

# A Compatibilist Account of the Beliefs Required for Rational Deliberation<sup>1</sup>

*Journal of Ethics* 12, 2008, pp. 287-307

Derk Pereboom, Cornell University

Penultimate version

## 1. Deliberation and the belief in openness.

Whenever we deliberate about what to do, we at least typically believe that we have more than one distinct option for which action to perform, each of which is available to us in the sense that we can perform each of these actions. That is, when we deliberate, we believe in the “openness” of more than one distinct option for what to do. It is often argued that belief in this sort of openness is actually required for deliberation, or at least for rational deliberation (Kant 1795/1981, Ak IV 448; Taylor 1966, chapter 12; Ginet 1966; van Inwagen 1983, 153-61).<sup>2</sup>

But according to some, such beliefs in openness would conflict with the belief that

---

<sup>1</sup> Thanks to Dana Nelkin, Louis deRosset, Joseph Campbell, Randolph Clarke, David Christensen, and Seth Shabo for comments and discussion. This paper also benefited from comments and discussion on the Second On-line Philosophy Conference, May 2007 – OPC2. Research on this article was facilitated by a generous Visiting Fellowship in the Centre for Consciousness of the Research School of Social Sciences at the Australian National University, October – December 2005.

<sup>2</sup> Peter van Inwagen writes: “one cannot deliberate about whether to perform a certain act unless one believes it is possible for one to perform it (1983, 154); and “if 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) For discussions of this type of view, see Kapitan 1986; Searle 2001; Nelkin 2004a, 2004b; Coffman and Warfield 2005).

determinism is true for the reason that, in any deliberative situation, the truth of determinism would evidently rule out the availability to us of all but one distinct option for what to do, and thus would evidently rule out openness about what to do. So then the content of a belief required for rational deliberation would be inconsistent with an evident consequence of determinism. Then an agent who believed that determinism was true, and had the belief required for rational deliberation, would have inconsistent beliefs. This line of reasoning supports an incompatibilist position about the relation between the belief that determinism is true and beliefs required for rational deliberation (Taylor 1966, Ginet 1966).

Here is a belief-in-openness requirement on rational deliberation, proposed by Dana Nelkin, which would secure this deliberation-incompatibilist result:

(I) Rational deliberators who deliberate about an action A must believe, in virtue of their nature as rational deliberators, that there exist no conditions that render either A or not-A inevitable. (Nelkin 2004b, 217),

An agent who rationally deliberates about an action A would then believe that there exist no conditions that render either A or not-A inevitable. But if she also believed in determinism and its evident consequences, she would believe that there do exist conditions that render either A or not-A inevitable. She would then have inconsistent beliefs. (Nelkin, as we shall see, goes on to challenge (I)).

Randolph Clarke agrees that when we deliberate about what to do, we typically presuppose that we have more than one distinct option for which action to perform, each of which is available to us in the sense that we can perform each of these actions, where can has

a metaphysical reading. But he also contends that deliberation need not have such a presupposition, and that there are alternative beliefs that are uncontroversially compatible with determinism that supply what is required for deliberation (2003, 112-7; 1992). I think that Clarke is right about this – and many others have agreed (e.g., Dennett 1984; Kapitan 1986). But existing proposals for the salient compatibilist belief have met with significant opposition. In what follows I contend that the preferable position has it that there are two such compatibilist beliefs, and that this dual-belief proposal meets what are in effect two distinct strands in the incompatibilist objections to such proposals. One of these is a belief in the *epistemic*, as opposed to metaphysical, openness for what to do, and the other is a belief in the efficacy of deliberation. Tomis Kapitan has argued for a requirement that, in effect, includes both these kinds of belief (1986). Here I amend his account by responding to concerns for his version of deliberation-compatibilism.<sup>3</sup>

A different sort of deliberation-compatibilism incorporates compatibilist-friendly metaphysical readings of ‘it is within S’s power to do each of A and B, or ‘S can do each of A or B,’ or ‘S has the ability to do each of A and B’ where A and B are distinct actions. One motivation for preferring the epistemic interpretation instead is the threat of the consequence argument against the metaphysical readings of these claims, given determinism. Another is that typical compatibilist metaphysical analyses of these notions are conditional on the

---

<sup>3</sup> I have defended incompatibilism about the relation between determinism and the sort of freedom required for moral responsibility, but compatibilism about many other features of agency and morality (Pereboom 2001). In particular, I have endorsed without elaboration Kapitan’s compatibilism about determinism and the beliefs required for deliberation (2001, 135-7). Here I present my more considered deliberation-compatibilist view.

following model:

S can do otherwise just in case if S had chosen otherwise, then S would have done otherwise.

and Roderick Chisholm has argued quite convincingly that such analyses are implausible (1964). In addition, as I shall contend later, openness would seem to be a categorical and not a conditional notion. The idea is that to rationally deliberate, I must believe that each of several options for what to do must in some sense be possible for me to secure, but not merely on the supposition that some condition is satisfied. A deliberation-compatibilism that incorporates epistemic readings of these claims derives plausibility from the fact that we are often not clear on whether in our thoughts 'possibility' and related notions are metaphysical or epistemic, and so in the beliefs that accompany deliberation, the epistemic readings are credible.

Let me here make two points by way of conceptual clarification. First, what notion of belief is at work here? I will assume the dispositional notion that Coffman and Warfield specify, which they characterize as 'S dispositionally believes p iff acceptance of p is accessible to S by way of a non-generative cognitive retrieval process.' (2005, 26) Another possibility is the weaker alternative of *rational commitment* that Nelkin suggests: "if the agent reflected on it, she would believe that her alternatives are open and her deliberation efficacious."<sup>4</sup> In the case of explication of rational commitment, it's open that there is no actual belief to *retrieve*; instead the belief might first come to be by way of agent's potential but non-actual reflection. Is mere rational commitment in this sense by contrast with dispositional belief, sufficient for

---

<sup>4</sup> On OPC2.

rational deliberation? Perhaps, it may be that our notion of rational deliberation is not sufficiently determinate to allow us to decide the issue.

