



**CARING AND THE BOUNDARY-DRIVEN  
STRUCTURE OF PRACTICAL DELIBERATION**

BY JEFFREY SEIDMAN

## Caring and the Boundary-Driven Structure of Practical Deliberation<sup>1</sup>

Jeffrey Seidman

### 1. Limited deliberative attention as a problem for the theory of practical reasoning

THE SORTS OF PRACTICAL QUESTIONS that agents seek to answer by deliberating are markedly diverse. An agent might deliberate in order to discover the best means to some already given, well-defined end. An agent might deliberate in order to “specify” a general or vague end – asking herself, for instance, “what would be the morally decent thing to do right now?” Or an agent might deliberate in order to answer questions as open-ended as “what shall I do today?” or “what shall I do with my life?” As different as these instrumental, specificatory and more open-ended practical questions are from one another, our deliberative attempts to answer them share a common feature – a feature which is so familiar and so evident to reflection that it is easy to fail to be puzzled by it. In each sort of case, the domain of potential *answers* to her question that an agent will typically entertain in deliberation is extremely limited. Indeed, if the “courses of action” between which a deliberating agent chooses are individuated so as to include complex, temporally extended plans, then, in most cases, there will be indefinitely many courses of action that a finite deliberating agent, with finite cognitive capacities, will never consider – potential answers to her question on which her deliberative attention will simply never alight. This fact raises a fundamental question for the theory of practical reasoning: what determines *which* options an agent will and will not consider in deliberation? Common sense suggests that the deliberation of an agent who was simply *arbitrary* in which possible answers to her practical question she ignored and which she considered would not count as reasonable. So, in order to make a place in the theory of practical reasoning for the fact of our limited deliberative attention, we need an account that at once explains the patterns of deliberative attention that reasonable agents typically display, and allows us to see why these patterns of deliberative attention are, in fact, reasonable.<sup>2</sup>

---

<sup>1</sup> This paper has benefited from the generous comments of Michael Bratman, Jennifer Church, Agnieszka Jaworska, David Hills, an audience at the SENT workshop at Stanford, and the graduate students in the seminar on practical reasoning that I taught at Stanford in 2007.

<sup>2</sup> In his *Change in View* (Cambridge, MA: MIT Press, 1986), Gilbert Harman was among the first clearly to distinguish the study of reasoning – a norm-governed psychological process that may result in changes to a subject’s beliefs or intentions – from the study of forms of argument. (For an earlier, less clear version of this distinction, see G.E.M. Anscombe, “Practical Inference,” in R. Hursthouse, G. Lawrence and W. Quinn, eds., *Virtues and Reasons* (Oxford: Oxford University Press, 1995), pp. 1-34.) This essay aims to contribute to the theory of practical reasoning.

In this essay, I will present a somewhat idealized model of the path that a reasonable agent's deliberative attention will take during the course of temporally extended deliberation. This model, the *Boundary-Driven Model* of deliberative attention, borrows an idea that is central to Michael Bratman's "Planning Theory" of intention.<sup>3</sup> According to the Planning Theory, a reasonable agent brings to the deliberative questions she faces a framework of settled, relatively stable, prior intentions. These intentions often remain "in the background" of her deliberation, outside the scope of her deliberative attention; but *from* the background, they help to frame what appears in the foreground. Intentions can do this because they are subject to strong rational requirements of consistency with one another (relative to an agent's beliefs). Because I intend to teach tomorrow, for instance, and because nothing in my current situation gives me sufficient reason to reconsider that intention, I will not regard courses of action that I believe to be inconsistent with teaching tomorrow as *admissible* options. This makes my deliberative situation, today, much easier than it otherwise might be. Even as I ask myself the apparently wide-open question "what shall I do today?" for instance, there are indefinitely many courses of action, such as catching a flight to Rio de Janeiro, that I will not entertain. Typically, such inadmissible courses of action simply will not cross my mind – even as I search energetically and imaginatively for something new and exciting to do on this dull, grey Sunday. If I do think of such a course of action (perhaps in a daydream), or if someone proposes it to me, I will be disposed to *exclude* it from the set of possibilities I entertain, without pausing to consider its particular merits and disadvantages. As I shall put this point, an agent's prior intentions help to frame her deliberative problems by establishing *deliberative boundaries* – boundaries on the landscape of possibilities, beyond which she will not look for an answer to the question she faces.<sup>4</sup>

While the Planning Theory provides a starting point, an adequate account of deliberative attention must go beyond it. It will be a central claim of the Boundary-Driven Model that the deliberative boundaries that a reasonable agent's *intentions* establish are not the only boundaries that constrain her deliberative attention. The mental state of *caring*, I shall argue, imposes its own deliberative boundaries on the thought of a reasonable agent – boundaries that serve to narrow the field of possibilities over which the agent's deliberation will range. (From here onward, I will use the noun "concern" to refer to the mental state that we ascribe to a person when we say that she *cares* about something.) An agent's concerns, moreover, play a second role, determining the *course* that a reasonable agent's deliberative attention takes through time: they help to determine the relative stabilities of the various

---

<sup>3</sup> See Bratman, *Intention, Plans, and Practical Reason* (Stanford, CA: CSLI Publications, 1999).

