please indicate your level of agreement with the following statement:

When deciding which way the sled will go, it feels like I can either go to the left or go to the right.

Participants were asked to indicate their agreement with this statement on a 7-point scale (1=disagree completely; 7=agree completely).

For condition 2, the vignette was the same except that participants were asked to imagine that the sledding episode occurred many years ago. And instead of getting a response regarding a phenomenology of the ability to do otherwise, we asked them to indicate agreement with a statement about a past ability to have done otherwise:

“I could have gone right instead of left.”

Part 2: Training on determinism

We wanted to focus on participants who had a phenomenology of the ability to do otherwise, so only participants who indicated a positive level of agreement to the first questions (5 or higher) were directed to the training section. Here, participants were given a detailed explanation of causal completeness, summed up as follows: “According to causal completeness, everything that happens is fully caused by what happened before it. This is true from the very beginning of the universe, so what happened in the beginning of the universe fully caused what happened next, and so on right up until the present. Causal completeness holds that everything is fully caused in this way, including people’s decisions.”

Participants were then given two kinds of cases to illustrate the phenomenon. In one case, they were told that an earthquake fully caused the volcanic eruption at Mt. St. Helens⁶, and they were then told, “According to causal completeness, if we could somehow replay the entire past right up until St. Helens erupted on May 18, 1980, then St. Helens would once again erupt at that time. Another way to put this is to say that all the events leading up to the eruption made it so that the eruption had to happen.” In another case past events, feelings and beliefs led to Obama’s decision to pick Joe Biden as his running mate, and participants were told “According to causal completeness, if we replayed the past right up until Obama’s decision—including everything that was going through Obama’s mind—then Obama would once again make exactly the same decision. That is, all the events leading up to Obama’s decision

⁶ This is an oversimplification of the geological facts which we adopted to ease the load on participants.

(including everything that was going through Obama’s mind), made it so that it had to happen that Obama would pick Biden.”⁷

We then tested comprehension of causal completeness. We first asked participants to indicate whether the following was true or false:

According to causal completeness, St. Helens would have erupted on May 18, 1980 even if there had been no earthquake.

Participants who answered “True” (the incorrect answer) were corrected, and given an explanation of the right answer. These participants were then given a similar question to see if they had absorbed the training. If they answered incorrectly yet again, they did not move on to answer the compatibility question, as they were deemed to have insufficient comprehension of causal completeness.

Participants who passed this first kind of question either on the first or second try were given another true/false question to test for comprehension:

⁷ In defining determinism—our causal completeness (CC)—as meaning “everything that happens is fully caused by what happened before it”, some might think this consistent with certain indeterminist conceptions of causation. But to say that events are *fully* caused is meant to avoid this reading—being fully caused suggests that nothing extra-causal is needed to help settle events. Our examples aid the preferred interpretation. We say, e.g., that under CC, “it had to happen that Obama would pick Biden”.

Draft, February 6, 2012

According to causal completeness, if a week from now Barack Obama decides to have soda with dinner, all the events leading up to that decision will make it the case that he has to decide to have a soda with dinner.

Our objective here was to test for and correct overly weak interpretations of causal completeness. Those who answered “False” (the incorrect answer) were corrected and given another chance at a similar question. If they answered incorrectly yet again they failed the training and did not answer the compatibility question. Participants who passed both kinds of questions either on the first or second try were deemed to have adequate comprehension of determinism, and these participants moved on to the third part of the study, the compatibility question.

Part 3: Consistency

After successful completion of the training, in both conditions participants were told to recall their agreement with the statement regarding the ability to do otherwise (from Part 1 of the survey). E.g., in condition 1, they were told:

Now, recall that you previously agreed with the following statement:

When deciding which way the sled will go, it feels like I can either go to the left or go to the right.

Following this, they were asked the compatibility question. In condition 1, this read as follows:

Compatibility Question

Considering this previous statement and your understanding of causal completeness, please indicate your level of agreement with the following:

Even though it felt like I could either go to the left or go to the right, if causal completeness is true there is something mistaken about how that decision felt to me.

In condition 2, the compatibility statement was:

Even though I said I could have gone right instead of left, if causal completeness is true there is something mistaken about what I said.

Agreement was indicated on the same 7-point Likert scale as was used for the Ability Question, and an answer above 4 was taken to be an incompatibilist answer.

Results:

Of the 34 participants who started condition 1, 33 completed it. Of these, 31 indicated a phenomenology of an ability to choose among possibilities and all but 4 of them passed the

training section.⁸ The remaining 27 participants gave a mean response of 4.93 on the compatibility question, which differed significantly from the midpoint of the scale, $t(26) = 2.65$, $p = .014$. That is, participants tended to interpret their agentic experience as being incompatible with determinism.