Second, how should deliberation be characterized in this discussion? Coffman and Warfield (2005, 28), who cite Searle (2001, 14) and van Inwagen (2004, 217), argue that to join the issue with key participants in the debate such as van Inwagen and Searle, we should adopt the following characterization: “to deliberate is to try to choose what to do from among a number of incompatible courses of action under certain conditions.” Furthermore, Coffman and Warfield contend that Searle and van Inwagen conceive of deliberation as occurring “after reasons for various actions have been weighed and evaluated,” and this is how they propose deliberation be understood for the purposes of this debate (2005, 28). A concern about this proposal is that seems to leave too little to count as deliberation. Perhaps it leaves the forming of an all-things-considered judgment about what it is best to do from among distinct alternatives, and the forming of an intention to act from such alternatives. But these functions are not obviously aspects of deliberation, and even if they are, we ordinarily think of deliberation as involving more than this. What we in fact think of as central to deliberation is *figuring out what to do from among distinct alternatives by considering and evaluating reasons*. So suppose that we characterize deliberation as an active mental process whose aim is to figure out what to do from among a number of distinct, i.e., mutually incompatible alternatives, a process understood as one that can include the weighing and evaluating of reasons for the options for what to do?

Coffman and Warfield are concerned that the issue with van Inwagen and Searle will

not be joined on a characterization that involves more than they specify. But, at least initially, adding in the consideration and evaluation of reasons does not appear to have this upshot. The quest at hand is for necessary conditions for rational deliberation. If having the belief that I can do A and that I can do B is a necessary condition for my instantiating an active mental process whose aim is to choose whether to do A or to do B, understood as one that does not include the agent's weighing and evaluating of reasons for the options, it would seem that having this belief will also be a necessary condition for my instantiating the active process understood as one that can include weighing and evaluating of the reasons for the options -- and vice versa. More generally, I suspect that any belief-condition at issue in this debate would be necessary for both the broader and the narrower process, if it is necessary for one of them.

But also, the relevant quotation Coffman and Warfield (2205, 27) cite from van Inwagen is: "serious deliberation occurs when one is choosing between alternatives and it does not seem to one (once all the purely factual questions have been settled) that the reasons that favor either alternative are clearly the stronger." (2004, 217) This characterization of serious deliberation would at very least appear consistent with the following: all the purely factual questions have been settled, and it does not seem to the agent at this point that the reasons favor either alternative. But now genuinely serious consideration and evaluation of the reasons, and hence genuinely serious deliberation, can begin. The quotation from Searle is: "there is the gap of rational decision making, where you try to make up your mind what you are going to do. Here the gap is between the reasons for making up your mind, and the actual decision that you make" (2001, 14). It's at least consistent with this quote that the gap is

between the noting or apprehending of reasons and the actual decision, and what happens in the gap is serious consideration and evaluation of the reasons.

So perhaps on this proposed amendment on the characterization of deliberation, the issue is still joined with Searle and van Inwagen. And I think it has the virtue of characterizing deliberation in a more natural way.

## 2. The requirement of belief in epistemic openness.

One way in which deliberation-compatibilists have responded to the deliberation-incompatibilist on this issue is by claiming that required for deliberation is merely a belief in and epistemic kind of openness – for example, a belief that one has more than one option for what to do relative to what one believes, presumes, or knows. Kapitan, for example, contends that minimally rational deliberation requires a presumption of open alternatives, and

(PC) an agent presumes that his  $\phi$ -ing is an open alternative for him only if he presumes that if  $S$  is any set of his beliefs then his  $\phi$ -ing is contingent relative to  $S$ . (Kapitan 1986, 240)

Nelkin considers a similar thesis, with the change that the options be consistent with what the agent knows:

(K) Rational deliberators [about which action to perform] must believe, in virtue of their nature as rational deliberators, that they have multiple alternatives from which to choose, where those alternatives are each consistent with what they know. (Nelkin

2004b, 221)<sup>5</sup>

However, Nelkin argues that there do seem to be examples of rational deliberation that are counterexamples to K, and she cites the following case of Clarke's:

Imagine that Edna is trying to decide where to spend her vacation this year. She mentions this fact to her friend Ed, who, as it happens, is in possession of information that Edna does not yet have. Ed knows that Edna will soon learn that she can, with less expense than she had expected, visit her friend Eddy in Edinburgh. And given what Ed knows about Edna and her other options, he knows that after she learns of this opportunity, she will eventually decide to take it. However, Ed is a playful fellow, and he doesn't tell Edna all of this. He tells her only that he knows that she will eventually learn something that will persuade her to spend her vacation with Eddy in Edinburgh. [Edna] knows, let us suppose, that whenever Ed says anything of this sort, he is right. She believes then, with justification, that she will spend her vacation in Edinburgh.

(Clarke 1992; Nelkin 2004b, 221-2)

This seems indeed to be a counterexample to (K) (and, *mutatis mutandis*, to Kapitan's (PC)) considered as a requirement for rational deliberation; it also appears to be a counterexample to Carl Ginet's claim that "it is conceptually impossible for a person to know what a decision of his is going to be before he makes it" (1962, 50-1). I propose that this is because it seems possible that one can know a proposition without being certain that it is true, and against Kapitan's formulation, that one can believe a proposition without being certain that it is true

---

<sup>5</sup>Louis deRosset objects (in conversation) that one might rationally deliberate while suspending judgment about whether one has multiple alternatives – this seems right to me.



(Stanley, manuscript), and if one is not certain that one will do A, then (with a qualification to be discussed later) one can deliberate about whether to do it. However, if one is certain that one will do A, one might still seek the best reasons for doing A, but it seems to me that then one cannot deliberate about *whether to do it*. Given our understanding that the aim of deliberation is to determine what to do, nothing remains for deliberation to accomplish if one is already certain of what one will do.