<sup>4</sup> The language of "excluding" and the metaphor of a "boundary" are mine; Bratman writes of intentions imposing a "filter of admissibility" on the options an agent will consider in deliberation (*ibid.*, p. 33).

boundaries by which an agent's thought is constrained; and so they help to explain in what order, and in response to what considerations, an agent will *cease* allowing her thought to be constrained by diverse deliberative boundaries which have until then constrained her practical thought. The Boundary-Driven Model reflects both of these ways in which an agent's concerns direct her deliberative attention.

The argument that follows has three parts – the first two descriptive (§§2-3 and §§4-6), the third normative (§§7-8). In §§2-3, I argue that in order to explain the deliberative boundaries that structure our thought, we need to look beyond the resources of the Planning Theory, to our concerns. I make two, central claims, *Exclusion* and *Stability*, about the ways in which concerns give rise to and sustain deliberative boundaries. In §§4-6, I present the Boundary-Driven Model as a modest extension of those twin claims to instances of temporally extended deliberation, and I show that the model illuminates the rich, temporal structure of an instance of deliberation that is detailed enough to be lifelike, and that will strike many people, prior to philosophical theory, as a paradigm of reasonableness. These sections aim to establish the Boundary-Driven Model's claim accurately to describe and elaborate one part of our pre-theoretical conception of reasonableness in practical thought.

The Boundary-Driven Model does not constitute a general theory of practical deliberation. While it explains some of the ways in which deliberative boundaries direct an agent's deliberative attention and thus structure her deliberative agenda, it is silent as to how she will solve the deliberative problems that her deliberative boundaries help to frame. Indeed, it is the task of the third, normative, part of this essay (§§7-8) to argue that boundary-driven deliberation (deliberation in accordance with the Boundary-Driven Model) not only accords with our pre-theoretical conception of reasonableness; it can in fact be seen to *be* reasonable in the context of strongly divergent theories of practical reason. The model's import, however, does vary significantly depending on the more general theory of practical reason in which it is situated. If the model is located in the context of a maximizing conception of practical reason (whether the maximand is conceived as preference-satisfaction, the realization of objective value, or something else), then boundary-driven deliberation can be understood and justified instrumentally, as a heuristic device, of the sort recognized by theorists of bounded rationality. On the other hand, if we suppose that there is no single index of value that successful practical choice maximizes, then boundary-driven deliberation can be understood and justified in a different, and perhaps more interesting, way: an agent whose deliberation is boundary-driven frames the deliberative problems she poses herself narrowly enough that, in conjunction with deliberative devices which are not part of the model but which are compatible with it, she may be able to give a determinate, verdictive sense to the phrase "the best courses of action available to me" in cases in which a determinate meaning for this phrase would otherwise be lacking. In the context of an in-

commensurabilist conception of practical reason, boundary-driven deliberation is of more than merely instrumental value: it is partly *constitutive* of reasonableness in practical thought.<sup>5</sup>

## 2. Explaining the stability of deliberative boundaries arising from intentions

An agent's settled intention can establish boundaries for her deliberative attention only so long as she does not reconsider that intention itself. (One of the most obvious places to start, in reconsidering some prior intention, is to look *beyond* the boundaries it establishes, by considering the advantages and disadvantages of the courses of action which sticking with that intention would preclude.) It follows that, in order to explain the fact that an agent's deliberative attention is constrained by some boundary established by her prior intention, we need to explain the *stability* of that intention. (Following Bratman, I will use the term "stability" to describe, not the extent to which an intention resists alteration when reconsidered, but rather the extent to which the intention resists *reconsideration*.)

In fact, Bratman offers a powerful, two-part explanation for the stability of our intentions – and thus of the deliberative boundaries to which they give rise. The first part of the explanation depends upon the fact, which I noted in §1, that we are finite creatures, with limited cognitive resources. If, at every moment, facing an unsurveyably large range of possible courses of action, we had to ask ourselves a question no more limited than "what is the best thing that I could do right now?" – starting, as it were, from scratch each time – our deliberation would never get far.<sup>6</sup> Our ability to commit ourselves to future courses of action, by forming stable intentions, allows us to think carefully and well about *some* possibilities, because it allows us to treat some practical questions as *settled* (at least for now, absent sufficient reason to reconsider).

The second part of Bratman's explanation would hold even for creatures whose cognitive capacities were not finite. *Any* deliberator, even a cognitively unlimited one, will often face "Buridan cases," in which no single course of action is clearly better than all of the others open to her. An agent who knew that she would face a Buridan case later on, but who was unable to commit herself in advance to some one course of action, would be unable to anticipate, in advance, which course of action she would eventually settle upon – no matter how extensive the cognitive resources she brought to bear on the question.<sup>7</sup> And so such an agent would be unable to coordinate her own activities, over time. Neither, by the same token, would other agents be able to

---

<sup>5</sup> Truth in advertising: I am a partisan incommensurabilist. I hope to show, however, that accepting the Boundary-Driven Model requires no commitments on this question.