In condition 2, of the 50 participants who started the survey, 49 completed it. Of these, 47 indicated an ability to choose among possibilities and all but 2 of them passed the training section.⁹ The remaining 45 participants gave a mean response of 5.24 on the compatibility question, which differed significantly from the midpoint of the scale, $t(44) = 5.05$, $p < .001$.

A t -test comparing conditions 1 and 2 showed no significant difference, $p = .448$. So participants tended to be just as incompatibilist about retrospective judgments of their ability to do otherwise as they are about their current experience as of being able to do otherwise. This first study provides evidence that people do indeed judge that their experience of deciding is inconsistent with determinism, in the sense that the experience is somehow mistaken or non-veridical if determinism is true. It also suggests that the effect is robust across retrospective and present-focused cases.

⁸ Nine participants required correction, and ultimately passed the training section. The responses of those who required correction did not differ from those who answered correctly without training ($p > .2$).

⁹ Seven participants required correction, and ultimately passed the training section. Again, there were no differences between those who required correction and those who didn't.

3 Study 2

3.1 Overview

One major limitation of study 1 is that it involved merely imagined choices. This inserts a distance between the actual phenomenology of decision-making and judgments about that phenomenology. As a result, in study 2 we introduce conditions in which agents actually make decisions. In addition, study 1 focused on decisions that have no moral weight. We thus added a condition in study 2 in which the decision *does* have a moral element. So this study comprises three conditions to test for any effect from actual choices or from morally salient choices. We also introduced two innovations to the study's design.

First, we wanted to make our vignette more “choicely”. It struck us that in situations such as sledding down a hill often we don't have a salient experience as of deciding which way to go. We just go one way or the other. Second, we wanted to address a potential worry about our use of the word “mistaken” in the compatibility question. For example, in condition 1 from our first study, we asked:

Even though it felt like I could either go to the left or go to the right, if causal completeness is true there is something mistaken about how that decision felt to me.
(Emphasis added.)

One worry about this wording was that participants might misinterpret it as asking whether they were mistaken in thinking that their experiences felt a certain way, rather than as asking

whether there would be something mistaken about the content of their felt experiences.¹⁰ Our solution was to replace the above wording with a wording of the following form:

Even though it felt like I could either choose to X or choose to Y, if causal completeness is true then I couldn't really have chosen differently than I did.

With these modifications, condition 1 presented participants with an *imagined* choice among two very similar charities, condition 2 presented participants with an *actual* choice among two similar charities, and condition 3 presented participants with an *actual morally salient* choice among two charities, one for endangered trees, another for children's cancer treatments.

Method:

Participants:

155 participants were initially recruited online through the Mechanical Turk (MTurk) website. The survey itself was conducted using SurveyMonkey. 21 participants did not complete the survey, or indicated that they had recently taken a "very similar" survey.¹¹ They were excluded from the analysis.

¹⁰ Thanks to Lucas Thorpe for this objection.

¹¹ Studies 2 and 3 were run after study 1, and some of the conditions in studies 2 and 3 were run serially, so we excluded participants who indicated they had taken a very similar survey to minimize the influence of having previously taken one of our surveys.

Materials:

As in study 1, this study had three parts.

Part 1: The ability to do otherwise

For condition 1 of this study, we asked participants to imagine deciding between two charities for endangered trees.

Imagine that you have \$0.50 to donate. You have two options:

Donate to a foundation that protects the endangered tree *Castanea Dentata*.

OR

Donate to a foundation that protects the endangered tree *Ulmus Dentata*.¹²

These are your only two options.

Condition 2 was similar except that participants were given an *actual* choice. Participants were told that they had \$0.50 to donate to one of the two tree charities. We informed participants (truly) that we would actually donate this money to whichever charity they chose. Participants read:

¹² *Castanea Dentata* and *Ulmus Dentata* are the names of the American Chestnut and the American Elm, respectively. They are endangered species in North America. The charities we used were The American Chestnut Foundation (<http://www.acf.org/>), and Trees Winnipeg: Coalition to Save the Elms (<http://www.savetheelms.mb.ca/>).

You have \$0.50 to donate. We, the researchers, will actually donate this money for you whichever way you decide.

Participants were then presented with the same option language as in the imagined condition, and each option appeared as a radio button at the bottom of the page.

Finally, in condition 3 we presented participants with a morally salient choice between a foundation that protects the tree *Castanea Dentata* or (2) and The Childhood Cancer Foundation.¹³, on the assumption that people tend to think that saving dying children has greater moral weight than saving endangered trees.