The notion of certainty should not be the epistemic one, according to which S's being certain that P entails S knows that P. Rather, it can be understood as having a credence of 1.0, without this epistemic dimension. Now Kapitan raises an apt concern for the 'knows'-version of the belief-in-openness requirement, which, according to him favors the 'believes'-version: "I may, for instance, believe I will not fly to Copenhagen tomorrow and thus I do not deliberate about so doing, yet I may not know what I believe (perhaps some unforeseen emergency will call me to Copenhagen). The action is impossible relative to what I believe and so does not appear open to me, though it is contingent with respect to what I actually know." (1986, 239) What I mean by certainty is a credence of 1.0, and I don't have in mind an epistemic notion of certainty that builds in knowledge. Consequently, my claim is not undermined by the sort of concern Kapitan raises.

Clarke suggests for there to be a point to Edna's deliberating, "it is not necessary that she reopen the question of what she will do," or "that she suspend her belief that she will visit Eddy in Edinburgh" (1992, 108-9). He gives two reasons for this. One is that in deliberating she attempts to discover reasons for and against that alternative. But if that is the sole point of the

mental process, it won't count as deliberation by the characterization I've adopted. The other is that she hasn't yet formed the intention to go to Edinburgh, and that deliberation can produce this intention. Given that she knows she will go to Edinburgh, deliberation can produce an intention to go if she is not yet certain that she will. But first, given my definition of deliberation, further deliberation cannot have this effect if she is already certain that she will go. For then deliberation aims at determining what to do, and so if an agent is already certain of what she will do, then deliberation's aim has unequivocally been realized. Second, setting aside my definition, nothing that we ordinarily think of as deliberation would have the role of producing an intention to do A in a case in which the agent is already certain that she will do A. For example, consideration of reasons for and against in this situation would not have the role of producing an intention in such a case. Rather, at least typically, the agent would simply just form the intention, independently of further deliberation.<sup>6</sup>

---

<sup>6</sup> Nelkin points out (on OPC2) that there is an ambiguity in the technical use of the term 'deliberation' that trades on two senses of 'determine.' Here are two possible readings of the characterization I assume:

(D1) deliberation is an active mental process whose aim is to determine what to do, in the causal sense of producing a choice or intention for what to do from among a number of mutually incompatible alternatives.

(D2) deliberation is an active mental process whose aim is to determine what to do, in the practical epistemic sense of figuring out what to do, from among a number of mutually incompatible alternatives.

At this point, Brad Westlake (on OPC2) poses two alternatives: Edna could (a) go ahead and immediately form the intention to visit Edinburgh based on the secure testimonial knowledge that there exists a decisive reason to do so; or (b) she could wait until she discovers the decisive reason herself, and then form the intention on that basis.

It seems to me that in both cases we don't have deliberation in sense (D2), since she is already certain of what she will do, so there is no choice or intention yet to be figured out in the

Taking into account Clarke's counterexample, Nelkin considers the following alternative condition:

(C) Rational deliberators [about which action to perform] must believe, in virtue of their nature as rational deliberators, that they have multiple alternatives from which to choose, where those alternatives are each consistent with what they are certain of (Nelkin 2004b, 222)

Against this, she argues that (C) purchases immunity from examples such as Clarke's at the cost of explanatory power, for it cannot explain why we cannot deliberate about some kinds of options despite our lack of certainty that they are options we might realize:

...as far as (C) allows, there is a very large number of things that can count as deliberative alternatives. And, yet, there are circumstances in which we seem unable to deliberate in certain situations, precisely because we lack deliberative alternatives. (C) does not have the resources to explain these, since it rules out so little in the way of deliberative alternatives. For example: we seem to be unable to deliberate about whether to jump out of window from a high floor and float on the air currents, despite perhaps lacking certainty about whether this is possible (perhaps we do not rule out a "miracle" or even a perfect sequence of wind gusts). (Nelkin 2004b, 222)

---

practical/epistemic way. As for (D1), on (a), if this immediate formation of the intention counts as an active mental process whose aim is to produce an intention, and this all by itself counts as deliberation (which I would resist, since it does not involve the consideration and evaluation of reasons) then there would be deliberation in this case. On the application of (D1) to (b), maybe it is clearer that forming an intention on the basis of the reason would count as the right sort of active mental process.

Nelkin's point is that there are circumstances in which we lack certainty about whether an option for action is available to us, but yet we cannot deliberate about whether to perform this action. This lack of certainty cannot explain our inability to deliberate here, so what does?

At this point in the discussion I proposed (in correspondence) an alternative belief-in-openness condition (which Nelkin discusses in 2004b, 222-4). Here is a revised version:

(S) In order to rationally deliberate among alternative actions  $A_1 \dots A_n$  (where  $A_1 \dots A_n$  are distinct and incompatible actions), (a) the agent must believe that perhaps she will (i.e., that it is epistemically possible that she will) -- choose from among  $A_1 \dots A_n$ , and (b) the proposition that she does  $A_1$  is consistent with every proposition she, in the present context, regards as settled; and similarly for the proposition that she does  $A_2$ , etc., ... $A_n$ . (Note that the condition does not require that the agent believe (b)).

Two points of clarification. First, how should 'perhaps she will choose A' be spelled out -- what sense of epistemic possibility should be invoked? I suggest: the possibility of S's choosing A is not closed off by any credence that the relevant agent has; that S chooses A has a non-zero credence for that agent. Second, what exactly is it for an option for action to be consistent with what, in the agent's context of deliberation, she regards as settled? Here is my suggestion:

(Settled) An agent regards a proposition as settled just in case she has no doubt that it is true, or else, she disregards any such doubt she has, e.g., for the purpose of deliberation.

Condition (S) together with (Settled) plausibly delivers the desired result for cases like Clarke's

and cases like Nelkin's. Edna can deliberate about whether to go to Edinburgh, because she believes that perhaps she will (i.e., it is epistemically possible that she will) choose between going and not going to Edinburgh; and her making either choice is consistent with what she regards as settled. That is, she is not certain that she will go, nor that she will not, and she does not disregard the doubts that render her making either of these choices uncertain. As a result, that she makes either of these choices is consistent with every proposition about which she has no doubts, and with every proposition about which she disregards any doubts she might have. But I am unable deliberate about whether to float out of the window, for even though I am not certain that I will not do so, the claim that I will is inconsistent with a proposition I regard as settled, since for the purposes of deliberation I disregard any doubt I have that I will not float out of the window (because I disregard any doubt I have that I cannot float out the window).