<sup>6</sup> Cf. *ibid.*, p. 28.

<sup>7</sup> Cf. *ibid.*, pp. 11-12.

coordinate their activities with her. Our ability to commit ourselves to future courses of action by forming stable, future-directed intentions frees us from such difficulties. Stable intentions are, in a nutshell, a solution to what would otherwise be insoluble problems of intrapersonal and interpersonal coordination.

Bratman's account of the stability of intentions is compelling. I want to argue, however, that it is incomplete as it stands. To see this, we need only note that some intentions are *more* stable than others. When I drive from Palo Alto to San Francisco this afternoon, I plan to take 280 rather than 101.<sup>8</sup> It's a prettier road, and it's less likely to be congested. My intention is stable: absent a good reason to reconsider it, I will take 280 without further thought. But it will take very little to make me reconsider, and indeed alter, this intention – and so disregard the deliberative boundaries that it establishes. News of congestion on 280, or a request by my wife that I stop by the IKEA off 101, would do the trick. It would take *weightier* considerations, however, to lead me to reconsider my intention to attend the philosophy department colloquium this Friday. It is on a topic close to my own work, and the speaker is a philosopher whom I greatly admire. Merely discovering that attending will be more inconvenient than I had anticipated, or that there is some family shopping that the talk will prevent me from doing, would not lead me to reconsider my intention. On the other hand, if at the last minute on Friday afternoon I learned that the travel agent with whom I had purchased airline tickets would be closed over the weekend, so that Friday was my last chance, before the Monday flight, to pick up the tickets for next week's family vacation, I would reluctantly reconsider my intention to attend the colloquium. It would take still weightier considerations than these, however, to lead me to reconsider, let alone alter, some other intentions of mine. I intend, for example, to repay the money I recently borrowed from my brother. Learning that repaying my brother will require me to forgo a vacation I had very much looked forward to, contrary to what I had supposed when I borrowed the money, will not lead me to reconsider this intention.

There is a natural way to explain the fact that my intention to repay my brother is more stable than my intention to attend the colloquium this Friday, and that this in turn is more stable than my intention to take 280. I don't really care which highway I take. I do care, however, about attending the colloquium. And I care about repaying my brother *much more deeply* than I care about attending the colloquium. I care deeply about repaying my brother because I care deeply about acting honestly in general, and also because I care deeply about interacting with my brother, in particular, on terms of mutual respect and (deserved) implicit trust. In this essay, I will presuppose no particular analysis of what caring consists in, nor therefore of what the "depth"

---

<sup>8</sup> Bratman uses this as an example of a "Buridan" case, in which both options are equally desirable. I have borrowed the example. In my use of it, however, it is not a Buridan case.

of a concern consists in.<sup>9</sup> Rather, I will trade upon whatever grip we have on the notion of caring prior to philosophical theory. I think that even in advance of any analysis, our pre-theoretical understanding of caring warrants, and gives us some purchase on, the suggestion that an agent may care about something more or less deeply. I will register two, minimal assumptions about the relation between our concerns and our reasons in §3, and I will say more about this relation in the context of competing theories of practical reason in §§7-8. My aim throughout this essay, however, is to rely only on claims that are compatible with a wide variety of views about the nature of caring and its connection to reasons.

In keeping with this philosophical ecumenicalism, I will leave it an open question what sort of *further* explanation or justification our concerns themselves admit of.<sup>10</sup> Perhaps, for instance, an agent's concerns reflect or ought to reflect her judgments of value, so that we will not have a *full* explanation of the stability of a reasonable agent's intentions until we understand the judgments of value that in fact "lie behind" her concerns.<sup>11</sup> Be that as it may, it seems clear that knowing the concerns to which an agent's intentions minister can help us to understand the stability of her intentions *better* than merely knowing her intentions can. And it seems clear, moreover, why this is so: concerns come in varying depths, while intentions do not.<sup>12</sup>

---

<sup>9</sup> I develop an account of caring in "Valuing and Caring," draft ms. That account has two, complementary components. On the one hand, caring involves a set of emotional dispositions, of roughly the sort identified by Agnieszka Jaworska in her "Caring and Internality," *Philosophy and Phenomenological Research* 74(3) (2007), pp. 529-568. On the other hand, in competent, rational adults, caring also involves a set of distinctive *cognitive* and *deliberative* dispositions. These include the dispositions that I will describe in the theses *Exclusion* and *Stability*, later in this essay. And they include a defeasible disposition to believe certain considerations to be reasons for action. A deeper concern, on this account, is a concern whose constitutive dispositions are tied in tighter, more multitudinous and more thoroughgoing ways to other elements in the subject's psychic economy than are those of a shallower concern, so that *altering* a deeper concern would involve more numerous and greater changes to the rest of her psychic economy. And a deeper concern disposes an agent to believe relevant consideration to be *stronger* reasons for action than does a shallower concern. Nothing in the current essay depends on this account of the nature of caring and of the depth of a concern, however.