In all conditions, after being given the imagined or actual choice, participants were asked a question about the ability to do otherwise. For instance, in condition 3, participants were asked to indicate their level of agreement (on a 7-point scale) with the following statement:

When deciding which option to choose, it feels like I can either choose to donate to the endangered tree *Castanea Dentata* or choose to donate to the Childhood Cancer Foundation.¹⁴

¹³ <http://www.candlelighters.ca/>

¹⁴ We want to forestall a potential concern about the phrasing here. Suppose I am determined to choose p. It follows that I can choose p. And you might think it further follows that I can choose p or choose q, for this follows from the simple logical principle of disjunction introduction. In that case the ability to choose p or choose q is clearly *compatible* with

(In conditions 2 and 3, participants were subsequently required to make a choice between the charities.) As in study 1, only participants who agreed with the ability-to-do-otherwise statement proceeded to the training.

Part 2: Training on determinism

The training section was the same as that used in study 1, and once again only those who passed the training proceeded to the compatibility question.

Part 3: Consistency

The compatibility question was adapted for the new cases. For example, in conditions 1 and 2, participants were asked to indicate agreement (on a 7-point scale) with this statement:

Even though it felt like I could either choose to donate to *Castanea Dentata* or choose to donate to *Ulmus Dentata*, if causal completeness is true then I couldn't really have chosen differently than I did.

determinism. However, participants report an incompatibilist phenomenology as of an ability to choose p or q, which suggests that they are not reading “can choose p or can choose q” in this compatibilist way.

Results:

Of the 50 participants who started condition 1, 42 completed it and had not recently taken a very similar survey (3 had). Of these, 38 indicated a phenomenology of an ability to choose among possibilities and all but 3 of them passed the training section.¹⁵ The remaining 35 participants gave a mean response of 5.60 on the compatibility question, which differed significantly from the midpoint of the scale, $t(34) = 6.08, p < .001$. The results of an imagined choice are consistent with the results of condition 1, study 1, if not stronger by virtue of the more “choiciness” vignette.

In condition 2, of the 48 participants who started the survey, 42 completed it and had not recently taken a very similar survey (4 had). Of these, 39 indicated a phenomenology of an ability to choose among possibilities and all but 2 of them passed the training section.¹⁶ The remaining 37 participants gave a mean response of 5.78 on the compatibility question, which differed significantly from the midpoint of the scale, $t(36) = 6.85, p < .001$. That is, participants were again incompatibilist about the phenomenology, this time of an actual choice.

¹⁵ Twelve participants required correction and successfully passed the training section. There was a significant difference in responses between those who required correction and those who didn't. Those who required correction reported that their phenomenology was incompatible with causal completeness ($M = 4.82$) but to a lesser degree than those who answered these questions correctly the first time ($M = 5.96$), $p = .038$.

¹⁶ Six participants required correction and successfully passed the training section. There was no difference between the responses of those who required correction and those who didn't ($p > .2$).

In condition 3, of the 57 participants who started the survey, 50 completed it and had not recently taken a very similar survey (3 had). Of these, 43 indicated a phenomenology of an ability to choose among possibilities and all but 3 of them passed the training section.¹⁷ Most of the remaining 40 participants (90%) opted to donate to the Childhood Cancer Foundation, as we expected on the the assumption that the cancer charity would be regarded as more morally salient. The 40 participants gave a mean response of 5.85 on the compatibility question, which differed significantly from the midpoint of the scale, $t(39) = 7.66, p < .001$. Once again we find incompatibilist phenomenology—this time with a morally salient choice.

ANOVA testing showed no overall effect of condition among conditions 1, 2, and 3, $F(2, 111) = .254, p = .776$. So there appears to be no effect produced by making the condition an actual choice, or by making the choice morally salient.

The results of study 2 show that people report incompatibilist phenomenology of agency for actual choices. Indeed, whether the decision is set up as an imagined one or an actual one does not affect the degree to which participants interpret their agentic experience as being incompatible with determinism. The results also show that whether or not the decision is morally salient doesn't affect the degree to which participants interpret their agentic experience as being incompatible with determinism. So the results of previous studies seem to extend to the moral domain, where issues of responsibility loom large.

¹⁷ Six participants required correction and successfully passed the training section. Again, we found no difference between the responses of those who required correction and those who didn't ($p > .2$).

4 Study 3

4.1 Overview

One possible concern with our previous studies stems from the way we have phrased the key compatibility question. Notice our use of an “even though” locution in the following:

Even though it felt like I could either choose to donate to *Castanea Dentata* or choose to donate to *Ulmus Dentata*, if causal completeness is true then I couldn't really have chosen differently than I did. (Emphasis added.)