Kapitan plausibly contends that an epistemic condition on openness should not require consistency with all of the agent's beliefs – for instance, not with all of those that are merely dispositional in the present context (1986, 239). For this reason, in his

(PC) an agent presumes that his  $\phi$ -ing is an open alternative for him only if he presumes

that if  $S$  is any set of his beliefs then his  $\phi$ -ing is contingent relative to  $S$

Kapitan puts the restriction 'contingent relative to  $S$ ' within the scope of the attitudinal operator 'presumes.' Alternatively, one might leave the relevant restriction outside the scope of the attitudinal operator, but then independently restrict the relevant class, as I do by specifying it as

“what she regards as settled.” My main reason for opting for an external reading of the quantifier is to avoid the excessive cognitive sophistication that Coffman and Warfield warn against (which we will discuss in section 5). But perhaps that the concern raised by following Kapitan and placing “every proposition she, in the present context, regards as settled” inside the scope of the belief operator isn’t severe, since even animal deliberators are sensitive to inconsistency, and the conditions we’re adducing are on rational deliberation.

To support his internal placement of the quantifier, Kapitan adduces a case in which at 10 a.m., Mr. Hawkins, having decided to take his son bowling at 3 p.m., acquires the belief that he will take his son bowling then. He takes it as settled and asks his secretary to remind him of his commitment at 2:30. At 2:29 p.m., temporarily overlooking his earlier resolve, he deliberates about playing golf at 3 p.m. Given all that he believes at 2:29 pm. or regards as settled, it is not true that he assumes it possible that he play golf at 3 p.m. One can forget something one believes without ceasing to have that belief. When Hawkins is reminded by the secretary at 2:30 pm of his commitment he does not acquire a new belief. Instead, his attention is refocused upon a belief he already has (1986).

But my formulation of (S) specifies that the proposition that she does A1 etc. is consistent with every proposition she in the present context regards as settled. In the Mr. Hawkins case, it might be intuitive that at 2.29 he does not regard it as settled in the present context that he will not play golf at 3.00, since he has temporarily overlooked his original resolve. If he has temporarily overlooked his original resolve to take his son bowling at 3.00, is it clear that at 2.29 he can be certain that he will not play golf at 3.00? Or if he was never

certain, but instead disregarded any relevant doubts he had, maybe we'd want to say that at 2.29 those doubts, or else beliefs incompatible with taking his son bowling at 3, are seriously in play in this context.

Will rational deliberators actually have beliefs, in our dispositional sense, of type (S) together with (Settled)? Suppose that we all agree that when a rational agent deliberates about whether to do A or refrain from doing A, she will believe that it's possible for her to do A and refrain from doing so. One philosophical interpretation of these beliefs is that the notion of possibility employed is metaphysical, and this interpretation is supposed by Ginet (1966), Taylor (1966), and van Inwagen (1983). But it is very often unclear whether the notion of possibility employed in a belief is metaphysical or epistemic. It may turn out (1) that for most rational agents, the notion of possibility that figures into these beliefs is epistemic. If that's so, then an epistemic condition such as (S) captures the actual openness beliefs that most rational deliberators have. Or it may turn out (2) that most rational agents have beliefs that employ a metaphysical notion of possibility. If that's so, then the advocate of (S) could argue (2a) that while these agents have the metaphysical-possibility beliefs, they typically will also have the epistemic-possibility beliefs, and the latter is all that's required for rational deliberation by way of beliefs in openness. Then, while most rational agents, while deliberating, actually have beliefs that might conflict with a belief in determinism, all that's required for rational deliberation are openness beliefs that they also have that would not so conflict, to which they could retreat. Or else, the advocate of (S) could argue (2b) that while most rational agents have the metaphysical-possibility beliefs, they do not also have the epistemic-possibility beliefs, but

adopting the epistemic-possibility beliefs is all that's required for rational deliberation by way of openness beliefs. Then, while most rational agents, while deliberating, actually have beliefs that might conflict with a belief in determinism, all that's required for rational deliberation are beliefs that they do not have, but might adopt, that would not so conflict. I would like it best if (1) were true, (2a) is my second choice, and (2b) my third.<sup>7</sup>

3. Belief in the efficacy of deliberation is required in addition.

If (S) is to provide a successful compatibilist account of deliberation, then agents who satisfy (S) in situations in which they believe that determinism (and its consequences) are true

---

<sup>7</sup> Is accepting (S) and (Settled) consistent the hard incompatibilism that I have defended (Pereboom 2001)? 'Free will,' as I apply this notion in the context of the philosophical debate, designates the sort of free will required for moral responsibility in the "basic desert" sense, or perhaps more precisely, whatever sort of control in deciding and acting is required for moral responsibility in this sense. What my view claims is that human actions are never freely willed in the sense required for moral responsibility in the "basic desert" sense.) This technical use of 'free will' is not intended to correspond with exactness to the use of the term 'free will' in ordinary language -- or in philosophy generally speaking, for that matter. It is consistent with hard incompatibilism that there are correct applications of the ordinary language term 'freely willed' to human actions. Frankfurt, for example, specifies that a person acts freely and of his own free will just in case he wills X and wants to will X, and wills X because he wants to will X (1971). The arguments for hard incompatibilism pose no challenge to the proposal that we sometimes act freely and of our own free will in Frankfurt's sense, and the hard incompatibilist position is consistent with it. And it may well be that the meaning of our ordinary language term 'free will,' and our ordinary thinking relevant to this concept, allows for Frankfurt's use.

So now consider this principle:

A person S deliberates about whether to perform exactly one action among a set of mutually excluding actions A1...An only if S believes it's possible for her do each of A1...An.