<sup>10</sup> Among the questions an adequate account will have to address are: how stable are our concerns themselves, and what accounts for *their* stability?

<sup>11</sup> For the record, I think that the concerns of a reasonable agent are sensitive to, but underdetermined by, her judgments of value. Insofar as an agent is reasonable, she will not care about something she believes to be valueless, or to possess insufficient value to be worthy of her care. But a reasonable agent may recognize that two objects are both valuable, and that it would not be unreasonable of her to care about either or both of them, and yet she may care about only one of them. While I think that these claims are true, nothing in this essay depends upon them.

<sup>12</sup> Intentions are often embedded in one another hierarchically, and differences in hierarchical status may involve differences in stability. An agent's specific intention to see a particular film may be less stable, for instance, than the more general intention, in which it is embedded, to do something fun this evening. But hierarchy is not the same as depth, and greater hierarchical status does not entail greater stability. Michael might intend to marry someone (a

So far, we have a two-level explanation for the stability of an agent's deliberative boundaries. An agent's intentions give rise to deliberative boundaries – that is, they dispose her to exclude certain courses of action from the range of possibilities she entertains in deliberation. These intentions, just as such, have a certain in-built stability; if they did not, they could not serve, as we need them to do, as solutions to problems of intrapersonal and interpersonal coordination. It is this in-built stability which constitutes the stability of my intention to take 280 this afternoon, and so explains the stability of the deliberative boundaries to which that intention gives rise. But the stability of some deliberative boundaries also admits of a second level of explanation, because some intentions are “backed” by concerns to which they minister. (A single intention may, on this picture, be backed by multiple concerns. I care about understanding practical reasoning, and I also care about supporting my family financially. My intention to continue going to work each day ministers to both of these concerns.) Where some intention of an agent is backed by the agent's concerns, her concerns add to the stability of that intention, and thus of the deliberative boundaries to which her intention gives rise. The deeper the concerns, the more stable the deliberative boundaries.<sup>13</sup>

---

general intention he has harbored since childhood), and he also might intend, now, to marry Susan. His intention to marry someone may be no more stable than his intention to marry Susan, and he may care less deeply about the former than about the latter.

<sup>13</sup> While these claims supplement Bratman's account of the stability of intention, they are not incompatible with it. Bratman stipulates that an agent's intentions would exhibit “ideal stability” just if the agent were disposed to reconsider them on all and only those occasions in which reconsidering the intentions would in fact lead the agent to alter them, and in which the expected advantages of altering them would outweigh the expected costs (in time, energy, etc.) of reconsidering (*ibid.*, pp. 72-73). Thus, Bratman could acknowledge that the “ideal stability” of an intention that is backed by a deep concern would be greater than the “ideal stability” of an intention not backed by any concern, precisely because if the agent *were* to reconsider the intention backed by the deep concern, she would be less likely to *alter* it.

The intentions of a real agent cannot exhibit ideal stability. In order to explain the *actual* stability that a reasonable agent's intentions display, Bratman makes the sensible point that an agent's responses to different features of her environment are typically conditioned by an array of enduring habits and propensities – habits and propensities which will bestow “reasonable stability” on an agent's intentions just if they tend, in general, to produce results which approximate sufficiently well to those of ideal stability. Such habits and propensities can explain the fact that one agent's intentions might exhibit greater stability overall than another agent's intentions; and they can explain the fact that a single agent's intentions might be more stable in response to one sort of environmental stimulus and less stable in response to another. Thus, one sort of agent might be particularly prone to notice possible danger in her environment, and particularly prone to reconsider her intentions in the face of it. A different sort of agent might not be very sensitive to danger at all, but might be particularly disposed to notice unexpected opportunities for bodily pleasure, and particularly prone to reconsider her intentions in the face of such temptations. The habits and propensities to which Bratman appeals explain the fact that an agent is the kind of person whose intentions are *generally* stable or instable in the face of one sort of environmental stimulus or another. My account, on the other hand, explains the fact that, for one and the same reasonable agent, different *particular* intentions (backed by different concerns) might exhibit differing stabilities, even in the face of similar environmental stimuli.

### 3. Deliberative boundaries issuing directly from an agent's concerns

In the two-level explanation of the stability of deliberative boundaries, intentions play an ineliminable role. An agent's concerns may contribute to the *stability* of some deliberative boundary, but an intention is always necessary to explain the *existence* of that boundary in the first place. I want to argue now, however, that an agent's concerns give rise to certain deliberative boundaries *directly*, without any intermediary role played by intentions. In particular, I want to argue that the following claim expresses a conceptual truth about caring:

*Exclusion:* For any end,  $e$ , if an agent cares about  $e$ , then (whether or not the agent intends to bring about or promote or sustain  $e$ ) the agent is disposed to exclude courses of action that she believes to be incompatible with  $e$  from the range of courses of action that she will entertain as options in practical deliberation.<sup>14</sup>

In what follows, I will use the phrase “the deliberative boundary established by an agent's concern with  $e$ ” to refer to that deliberative boundary manifested in the disposition that *Exclusion* describes: the disposition to exclude precisely those courses of action which the agent believes to be incompatible with  $e$  from the range of courses of action which she will entertain as options in practical deliberation.