Although we took this to be a natural phrasing of the question, one might think that “even though” primes the participant to agree with the statement, which in this case is an incompatibilist response. Our final study drops this potentially troublesome phrase and also tests whether the phenomenology of epistemic uncertainty differs from the phenomenology of being able to do otherwise in terms of compatibility with determinism. In condition 1, we once again presented participants with an actual choice among two options and tested whether they would continue to report having an incompatibilist phenomenology as of being able to do otherwise. In condition 2, we focused on epistemic phenomenology.

Method:

Participants:

106 participants were initially recruited online through the Mechanical Turk (MTurk) website. The survey itself was conducted using SurveyMonkey. 15 participants did not complete the survey, or indicated that they had recently taken a “very similar” survey. They were excluded from the analysis.

Materials:

The vignette and first question for condition 1 read as follows.

Part 1: The ability to do otherwise

In both conditions, participants were told that they would have a chance to win 5 cents if they picked the right button. The text went as follows:

At the bottom of this page, there are two buttons, labeled H and V. Each option is currently available for you to choose. In a moment, we’ll ask you to choose just one of them. For this survey, only one of the buttons will give you an extra \$0.05 (as bonus payment on MTurk) if you choose it. But we won’t tell you which button it is—you’ll have to make a choice and find out.

But don’t decide just yet.

First, consider how things seem to you as you face your decision. In particular, consider what it's like to decide which option to choose.

In condition 1, participants were asked to indicate agreement (on a 1-7 scale) with the following statement:

When deciding which option to choose, it feels like I can either choose H or choose V.

Condition 2 was the same except that we dropped the modal “can” and asked participants to “consider what it's like to wonder which option you'll choose”. Participants were then asked to indicate their level of agreement with a statement about *epistemic* phenomenology:

When wondering which option I'll choose, it feels like I don't know for sure before I select a button which button is the bonus button.

As in study 1, only participants who agreed with the ability-to-do-otherwise statement proceeded to the training.

The two available options—H and V—appeared at the bottom of the screen, with a radio button representing each option. Participants were not told whether they had chosen the bonus button (H) until after they had answered the compatibility question.

Part 2: Training on determinism

The training section was the same as that used in study 1, and again participants only proceeded to the compatibility question if they passed the training.

Part 3: Consistency

The compatibility question was adjusted for the new cases. In condition 1, participants were told:

Now, recall the button-choosing situation. You previously agreed with the following statement:

When deciding which option to choose, it feels like I can either choose H or choose V.

Considering this previous statement about how things felt to you before your choice and your understanding of causal completeness, please indicate your level of agreement with the following:

If causal completeness is true, then I couldn't really have chosen differently than I did.

In condition 2, participants were reminded that they agreed with this statement:

When wondering which option I'll choose, it feels like I don't know for sure before I select a button which button is the bonus button.

They were then asked to indicate their level of agreement with the following:

If causal completeness is true, then I knew for sure before I selected a button which button was the bonus button.

Our aim was to test whether participants distinguish the sort of alternative possibilities they reported themselves as experiencing in other conditions from clearly compatibilist alternative possibilities, which have to do simply with our ignorance of the future.

Results:

Of the 53 participants who started condition 1, 47 completed it and had not recently taken a very similar survey (2 had). Of these, 44 indicated a phenomenology as of there being alternative possibilities in the situation and all but 3 of them passed the training section.¹⁸ The remaining 41 participants gave a mean response of 5.34 on the compatibility question, which differed significantly from the midpoint of the scale, $t(40) = 4.54, p < .001$. That is, participants

¹⁸ Eight participants required correction, and passed the training section. There was a significant difference in responses between those who required correction and those who didn't. Those who required correction reported that their phenomenology was incompatible with CC ($M = 4.25$) but to a lesser degree than those who answered these questions correctly the first time ($M = 5.60$), $p = .057$.

once again demonstrated a strong tendency to interpret their agentic experience as being incompatible with determinism.

In condition 2, of the 53 participants who started the survey, 44 completed it and had not recently taken a very similar survey (8 had). Of these, 39 indicated a phenomenology of an ability to choose among possibilities and all but 1 of them passed the training section.¹⁹ The remaining 38 participants gave a mean response of 2.66 on the compatibility question, which differed significantly from the midpoint of the scale, $t(37) = -5.23, p < .001$. That is, participants tended to regard their phenomenology of uncertainty about the future as compatible with determinism. A *t*-test between conditions 1 and 2 showed that results differed significantly between these two conditions, $t(76) = 6.85, p < .001$.