On an epistemic or compatibilist reading of 'possible,' my view has no worries about endorsing the claim that for a set of mutually excluding actions A1...An, (an actual human agent) S can truly believe that it's possible for her do each of A1...An. For, in my view, the kind of freedom



should be capable of rational deliberation – in particular, they should be able to deliberate without having inconsistent beliefs. Against this, there is a type of situation, first brought to our attention by van Inwagen, in which an agent who satisfies (S) would be incapable of deliberation. He illustrates it with the following example:

...imagine that [an agent] is in a room with two doors and that he believes one of the doors to be unlocked and the other door to be locked and impassable, though he has no idea which is which; let him then attempt to imagine himself deliberating about which door to leave by (1983, 154)

About this example, Nelkin remarks, to my mind correctly: “while it seems that I can deliberate about which door to decide to try to open and even which door handle to decide to jiggle, it also seems that I cannot deliberate about which door to open” (cf. Kapitan 1986, 247). But my opening door #1 is consistent with what I regard as settled in the sense specified above, as is my opening door #2. Thus this example poses a threat to (S) together with (Settled) as a compatibilist account for beliefs required for deliberation. What’s more, if one believed determinism and its consequences, then in any deliberative situation one would believe that all but one option for what to do was closed off; “locked and impassible,” so to speak. If in the example one cannot deliberate about which door to open, and one believed determinism and its consequences, then it seems that one would never be able to deliberate about what to do.

Kapitan (1986, 247) suggests, and Nelkin agrees, that this case indicates that rational deliberation requires a belief in the efficacy of deliberation: rational deliberators must believe that for each of the options for action under consideration, deliberation about it would, under

---

involved here would not be enough for moral responsibility.

normal conditions, be efficacious in producing the choice for that action and the action itself. The key insight is that it is not the absence of a belief in openness that precludes deliberation about which door to open. Rather, what precludes such deliberation is that given the agent's belief that one of the two doors is locked, if she is rational she will believe that her deliberation would not ultimately be efficacious in her opening of one of the doors.

Nelkin then correctly points out that (S) does not capture a requirement of a belief in the efficacy of one's deliberation. However, belief in deliberative efficacy is plausibly distinct from belief in openness – belief in the possibility of more than one option from which to choose. So it makes sense for (S) to stand as a compatibilist condition designed to cover belief in openness. But now the question is: can a distinct compatibilist condition be formulated to capture the requirement for a belief in deliberative efficacy? If so, then we needn't think of (S) as having to bear the compatibilist's burden of answering the two-door problem, but only as accounting for openness, since the issue of deliberative efficacy raised by this problem would be addressed by a separate condition.

So how might the compatibilist belief-in-deliberative-efficacy condition be formulated?

Kapitan suggests the following:

(PE) an agent presumes that his  $\phi$ -ing is an open alternative for him only if he presumes that he would  $\phi$  if and only if he were to choose to  $\phi$ . (Kapitan 1986, 234)

One concern for this proposal is that, as Nelkin persuasively argues, the key belief is not in efficacy between choice and action, but in efficacy between the agent's deliberation on the one hand, and her choice and action on the other (2004b). One might imagine a case in which

my deliberation could not produce a choice because it is psychologically impossible for me to make that choice, but nonetheless if it were made, it would result in action. Then, intuitively, deliberation would not be efficacious. A second concern is that deliberative efficacy is a causal notion (as Kapitan himself notes, 234). Causation is unidirectional, here from deliberation to choice and then to action, and, for instance, not from action back to choice. But in Kapitan's formulation the relation between choice and action is bidirectional – it is expressed biconditionally. Moreover, the causal nature of efficacy should be explicit, while Kapitan's biconditional characterization does not satisfy this desideratum.

Clarke suggests the following belief-in-deliberative-efficacy requirement (without endorsing it – he is not committed to the claim that deliberation requires any beliefs about one's abilities) (1992, 103):

(CF') In order to deliberate, an agent must believe that there are at least two distinct actions, A and B, such that (i) were she to choose to A (B), she can A (B) on the basis of that deliberation.

We might specify that 'on the basis of' should be read causally – (i) would then be interpreted as 'were she to choose A, her deliberation would (also) cause her to A.' In addition, Clarke speculates that an agent must, at most, believe that

(ii) if she finds better reason to do A (B), she can decide to A (B) and (iii) if there is better reason to A (B), she can find it.

I think that this is close to a correct set of requirements. But I'm not convinced that in order to deliberate about whether to do A or B, an agent must have a belief that if there is a better

reason to A, she *can* find it. For it seems that an agent can deliberate about whether to A if she only has the belief that she *might* be able to find a better reason to A if there is one – I would not be surprised if this situation were more common than one in which the agent believes that she actually can find the better reason if there is one. Moreover, suppose someone asks me to choose between one of two doors, behind one of which there is a prize which I would win if I chose that door. I know that there is a reason to choose one door over the other – there is a prize behind that door, and not the other – but I also believe that the reason is not accessible to me. Nonetheless, I can engage in an active mental process whose aim is to choose one of the two doors. So I can deliberate, given our characterization of deliberation, about which door to choose, even though I believe that if there is a better reason to choose one over the other, I cannot find it.

Assuming that the pre-decision result of deliberating about whether to do A1 or A2 is judging which it would be best to do, here is my proposal for a condition for belief in deliberative efficacy:

(DE) In order to rationally deliberate about whether to do A1 or A2 (where A1 and A2 are distinct and incompatible actions), an agent must believe that if as a result of her deliberating about whether to do A1 or A2 she were to judge that it would be best to do A1, then, under normal conditions, she would also, on the basis of this deliberation, decide to do A1, and do A1; and similarly for A2.

It might be that the immediate result of deliberation is often something other than explicitly judging that it would be best to do something – perhaps instead we sometimes fix on an option

for action as preferred without explicitly judging it best. Animals and small children might fix on an option as preferred with the use of demonstrative concepts. (DE) can be amended to allow for an immediate result of this sort. Now notice that (DE) is not met by the agent in the two-door situation, but it is satisfied by someone in an ordinary deliberative situation in which she believes that determinism is true and that she therefore has only one possibility for decision and action -- but she doesn't know which. If she believes that one of either

(i) deciding to do A1 on the basis of deliberation

and

(ii) deciding to do A2 on the basis of deliberation

is such that she cannot do it because determinism is true, but she doesn't know which, she can still meet condition (DE).