---

We could put my account in Bratman's terms by saying that, in order for an agent's intentions to exhibit reasonable stability, her dispositions to reconsider them must be conditioned not only by appropriate sensitivities to threats and opportunities afforded by her external environment, but also by an appropriate sensitivity to the relations between her own intentions and her concerns.

<sup>14</sup> *Exclusion* describes a conceptual connection between an agent's concern with some *end* and the agent's deliberative dispositions. I do not mean to suggest that all of an agent's concerns have ends as their objects; in fact, most paradigm instances of concern do not. An agent may care about a person, or a work of art, or an institution, or an aesthetic or moral ideal. Moreover, often (though not always), where an agent does care about some end (that her child should develop a sound moral character, or that a particular gallery should have the funds to continue purchasing new works of art, for instance), her end-directed concern derives from or manifests an explanatorily prior concern with some non-propositional object (her child, the art gallery). Nor can an agent's concerns with such non-propositional objects be *reduced* to concerns with ends. (For an argument to this effect, cf. Agnieszka Jaworska, “Desire, Caring, and the Structure of Practical Reasons,” draft ms.) Nonetheless, in order to explain the impact of an agent's concerns with non-propositional objects on her practical deliberation, we must often look to end-directed concerns that those concerns generate. Thus, we can explain the fact that an agent who cares about an art gallery is disposed to exclude from her deliberation courses of action which would cause the art gallery to go bankrupt by noting that an agent who cares about an art gallery will care that the art gallery be in sound financial health, and then noting that this latter, end-directed concern involves the kind of disposition to exclude that *Exclusion* identifies.

The truth of *Exclusion* is easily obscured, for the following reason: often, when an agent cares about some end, her concern will be manifested in *intentions* with respect to that end, and these intentions will themselves give rise to deliberative boundaries, evident in corresponding dispositions to exclude courses of action which lie beyond these boundaries. Because I care about finishing this article, for instance, I *intend* to finish this article; and this intention disposes me to exclude from my deliberation courses of action incompatible with finishing it. So, in order to see that concerns give rise to deliberative boundaries *directly*, we need to consider a case in which an agent cares about something, but in which she has no corresponding intention – for instance, because it seems to her that there is nothing she can do that will have any causal bearing on the object of her concern.

The following case will do: I care, let us suppose, about the continuance of Tibetan culture, even as I believe (perhaps over-pessimistically) that there is nothing at all I can do about its ongoing decimation as a result of Chinese government policy. Because I believe there is nothing I can do about the continuance of Tibetan culture, there is nothing I intend to do or try to do about it. That is, there is no positive intention associated with my concern. Nonetheless, my concern generates a deliberative boundary. The boundary would become evident if, contrary to my current supposition, I discovered that my actions might have some indirect, causal bearing on the fate of Tibet. If, for instance, in the course of searching for a job, I found that one of the possibilities open to me was at a public relations firm that works on contract for the Chinese government, seeking to counter criticism of China's role in Tibet in the international media, I would, in normal circumstances,<sup>15</sup> exclude this option from the set of possibilities I considered. If someone were to ask me why I refused to consider this job, I would explain myself (and attempt to justify myself) by noting that I care about the fate of Tibet.

With this example before us, I can offer the following, simple argument for the truth of *Exclusion*. Suppose, to revert to the example, that I had *no* disposition, however defeasible, to exclude from my deliberation courses of action which aided in the destruction of Tibetan culture. Insofar as an observer could know this, she would be warranted in taking it as *conclusive* evidence that I did not, after all, care about the continuance of Tibetan culture. I think that the example generalizes: in any similar case, we would be warranted in taking the absence of an exclusionary disposition as conclusive evidence for the absence of caring. The connection that *Exclusion* purports to identify between caring and this sort of exclusionary disposition is a tight, conceptual connection. I do not have a more elaborate argument to offer for the truth of *Exclusion*. But I think that when we frame examples so as to isolate the concepts in question, and then test our intuitions in these cases, the verdict is clear. Moreover, *Exclusion* will play a central role in the Boundary-

---

<sup>15</sup> Not in every circumstance. I will argue in §4 that the deliberative boundaries that concerns impose on our thought are provisional and defeasible.

Driven Model of deliberative attention. In §§4-6, I will try to show that the model describes the way in which agents whom we ordinarily regard as reasonable, prior to philosophical theory, deliberate. And in §§7-8, I will argue that deliberation in accordance with the model is in fact reasonable. Insofar as that Boundary-Driven Model proves plausible as a whole, it will confer further credibility on *Exclusion*.