This final study provides yet further evidence that people do indeed judge that their experience of deciding is inconsistent with determinism, in the sense that the experience is non-veridical if determinism is true. At the same time, people tend to think that the feeling of not knowing what will happen is perfectly consistent with determinism. This suggests an appropriate sensitivity to the fact that ignorance is not incompatible with determinism.

¹⁹ Nine participants required correction, and passed the training section. There were no statistically significant differences in responses between those who required extra training and those who didn't ($p = .2$).

5 General Discussion

5.1 Incompatibilism

Our results have implications for several issues concerning free will. Perhaps most importantly, our studies seem to vindicate the incompatibilist descriptions of our experience as of being able to do otherwise suggested by Campbell, O'Connor, and Searle. By the same token, our results run counter to the compatibilist descriptions of our experience suggested by Mill, Grünbaum, and Nahmias and colleagues. The design of our studies left it open for participants to describe their experience as involving the ability to do otherwise, while allowing them to interpret this ability however they wished. The results indicate that the people in the population we tested tended to judge that their experience was *incompatible* with determinism.

The results also address a concern that has plagued recent work on intuitions about free will. Nahmias & Murray (2011) contend that people give incompatibilist responses in previous experiments simply because people misunderstand determinism. This is an important concern. But rather than merely testing to see whether people misunderstand determinism, we attacked the comprehension issue directly by exploiting the familiar technique of training to criterion. And we did not find any widespread confusion of determinism and bypassing. Part 1 of our training controls for confusion between determinism and fatalism. And the majority of our participants reported that the accuracy of their experience as of being able to do otherwise is inconsistent with *determinism*, correctly understood. Across all our studies, the percentage of participants who didn't make it to the compatibility question due to failing the training section was small, at 6.15%. When we look at those participants who answered part 1 of the training

incorrectly—that is, at those who *did* initially confuse determinism with fatalism, and who were directed to the follow-up training question—the percentage was small compared with Nahmias & Murray’s results: Only 20.68% of participants initially made this mistake. Of those who initially made the mistake, 85.71% answered the follow-up training question correctly. Thus, fewer than 3% of participants continued to confuse determinism and fatalism after training. And those who required correction did not respond in any significant way differently from those who didn’t.²⁰

5.2. The ability to do otherwise

Much of the free-will debate, since at least Hobbes, has been about an *ability to do otherwise*. One influential compatibilist thought is that the notion of the ability to do otherwise should be understood in contrast to constraint or coercion. The idea is that an agent is able to do otherwise just in case, if she had chosen, or wanted, or tried to do otherwise, then she would have done so (cf. Moore 1912). There are also recent versions of such a “conditional analysis” of the ability to do otherwise. According to Kadri Vihvelin (2004), for instance, an ability to act (or not to act, which is simply to be able to act in another way) is analyzable along something like the following lines: an agent can Φ at t_1 (say, raise her hand at t_1) just in case were she to choose to Φ at t_2 , and her body stayed working normally and nothing interfered with her, she would Φ at t_2 .²¹ In other words, Vihvelin holds that “persons have abilities by *having intrinsic properties that are the causal basis of the ability*” (2004: 438). So Vihvelin thinks that an

²⁰Again, with the exception of study 2, condition 1, and study 3, condition 1. (See footnotes 15 and 18.)

ability to act is a disposition, or a bundle of dispositions. And, as she points out, “no one denies that dispositions are compatible with determinism” (2004: 429). After all, even if determinism is true, glass is still fragile—i.e., it has the disposition to break if struck.²²

Other compatibilists embrace an epistemic reading of “can do otherwise.” On this view, to maintain that I can go left or right is simply to note that it is epistemically open whether I will go left or right. J. J. C. Smart argues that this is a natural way to interpret the expression “could have done otherwise” even outside the sphere of action. When I say, “the plate fell, and it could have broken”, I am not, says Smart, committing myself to any claim about determinism. Rather, what I am saying is that, before the plate completed its fall, for all I knew, the plate would break (1961: 298). Similarly, perhaps when I say that Oswald could have done otherwise, all I’m saying is that, before Oswald pulled the trigger, for all anyone knew, he wouldn’t pull the trigger. If I’m merely making a claim about epistemic possibilities, then there is no conflict with determinism.

²¹ Vihvelin’s exact formulation is as follows: “*S* has the ability at time *t* to do *X* iff, for some intrinsic property or set of properties *B* that *S* has at *t*, for some time *t*’ after *t*, if *S* chose (decided, intended, or tried) at *t* to do *X*, and *S* were to retain *B* until *t*’, *S*’s choosing (deciding, intending, or trying) to do *X* and *S*’s having of *B* would jointly be an *S*-complete cause of *S*’s doing *X*” (2004: 438).