(DE) could be made more precise in a crucial respect. Nelkin suggests that one can deliberate about whether to do A while only believing that deliberation *might* be effective.<sup>8</sup> Kapitan points out that I might be aware of my occasional weakness of will, akrasia, and this does not keep me from deliberating. In my definition, there is already wiggle room of the sort required to respond to these concerns: it specifies that "under normal conditions, she would also, on the basis of this deliberation, decide to do A1, and do A1..."<sup>9</sup> Wiggle-room is needed for certain types of non-optimal functioning, and akrasia is perhaps the most significant of these in the present context. But perhaps it would be better to make explicit mention of

---

<sup>8</sup> In her commentary on this paper for OPC2.

<sup>9</sup> On OPC2.

akrasia in DE. However, a concern with adding sophisticated notions to belief conditions is highlighted by Coffman and Warfield (2005) – one would not to endorse conditions that require excessive cognitive sophistication on the part of rational deliberators (I will address this issue shortly).

Another possibility is: “under normal conditions, on the basis of this deliberation, she might decide to do A1, and she might do A1...” But this condition cannot handle a slightly revised version of the two doors case, in which one is aware of a 1/1000 chance that the locked door will spontaneously become unlocked – Louis deRosset has raised such probabilistic two-doors cases in conversation. Even then it seems that it is not possible to deliberate about which door to open, despite one’s believing that one might open each.

Kapitan (in correspondence) proposes an efficacy condition that allows for some appropriate wiggle room

At t1, S presumes that his A-ing at t2 is open only if S presumes at t1 that (i) it is more likely than not that he would A at t2 were he to undertake A-ing at t2, and (ii) it is more likely than not he would refrain from A-ing at t2 were he to undertake refraining from A-ing at t2.

But it seems to me that this condition is challenged by a case of the sort that deRosset suggests – here one door will do. Imagine that you believe that the probability the door being unlocked is .51. I think that here one cannot deliberate about whether to open it, even though you presume that it is more likely than not that you would open the door if you were to undertake doing so, and it is more likely than not that you would refrain from opening the door if you

were to undertake refraining from opening it. On the other hand, it's also clear that if you believe that the probability of the door being unlocked is .99, you can deliberate about whether to open it. Many ordinary cases of deliberating about whether to do something fit this model (more or less). So between .51 and .99 there would seem to be a threshold range. Understanding the nature or structure of this threshold would subserve a more precise formulation of (DE).

4. Both the belief-in-openness condition (S) and the belief-in-deliberative efficacy condition (DE) are needed.

One might suspect that because (DE) yields the right result for the two-door example, and since this example features an absence of openness, that no belief-in-openness condition, such as (S), is required in addition to (DE). But this is incorrect – both sorts of conditions are required. For an agent might satisfy (DE) in a case in which an agent cannot deliberate because a belief-in-openness condition is not satisfied. For instance, I cannot now deliberate about whether to drop everything to become a mercenary in Africa. But my so deliberating does satisfy (DE), since I believe that if as a result of my deliberating about whether to drop everything to become a mercenary in Africa, I were to judge that it would be best to do so, then I would also, on the basis of this deliberation, decide to do so, and do so; and similarly for refraining from this course of action. So (DE) all by itself is incomplete. But in this example (S) together with (Settled) is not satisfied. For my dropping everything to become a mercenary in Africa is inconsistent with a proposition for which in the present context I disregard any

doubts I might have, viz., that I will not drop everything to become a mercenary in Africa. The moral is that (DE) needs to be supplemented by a belief-in-openness condition; and I am recommending (S) together with (Settled).

This example also indicates that both disjuncts of:

(Settled) An agent regards a proposition as settled just in case she has no doubt that it is true, or else, she disregards any such doubt she has, e.g., for the purposes of deliberation.

are needed – that is, both the no-doubt, and the disregard-doubt disjuncts are required. One might suggest that all cases in which an agent cannot deliberate, and in which she is not certain that a relevant proposition is true, are cases in which a belief-in-deliberative-efficacy requirement is not satisfied. However, the example indicates why this is not so. For here (DE) is satisfied, and I am not certain that I will not drop everything to become a mercenary in Africa, and yet I cannot deliberate. But in the present context I do regard this proposition as settled, since I now disregard any doubt I have that this proposition is true. In this case, it is my disregarding or bracketing certain aspects of my epistemic state that keeps me from being able to deliberate, and this is accounted for by the disregard-doubt disjunct of (Settled). This disregard is practical in nature; while epistemically, the doubts are in place, these doubts fail to make deliberation possible, because they have been bracketed for my practical purposes in the present context.

Furthermore, it might be argued that (DE) is deficient, since an agent might satisfy it even though he cannot deliberate due to his being convinced that doing so would be



ineffective since he believes, specifically, that he cannot judge it best to do A.<sup>10</sup> I am convinced that in the present context it is psychologically impossible for me to judge that it would be best to drop everything to become a mercenary in Africa on the basis of deliberating about whether to do so, as a result of which I am unable to deliberate about whether to do so. Still, my so deliberating satisfies (DE). But at this point (S) together with (Settled) can again be brought to bear. My dropping everything to become a mercenary in Africa is inconsistent with a proposition for which in the present context I disregard any doubts I might have, viz., that I will not drop everything to become a mercenary in Africa. So (S) together with (Settled) is not satisfied, and this can explain why deliberation is not possible for me here, even though (DE) cannot.

One might think that the belief-in-openness and belief-in-deliberative-efficacy requirements not really distinct because the former can absorb the content of the latter or vice versa. There is a general reason to think that this is not so. Efficacy of deliberation is a matter of deliberation's causally efficacy in producing choice and action. For this reason, the content of the belief in a belief-in-deliberative-efficacy requirement is appropriately expressed in hypothetical or conditional form: to deliberate rationally, I must believe that if my deliberation has such-and-such an immediate result, I can produce, on the basis of this deliberation, choice and action corresponding to this result. By contrast, as Kapitan in effect argues, openness is plausibly categorical and not conditional: each of several options for what to do must in some sense be possible for me to secure, and not merely on the supposition that some condition is

---

<sup>10</sup> Clarke suggests this line of argument in correspondence.

satisfied (Kapitan, 1986, 241). Consider, for instance, the following proposal for a conditional belief-in-openness requirement:

(CO) In order to rationally deliberate about whether to do A or B, the agent must believe that if she chooses A she can do A, and if she chooses B, she can do B.