One might grant that there is a conceptual link, of the sort that *Exclusion* identifies, between my concern with Tibet and the deliberative boundary by which my thought is constrained, and still insist that intentions must play the intermediary role that the two-level explanation presupposes. For it might be claimed that, in the light of my explanation for refusing to consider the job, it is legitimate to ascribe to me a “negative intention” – an intention *not* to aid in the destruction of Tibetan culture – even if there is no positive intention associated with my concern for Tibet.<sup>16</sup> This negative intention, it might be claimed, is itself conceptually bound up with my concern: if we could not ascribe the intention to me, we could not ascribe the concern. But it is the negative intention that must explain the deliberative boundary, it might be argued. After all, we understand *why* intentions give rise to deliberative boundaries: they are subject to rational requirements of consistency with one another, without which they could not play the coordinating roles they do in our practical lives. If concerns did not involve negative intentions, with their associated norms of consistency, the two-level theorist might claim, it would be mysterious why they should give rise to deliberative boundaries, as *Exclusion* rightly claims they do.

The two-level theorist I am envisaging is partly right: in virtue of my concern with the continuance of Tibetan culture, and the associated disposition to exclude courses of action which are incompatible with its continuance, it is legitimate to ascribe to me the *aim* of not aiding in the destruction of Tibetan culture. We can borrow (and alter) a term from Kant to express this: in virtue of my concern with Tibetan culture, the continuance of Tibetan culture is for me a *negative end* – an end that I *aim* never to “act against.”<sup>17</sup> This negative end is manifest in the deliberative boundaries by which my thought is constrained, and is as conceptually bound up with my concern as are they. But does having this negative end amount to having a negative *intention*? It does not. For the negative ends that our concerns give us

---

<sup>16</sup> For the term “negative intention,” see Gilbert Harman, “Practical Reasoning,” in A. Mele (ed.) *The Philosophy of Action* (Oxford: Oxford University Press, 1997), pp. 149-177, at p. 157.

<sup>17</sup> In explaining the Formula of Humanity, Kant describes the rational nature of persons as an end that can “be conceived only negatively,” in the sense that it “must never be acted against” (*Groundwork of the Metaphysic of Morals*, trans. H.J. Paton (New York: Harper Torchbooks, 1964), at p. 437 in the *Akademie* pagination). My usage differs from Kant’s in two respects: (1) On my usage, but not on Kant’s, the status of negative ends is *defeasible*: sometimes, a rational agent will act against her negative ends. (2) On my usage, but not on Kant’s, an end may be a negative end for an agent on one occasion, but a positive end, which she seeks actively to promote or bring about, on another.

are not subject to rational requirements of consistency of the sort that govern intention – requirements which, on the two-level view, are supposed to explain our deliberative boundaries.

If I intend to X and I intend to Y, and yet I know that X and Y are not co-realizable, I am subject to rational criticism.<sup>18</sup> But suppose that I care about publishing enough to get tenure, and that I also care about spending significant time with my family. We can make sense of the idea that I have negative ends associated with each of these concerns: I aim not to “act against” either one of these ends, and so am *disposed* to exclude from my deliberation courses of action which would be incompatible with either of them, as *Exclusion* claims. If my circumstances are unfortunate, however, I may recognize that these two negative ends will be in frequent, recurring conflict, so that I will often have to act in a manner incompatible with one of these negative ends in order to avoid “acting against” the other, choosing a course of action which one of my concerns *disposes* me to exclude from my thought. This would be an uncomfortable position to be in, and it might present me with severe deliberative problems. I might respond by giving priority to different ends on different occasions, depending on the context; or, I might fairly consistently side with one of these ends over the other. But even if I consistently decide to give short shrift to one of these ends, I am under no rational pressure to cease caring about it, or to cease *aiming* to accommodate it – and doing so insofar as circumstances and the demands of my other concerns permit.<sup>19</sup> This “aiming,” therefore, cannot be intending.<sup>20</sup>

---

<sup>18</sup> Not all philosophers accept this. Niko Kolodny has recently argued, for instance, that intentions are not subject to (“wide-scope”) rational requirements of consistency at all. (See his “The Myth of Practical Consistency,” *European Journal of Philosophy* 16:3 (2008), pp. 366–402.) On Kolodny’s view, there is no difference between intending and what I call “aiming” (or what Bratman, at p. 129 of his *op. cit.*, terms “endeavoring”). While my argument in the main text presupposes that Kolodny’s view is false, his view does not in fact threaten the claim for which I am arguing. I mean to argue that concerns can give rise to deliberative boundaries “directly,” rather than by way of intermediating intentions. If Kolodny is right, and intentions are not in fact subject to strong rational requirements of consistency, then intentions do not give rise to deliberative boundaries at all. So, if Kolodny’s view is correct, and if concerns do give rise to deliberative boundaries, they do not do so by way of intermediating intentions.

<sup>19</sup> There is nothing exceptional about the example. The concerns of a normal human being are *numerous*, and actual and potential conflicts between the negative ends they each establish are a persistent, pervasive feature of our lives. If we were rationally required to cease *caring* about something whenever we found that we could not avoid acting against the negative end it establishes, our practical and affective lives would (insofar as we heeded this requirement) be greatly impoverished.