²² For similar accounts, see Smith (2004) and Fara (2008). Questions persist (see e.g., Clarke 2009) about whether any “dispositionalist” account is an adequate analysis of the ability to act, and thus of the ability to act otherwise.

By contrast, incompatibilists think being able to do otherwise (in the relevant contexts) means being able to do something other than what one does, all prior conditions (including one's desires) remaining the same. This ability is presumed to be a matter of fact, not something about our epistemic access to facts.

At least insofar as the relevant notion of the ability to do otherwise is reflected in the experience as of being able to do otherwise, our results suggest that the compatibilist accounts fail. Across three studies, participants tended to interpret their agentic experience in terms of an ability to do otherwise, and they interpreted that ability incompatibilistically. Concerning the traditional compatibilist analysis, our results equally undercut old and new versions. After all, we allowed participants to describe their experiences as involving the ability to do otherwise or not, where they were free to interpret this ability however they wished. Participants then judged that *this* ability—the one they had been allowed to interpret however they wished—was incompatible with determinism. The epistemic compatibilist account is also undermined by these results. Participants gave compatibilist judgments about the case of ignorance about the future (study 3, condition 2), indicating that they do have an appreciation that the feeling of uncertainty is consistent with determinism.

We conclude that the notion of “can do otherwise”, at least with respect to one's decisions, is naturally interpreted in ways that contravene the most familiar compatibilist approaches in the philosophical literature. When participants attend to their experience while they consider future events, their usage of “can” tends to reflect a sense of *metaphysical* openness that is incompatible with determinism.

5.3. Misinterpreting one's agentic experience

Obviously, the fact that people interpret their agentic experience as incompatibilist doesn't show that people actually have an incompatibilist ability to do otherwise. Terry Horgan argues that people might be mistaken in their interpretation of their own phenomenology. He allows that people might regard their agentic experience as incompatibilist:

When one attends introspectively to one's agentic phenomenology, with its... [representational]... aspects of freedom... and when one simultaneously asks reflectively whether the veridicality of this phenomenology is compatible with causal determinism... , one feels *some* tendency to judge that the answer to such compatibility questions is No. (In press)²³

But Horgan notes that we must distinguish between the content of our experience and the content of judgments. The former kind of content Horgan dubs "presentational content", and it ... is the kind that accrues to phenomenology directly—apart from whether or not one has the capacity to articulate this content linguistically and understand what one is thus articulating, and apart from whether or not one has the kind of sophisticated conceptual repertoire that would be required to understand such an articulation. (Horgan, in press)

By contrast, "judgmental content" is the kind of content associated with linguistic articulations. Of course, we make judgments about our phenomenology, and so we can have judgmental content that aims to capture our presentational content. The key point here is that it is possible for our judgments about the (presentational) content of our experience to go awry.

²³ For Horgan, this representational "aspect of freedom" is what we have been calling the experience as of being able to do otherwise.

That said, those judgments are at least *prima facie* evidence of the nature of the presentational-cum-phenomenal content, we maintain, so we would need some positive reason to think that participants have systematically misinterpreted the nature of their phenomenology. Further, even if we grant *arguendo* that the presentational content of agitive experience is (in the first instance) compatible with determinism, and that reports to the contrary count as mistaken interpretations, still, the fact that people judge the experience incompatibilist would be significant. For one thing, when considering how best to understand the notion of the “ability to do otherwise”, in many cases what will be of primary importance is how people *think* about their ability to do otherwise, and that is clearly judgmental. Second and more interestingly, judgmental content can feed back into presentational content. It is well known that what one judges about a situation can affect one’s perception of the situation. Horgan recognizes this, and he notes that the distinction between presentational and judgmental content isn’t always sharp: “it may very well be that the two kinds of content can interpenetrate to a substantial extent” (in press). As a result, even if the presentational content of agitive experience is, in the first instance, compatibilist, that doesn’t mean that the presentational content *remains* compatibilist. It might be that the incompatibilist judgment shapes the presentational content.²⁴

²⁴ Note that Horgan’s notion of “presentational content” is not simply “raw feels” with no propositional content. For everyone would concede that insofar as we have incompatibilist phenomenology, it must be presented at a level with greater conceptual sophistication than is provided by raw feels. Horgan is explicit about the possibility of rich conceptual resources being implicated in presentational content: “It is plausible... that humans can have

5.4 Deliberation compatibilism

A final issue that might be illuminated by our results is the debate over the presuppositions of deliberation. Some philosophers have maintained that deliberation carries with it a presumption of genuinely open possibilities of an incompatibilist variety. Richard Taylor writes, “I cannot deliberate about what to do, even though I may not know what I am going to do, unless I believe that it is up to me what I am going to do” (1983: 38–39). And this “up to me” is incompatible with determinism. Van Inwagen makes a similar point: “[I]f someone deliberates about whether to do A or to do B, it follows that his behavior manifests a belief that it is *possible* for him to do A—that he *can* do A, that he has it within his power to do A—and a belief that it is possible for him to do B” (1983: 155).