Suppose that an agent believes that because determinism is true, one of: (choosing A, choosing B) is precluded for her by the Past and the Laws. Then, it seems intuitive that she does not believe that doing A and doing B are both open to her. But she might still satisfy (CO). Any conditional belief-in-openness requirement could be challenged by a similar argument (cf. Chisholm 1964). Hence, the content of the belief in a belief-in-openness requirement is appropriately expressed in categorical terms. Since the beliefs in deliberative efficacy and openness plausibly differ in logical structure, we have good reason to believe that the content of one cannot be absorbed by the other.

## 5. Objections.

Coffman and Warfield (2005, 40-1) express a concern about certain compatibilist conditions on beliefs required for deliberation requiring too much by way of conceptual or cognitive sophistication. They take one existing proposal, the 'belief in counterfactuals about choices' thesis,<sup>11</sup> to task for requiring the deliberator to handle counterfactuals. Here is their formulation of this proposal:

(BCC) S deliberates among some different courses of action only if S believes of each of

---

<sup>11</sup> As Coffman and Warfield point out, this thesis is endorsed by Kapitan (1986, 241) and Bok (1998, 112-3).

those actions that she would perform it if she were to choose to perform it.

They object as follows:

BCC entails that one deliberates only if one has the cognitive ability to handle counterfactuals. This seems excessive: it seems that there could be creatures that deliberate yet lack the cognitive ability to handle counterfactuals (which perhaps involves something like the ability to “mentally simulate” the obtaining of certain conditions and subsequently make a judgment about a distinct proposition’s truth value under those “mentally simulated” conditions). Certain higher non-human animals may be deliberators that lack the ability to handle counterfactuals. Small children may also be deliberators who lack the ability to handle counterfactuals. (2005, 40-1)

Note first that it is rational deliberation for which conditions are being tested, and the rationality of deliberation itself plausibly requires some cognitive sophistication. Indeed, the belief-type specified by (DE) would involve the ability to handle subjunctive conditionals. The conditionals at issue are not counterfactual, exactly, since they don’t involve suppositions that are contrary to fact, but rather suppositions of various unrealized options for action (one of which may turn out to become factual). This may be significant, for the reason that it may be more likely that young children and animals can simulate unrealized options for action -- even as such -- than for them to represent suppositions as contrary to fact. However, it is plausible to hold that if the deliberators at issue have cognitive abilities at all (and are not merely stimulus-response mechanisms, for example), one would expect them to have the ability to handle subjunctive conditionals. The reason is this: all the actual examples of deliberation

plausibly exhibit causal reasoning at a rudimentary level that involves the ability to handle subjunctive conditionals. A cat might well be in a mental state that we could represent as “if I jumped to the left, that mouse would get away, but if I jumped to the right, it would be trapped against the wall.” To be sure, the cat does not have the means to report these conditionals linguistically, but mental simulation and non-linguistic or non-conceptual representation might well suffice. So it may well be that even causal reasoning at a rudimentary level involves the ability to handle subjunctive conditionals, and young children and deliberating animals have such a capacity for causal reasoning.<sup>12</sup>

Coffman and Warfield have an argument in reserve: “even if every actual deliberator has the cognitive sophistication BCC requires for deliberation, it seems possible that there be a deliberator that lacks such sophistication (2005, 41). However, it might well be impossible to deliberate without thinking of oneself as causally efficacious in realizing various options for what to do – as in the cat example just cited -- and it might well be impossible to think of oneself as causally efficacious in this way without being able to handle the sorts of subjunctive conditionals that this example features.

Notice that, unlike some of its relatives, (S) together with (Settled) does not require the agent to have the concept of consistency.<sup>13</sup> It says that the proposition that she does A1 is

---

<sup>12</sup> For a general defense of animal cognition, see Kornblith (2002), esp. Chapter 2, pp. 28-69.

<sup>13</sup> Coffman and Warfield (2005, 38) formulate the “belief in epistemic possibility” thesis as follows:

S deliberates among some different courses of action only if S believes of each of those actions that her performing it is consistent with certain other propositions she believes. citing Dennett (1984), Jones (1968, 260), Kapitan (1986, 241), and Mele (2002, 906-7).

consistent with every proposition she, in the present context, regards as settled, and similarly for the proposition that she does A2, etc. It does not require that *she believe* that these propositions be consistent with what she regards as settled. That said, I suspect that a good case can be made that all rational deliberators will have sensitivity to consistency and inconsistency, even if they need not have the concept 'consistency.'<sup>14</sup>

Coffman and Warfield take another suggestion to task for ruling out what they call "double-minded deliberation," that is, deliberation about an alternative which one both believes one has, one believes one does not have (2005, 37). This suggestion is the "no belief in ability thesis":

(NBI) S deliberates among some different courses of action only if S does not believe of some of those actions that she can't perform them.<sup>15</sup>

They argue that the possibility of double-minded deliberation rules out proposals according to which the deliberator must lack a belief that he does not have the alternative at issue. But note first that (S) together with (Settled) do not rule out double-minded deliberation, since these conditions specify only that the deliberator must believe that it's epistemically possible that she choose from among various options for action, where it is in fact true that for each option, her performing that action is consistent with what she regards as settled. The conditions do not rule out deliberation in a case in which she also has a belief that she lacks

---

<sup>14</sup> Kornblith (2006) points out that animals and young children have sensitivity to logical notions, but may well lack corresponding concepts.