<sup>20</sup> Could this aiming consist in having a “policy” of not acting against my negative end? Policies, after all, are defeasible on particular occasions. But policies are still subject to rational requirements of consistency, even if these are weaker than the requirements governing simple intentions. To borrow an example from Bratman: If I have the policy of playing basketball each night and the policy of reading German each night, even though I know the former

(It might seem that, in arguing that our negative ends are not subject to rational requirements of consistency, because we can “act against” a negative end without incurring a charge of irrationality, I have undermined my claim that an agent’s concerns give rise to genuine *deliberative boundaries*. A promissory note: in §§4-6, I will show, in some detail, why this is the wrong conclusion to draw. A deliberative boundary may structure an agent’s thought even if it is defeasible and, in fact, defeated.)

If concerns need not involve negative intentions, what does explain the deliberative boundaries to which, as *Exclusion* claims, they are conceptually tied? It would be unsatisfying simply to assert as a brute, conceptual fact about concerns that they involve dispositions to exclude certain courses of action from deliberation. That would raise the question: why have concerns at all, rather than, say, *concerns\** – mental states that are exactly like concerns except that they have no conceptual connection to exclusionary dispositions?<sup>21</sup> Even if, as I shall claim in §§4-6, we ordinarily take agents whose practical thought is structured by deliberative boundaries that their concerns establish to be reasonable, why is this not an idiosyncratic, and perhaps ultimately irrational, fact about us and our pre-theoretical conception of reasonableness? We do not, after all, regard as reasonable the practical thought of agents whose thought is shaped by deliberative boundaries that their phobias establish. Why are concerns different?

I will not be in a position to give a satisfying answer to this question until §§7-8. There, I will in fact offer two different answers, corresponding to two different theoretical frameworks in which the account of deliberative attention that I am offering might be located. But we can sharpen the question now, and lay out some material these answers will make use of, by registering two minimal normative assumptions:

*Reason:* For any end, *e*, if an agent cares about *e* and it is not unreasonable of the agent to care about *e*, then (whether or not the agent intends to bring about or promote or sustain *e*) the agent has a *pro tanto* reason not take a course of action that is incompatible with *e*.<sup>22</sup>

---

will leave me too tired for the latter, I am subject to rational criticism (Bratman, *op. cit.*, p. 89).

What about a policy of giving *weight* to certain considerations in deliberation? (This proposal would identify this aspect of caring with “valuing,” as Bratman has attempted to understand that notion in several of the recent papers collected in his *Structures of Agency* (Oxford: Oxford University Press, 2007).) After all, I violate no requirement of consistency in giving weight to considerations pertaining to time with my family and also giving weight to considerations pertaining to my work. But precisely because there is no inconsistency here, this proposal does not explain the fact that negative ends bring with them *deliberative boundaries*.

<sup>21</sup> Thanks to Michael Bratman for pressing this question in conversation.

<sup>22</sup> In addition to the reasons that caring agents have *not* to take certain courses of action, caring agents also have positive reasons to take certain courses of action that help to sustain or realize ends they care about. If I care about understanding practical reasoning, for instance, and it is not unreasonable of me to do so, I may have a *pro tanto* reason to read a new book

*Strength*: For any ends, *e* and *f*, if an agent cares about *e* more deeply than she cares about *f* and it is not unreasonable of the agent to care more deeply about *e* than she cares about *f*, then the agent has a stronger *pro tanto* reason not to take a course of action that is incompatible with *e* than she has not to take a course of action that is incompatible with *f*.

Both assumptions are very weak, and are compatible with most theories of practical reason, precisely because they say nothing about what makes it reasonable or unreasonable of an agent to care about something, or to care more deeply about one thing than another. Perhaps concerns are answerable to no rational standards at all, so that there is nothing that it would be unreasonable of an agent to care about, or to care more deeply about than she cares about something else. In this case, *Reason* implies that an agent has a reason not to take a course of action incompatible with an end she cares about *just because* she cares about it, and *Strength* implies that the strength of this reason simply derives from the depth of the concern. Or perhaps, in order not to count as unreasonable, an agent's concerns must track independent facts about what is valuable, or about what is important for her own wellbeing, or something else. In this case, it might be these independent facts that at once give an agent reason to care about some end more or less deeply and give her a reason of some particular strength not to take a course of action incompatible with that end. Or perhaps the concerns that it is not unreasonable of an agent to have stand in some more complex relation to facts about value, or importance, or something else.

Weak as *Reason* and *Strength* are, there may be theories of practical reason with which they are incompatible. If so, then the normative ecumenicalism to which I have so far aspired must be restricted: the account I will offer in §§7-8 is committed to the assumption that *Reason* and *Strength* are true.

Granting that an agent has a *pro tanto* reason not to take a course of action incompatible with an end she cares about (and that it is not unreasonable of her to care about), why is it not an adequate deliberative response to this *pro tanto* reason for the agent simply to attach a quantum of disvalue to courses of action incompatible with that end, proportional to the strength of the reason she has not to take it – to be weighed in deliberation against the positive value of those courses of action, so that the net value of those courses of action may be compared to the net value of the other courses of action open to her? The task of §§7-8 is to bridge the gap between the claim that an agent's concerns give her reasons not to take some courses of action, and the very different claim that it is reasonable of her to *exclude* those courses of action from the range of options she entertains in deliberation. Before we can see the reasonableness of doing so, however, we need to see,

---

on the subject by a good philosopher. Such reasons do not figure directly in the rationale I will offer for *Exclusion*, however, so I do not discuss them in the main text.

in §§4-6, what doing so amounts to, and what sort of structure doing so gives to practical deliberation.