On the other side, we find “deliberation compatibilists”, who maintain that deliberation contains no such presuppositions. Tomis Kapitan begins his paper (which would become the locus classicus for deliberation compatibilism) thus:

By *deliberation* we understand practical reasoning with an end in view of choosing some course of action. Integral to it is the agent’s sense of alternative possibilities, that is, of two or more courses of action he presumes are *open* for him to undertake or not. (1986: 230)

presentational contents the possession of which require (at least causally) a fairly rich repertoire of background concepts that can figure in judgmental states.” For instance, “One can have presentational experiences, for instance, as-of computers, automobiles, airplanes, train stations” (Horgan in press).

Kapitan goes on to argue that the presumption of openness does not require *metaphysical* openness, but only *epistemic* openness.²⁵ A number of philosophers have followed Kapitan in developing compatibilist accounts of the presuppositions behind deliberation (e.g., Nelkin 2004, Pereboom 2008).

Insofar as deliberation compatibilism claims that deliberation is not *as a matter of fact* experienced as having incompatibilist presuppositions, our studies indicate this position is mistaken. This does not decide the dispute concerning deliberation compatibilism, but it does show that we need to distinguish three versions of deliberation-compatibilism:

- (1) people's beliefs about their current deliberations are compatible with determinism;
- (2) people's beliefs about their current deliberations are not compatible with determinism, but they can be adjusted to be compatible;
- (3) people's beliefs about their current deliberations are not, and cannot be adjusted to be, compatible with determinism, but we can conceive of a rational being whose beliefs about deliberation are compatible with determinism.

Our results suggest that the first version of deliberation compatibilism is false. People's beliefs about their deliberations are incompatibilist. The second version—that our actual experiences

²⁵ According to Kapitan and other deliberation compatibilists, there are other conditions, too. In particular, Kapitan maintains that deliberation carries a presupposition of *efficacy*, which he characterizes roughly as follows: “an agent presumes that his Φ -ing is an open alternative for him *only if* he presumes that he would Φ if and only if he were to choose to Φ ” (1986: 234). See also Pereboom (2008: 288). We leave this complication aside since it doesn't affect our point.

are incompatibilist but revisable—is an interesting possibility, but it remains an open question whether it *is* possible to revise this aspect of our experience. Until we know more about what generates the incompatibilist experience, it is hard to know whether it can be modified. One possibility is that the incompatibilist experience is generated in a way that is not cognitively penetrable (see e.g., Bayne 2011). That is, it might be that even if we form the explicit high-level belief that deliberation is theoretically compatible with determinism, this will not eradicate our experience of our deliberation as incompatibilist. The third version of deliberation compatibilism—that we can conceive of rational creatures who deliberate as determinists—is not under any threat from our results. But if it turns out to be impossible for us to *be* such rational animals, that might undercut some of the interest of deliberation compatibilism.

6 Conclusion

The experience as of *being able to do otherwise* has long been central to debates about agency and free will. Libertarians appeal to this experience as evidence that determinism is false; compatibilists reject the libertarian accounts of the character of the experience. Despite the pivotal role of experience in these arguments, the experience itself has received scant attention. Our studies are an attempt to advance the issue. We find consistently incompatibilist judgments about the nature of the experience as of being able to do otherwise. This lends support to the phenomenological claim of libertarians, though we ourselves are not inclined to take the phenomenology of indeterminism as evidence that agency isn't determined. Our results also

Draft, February 6, 2012

suggest that an incompatibilist interpretation of the notion of “ability to do otherwise” is the best interpretation of that notion, at least insofar as that notion is supposed to reflect our experience as of being able to do otherwise. Finally, our results also speak to the presuppositions of deliberation. What our studies indicate is that *as a matter of fact* our experience of deliberation features metaphysical openness (that is inconsistent with determinism). While this does not decide the dispute between deliberation compatibilists and deliberation incompatibilists, it does make salient the possibility that deliberation compatibilism requires an account of deliberation that is explicitly revisionist with respect to our actual experience of deliberation.