<sup>15</sup> Coffmann and Warfield note that NBI is endorsed by Bok (1998, 110), Clarke (1992, 110), Kapitan (1986, 235-41), Pettit (1989, 43), Searle (2001); Taylor (1966, chap. 12), and Waller ((1985, 49).

one of the alternatives if at the same time she does not regard it as settled that she lacks this alternative. At the same time, double-minded deliberation about A may be impossible if the agent actually regards it as settled that she lacks one of two relevant alternatives. Finally, it may be that double-minded deliberation does not qualify as rational deliberation. Then, since the proposed conditions at issue are intended as necessary conditions on rational deliberation, it will not be a problem if they fail to be satisfied in cases of non-rational deliberation.<sup>16</sup>

## 6. Final words.

In summary, rational deliberation plausibly requires an assumption of openness, a belief that one has more than one option from which she might choose, and in addition, a belief that one's deliberation can be efficacious in producing choice and action. The deliberation-compatibilist condition:

(S) In order to rationally deliberate among alternative actions A1...An (where A1...An are distinct and incompatible actions), (a) the agent must believe that perhaps she will (i.e., that it is epistemically possible that she will) choose from among A1...An, and (b) the proposition that she does A1 is consistent with every proposition she, in the present context, regards as settled; and similarly for the proposition that she does A2, etc., ... An.

---

<sup>16</sup> Thanks to Louis deRosset for this point. One might add that if the deliberation-incompatibilist's concern is that deliberators who believe that determinism is true will have inconsistent beliefs, and the belief-compatibilist proposes belief-conditions that avoid such inconsistent beliefs, it is dialectically questionable for the belief-incompatibilist to object that the compatibilist proposal fails to allow for deliberation that involves inconsistent beliefs. See

together with

(Settled) An agent regards a proposition as settled just in case she has no doubt that it is true, or else, she disregards any such doubt she has, e.g., for the purposes of deliberation.

provides the requisite account for a belief in openness. While (S) together with (Settled) does not also yield a compatibilist account for the supposition that one's deliberation be efficacious, deliberative efficacy is sufficiently different from openness to warrant a separate condition.

Here I have proposed:

(DE) In order to rationally deliberate about whether to do A1 or A2 (where A1 and A2 are distinct and incompatible actions), an agent must believe that if as a result of her deliberating about whether to do A1 or A2 she were to judge that it would be best to do A1, then, under normal conditions, she would also, on the basis of this deliberation, decide to do A1, and do A1; and similarly for A2.

(S) together with (Settled), and (DE) appear not to be vulnerable to objections that have been raised against other compatibilist proposals for the beliefs required for deliberation, and this in turn provides reason to think that a deliberation-incompatibilism, for example, one that incorporates (I) can be resisted. It might well be that when we rationally deliberate we usually believe, in the dispositional sense, that we have more than one distinct option for which action to perform, each of which is available to us in the sense that we can, in a metaphysical sense, perform each of these actions, and in most cases the specific content of this belief would conflict with a belief in determinism. But when we rationally deliberate we also believe, in the

---

also Neil Levy's discussion of this issue (2006).

dispositional sense, (S) together with (Settled), and (DE), or something close, and retaining these beliefs while relinquishing those that conflict with determinism would suffice.



## References

Bok, H. (1998). *Freedom and Responsibility*, Princeton: Princeton University Press.

Chisholm, R. (1964). "Human Freedom and the Self," The Lindley Lecture, Department of Philosophy, University of Kansas, 1964; reprinted in *Free Will*, Gary Watson, ed., Oxford, Oxford University Press, 1982.

Clarke, R. (1992). "Deliberation and Beliefs About One's Abilities," *Pacific Philosophical Quarterly* 73, 101-13.

Coffman, E. J., and Warfield, T. (2005). "Deliberation and Metaphysical Freedom," *Midwest Studies in Philosophy* 29, 25-44.

Dennett, D. (1984). *Elbow Room*. Cambridge: MIT Press.

Jones, D. (1968). "Deliberation and Determinism," *Southern Journal of Philosophy* 6, 255-64.

Frankfurt, Harry G. (1971). "Freedom of the Will and the Concept of a Person," *Journal of Philosophy* 68, pp. 5-20.

Ginet, C. (1962). "Can the Will be Caused?" *The Philosophical Review* 71, 49-55.

Ginet, C. (1966). "Might we have no Choice?" in *Freedom and Determinism*, Keith Lehrer, ed.  
New York: Random House.

Kant, I. (1785/1981). *Grounding for the Metaphysics of Morals*. J. Ellington, tr. Indianapolis:  
Hackett.

Kapitan, T. (1986). "Deliberation and the Presumption of Open Alternatives," *The  
Philosophical Quarterly* 36, 230-51.

Kornblith, H. (2002). *Knowledge and its Place in Nature*. Oxford: Oxford University Press.

Kornblith, H. (2006). "Reply to Bermudez and Bonjour,"

Levy, N. (2006). "Determinist Deliberations," *Dialectica* 60, 453-59.

Mele, A. (2002). "Review of John Searle's Rationality in Action." *Mind* 111, 905-9.

Nelkin, D. (2004a). "The Sense of Freedom," in *Freedom and Determinism*, J. Campbell, M.  
O'Rourke, and D. Shier, eds., Cambridge: MIT Press, 105-34.

Nelkin, D. (2004b). "Deliberative Alternatives," *Philosophical Topics* 32, 215-40.

Pereboom, D. (2001). *Living Without Free Will*. Cambridge, Cambridge University Press.

Pettit, P. (1989). "Determinism with Deliberation," *Analysis* 49, 42-4.

Searle, J. (2001). *Rationality in Action*. Cambridge: MIT Press.

Stanley, J. (manuscript). "Knowledge and Certainty."

Taylor, R. (1966). *Action and Purpose*, Englewood Cliffs: Prentice Hall.

van Inwagen, P. (1983). *An Essay on Free Will*, Oxford: Oxford University Press.

van Inwagen P. (2004) "Van Inwagen on Free Will," in *Freedom and Determinism*, J. Campbell, M. O'Rourke, and D. Shier, eds., Cambridge: MIT Press, 213-30

Waller, B. (1985). "Deliberating about the Inevitable," *Analysis* 45, 48-52.