Works Cited:

Bayne, Tim. 2011. “The Sense of Agency”. In F. Macpherson (ed.) *The Senses*. Oxford: Oxford University Press. 355–74.

Berofsky, Bernard. 2002. “Ifs, Cans, and Free Will: The Issues”. In R. Kane (ed.) *The Oxford Handbook of Free Will*. New York: Oxford University Press. 181–201.

Buhrmester, M., T. Kwang, and S. Gosling. Forthcoming “Amazon’s Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data?” *Psychological Science*.

Draft, February 6, 2012

Campbell, C. A. 1951. "Is 'freewill' a pseudo-problem?" *Mind* 60/240: 441–65.

Campbell, Joseph. 2005. "Compatibilist Alternatives", *Canadian Journal of Philosophy* 35: 387–406.

Clarke, Randolph. 2009. "Dispositions, Abilities to Act, and Free Will: The New Dispositionalism", *Mind* 118: 323–51

Fara, Michael. 2008. "Masked Abilities and Compatibilism", *Mind* 117/468: 843–65.

Fischer, John M. 1994. *The Metaphysics of Free Will*. Oxford: Blackwell Publishing.

Fischer, John M. 2008. "Freedom, Foreknowledge, and Frankfurt: A Reply to Vihvelin", *Canadian Journal of Philosophy* 38/3: 327–42.

Frankfurt, Harry. 2003/1969. "Alternate possibilities and moral responsibility. In G. Watson (ed.) *Free Will*. Oxford: Oxford University Press. 167–176.

Grünbaum, Adolf. 1952. "Causality and the Science of Human Behavior," *American Scientist* 40(4): 665–676

Draft, February 6, 2012

Horgan, Terry. In press. "Causal Compatibilism about Agentive Phenomenology". For a festschrift for J. Kim, edited by M. Sabates, D. Sosa, and T. Horgan.

Hume, David. 1960/1739. *A Treatise on Human Nature*, L. A. Selby-Bigge (ed.). Oxford: Clarendon Press.

Kapitan, Tomis. 1986. "Deliberation and the presumption of open alternatives", *The Philosophical Quarterly* 36/143: 230–251.

Lehrer, Keith. 1960. "Can we know that we have free will by introspection?", *The Journal of Philosophy* 57/5: 145–57.

Mill, John Stuart 1865. *An Examination of William Hamilton's Philosophy*. Boston: William V. Spencer.

Moore, G. E. 1912. "Free Will". From *Ethics*. Oxford: Oxford University Press. Chapter 6, 122–37.

Nahmias, Eddy, Stephen Morris, Thomas Nadelhoffer & Jason Turner. 2004. "The Phenomenology of Free Will", *Journal of Consciousness Studies* 11/7–8: 162–79.

Draft, February 6, 2012

Nahmias, Eddy & Dylan Murray. 2011. "Experimental Philosophy on Free Will: An Error Theory for Incompatibilist Intuitions". In *New Waves in Philosophy of Action*, Aguilar, Jesús, Andrei Buckareff and Keith Frankish (eds.). Palgrave Macmillan. 189-216.

Nelkin, Dana. 2004. "Deliberative alternatives", *Philosophical Topics* 32: 215–240.

O'Connor, Timothy. 1995. "Agent causation". In T. O'Connor (ed.), *Agents, Causes, and Events: Essays on Indeterminism and Free Will*. New York: Oxford University Press. 173–200.

Pereboom, Derk. 2008. "A Compatibilist Account of the Epistemic Conditions on Rational Deliberation", *The Journal of Ethics* 12/3: 287–306.

Perry, John. 2004. "Compatibilist Options". In J. Campbell, M. O'Rourke, and D. Shier (eds.) *Freedom and Determinism*. MIT Press. 231–254.

Searle, John. 1984. *Minds, Brains, and Science*. Cambridge, MA: Harvard University Press.

Smart, J. J. C. 1961. "Free Will, Praise and Blame." *Mind* 70: 291–306.

Smith, Michael. 2004. "Rational Capacities", in *Ethics and the A Priori: Selected Essays on Moral Psychology and Meta-Ethics* by M. Smith. New York: Cambridge University Press. 114–135.

Draft, February 6, 2012

Taylor, Richard. 1983. *Metaphysics*, 3rd edition. Englewood Cliffs, N.J.: Prentice Hall.

van Inwagen, Peter. 1983. *An Essay on Free Will*. Oxford: Oxford University Press.

Vihvelin, Kadri. 2004. "Free Will Demystified: A Dispositional Account", *Philosophical Topics* 32: 427–450.