I have sought to establish the claim that an agent's concerns give rise to some deliberative boundaries "directly," without an intermediary role for intentions, by appealing to cases in which there is no positive intention associated with the agent's concern. But if the conceptual link between an agent's concerns and her deliberative boundaries is real, there is no reason to suppose that it does not hold quite generally. As I noted above, this is easily obscured because, in many cases, concerns also give rise to positive intentions. Because I care about finishing this article, I intend to finish it. And because intentions are subject to rational requirements of consistency, this positive intention will bring with it a deliberative boundary (I will not consider courses of action which are incompatible with finishing the article), and so a complementary negative end: I will aim not to act in a way incompatible with finishing the article. But if the claims of this section are correct, the deliberative boundary and the corresponding negative end are over-determined. This deliberative boundary, and the corresponding negative end, can *also* be ascribed to me simply in virtue of the concern: if my thought were not shaped by this deliberative boundary, this would be conceptual warrant for the conclusion that I do not, after all, care about finishing the article.

Thus we have overlapping explanations for the existence and stability of the deliberative boundaries that constrain our thought. Some of our deliberative boundaries are explained by intentions that are not backed by any concerns. Some of our deliberative boundaries are explained solely by concerns, with no corresponding intentions. And some deliberative boundaries are backed both by intentions and by concerns to which those intentions minister.

I claimed at the end of §2 that, in cases where some deliberative boundary issues from an intention, the stability of that boundary increases or decreases with the depth of the concerns by which that intention is "backed." I want now to suggest that this claim can be generalized, so that it covers cases in which a concern gives rise to a deliberative boundary directly, without intermediating intentions. That is, I want to suggest that the following claim is true:

*Stability:* For any concern of an agent and any deliberative boundary explained by that concern (whether directly, indirectly via intermediating intentions, or both), the stability of the boundary will increase or decrease with the depth of the concern.<sup>23</sup>

---

<sup>23</sup> This formulation is meant to allow that a single deliberative boundary may be explained by multiple concerns, and that *each* may contribute to its stability.

Of course, a concern might explain some deliberative boundary in a "deviant" or non-standard way, and in this sort of case, the depth of the concern might not contribute to the stability of the boundary. There is not much to be gained, except verbosity, in tidying up *Stability* so as to avoid such difficulties.

I will take *Stability* as a working-hypothesis in the coming sections, but I will not argue for it directly. Along with *Exclusion*, this hypothesis will, however, play a crucial role in motivating the Boundary-Driven Model of deliberative attention. Insofar as the Boundary-Driven Model proves to be descriptively adequate and normatively plausible, this should go some way toward confirming the truth of *Stability*.

#### 4. Deliberative boundaries and the temporal structure of deliberation: an example

As I have presented it so far, the idea that an agent's concerns explain the stability of the boundaries by which her deliberation is constrained may seem to suggest a quasi-mechanical picture, according to which the reasons an agent has to transgress some deliberative boundary are felt as psychological forces. On this picture, the deeper the concerns in which a deliberative boundary is rooted, the stronger the countervailing force that will be necessary in order to push through or over that boundary. But such a quasi-mechanical picture cannot accommodate the phenomena, for at least two reasons.

First, agents do not regard the deliberative boundaries that their concerns establish simply as brute, psychological facts about themselves. Agents take themselves to have good *reasons* to exclude certain courses of action from their deliberative options – reasons that they suppose may *justify* them in doing so.<sup>24</sup> If someone were to invite me to a movie on Friday, I would explain that I cannot go, because I intend to attend the department colloquium. If he were to press me to reconsider or abandon my intention, I would attempt to justify my unwillingness to do so by noting that (perhaps contrary to what he supposed) I in fact care about hearing this talk. If my interlocutor is able to understand my concern and to see it as not unreasonable, he will likely accept this justification. On the other hand, if he were able to convince me that it is unreasonable to care about hearing this talk (for instance, by informing me that the paper the speaker intends to give on Friday is the very same one I recently heard her give at a conference), he would rob me of the justification I have taken myself to have for refusing to reconsider my intention, and I would reconsider it. I will argue in §§7-8 that we are *right* to suppose that those of our concerns that are not unreasonable justify us in excluding certain courses of action from our deliberation. For now, however, I will rely on the weaker claim that we ordinarily *take* our concerns to justify us in excluding certain courses of action, at least when we do not suppose our own concerns to be unreasonable. It seems to us, rightly or wrongly, that the “force” with which reasonable concerns sustain the boundaries by which

---

<sup>24</sup> This does not mean that an agent's response to such a reason need be conscious or *deliberate*.

our deliberation is constrained is rational force, not the brute force of psychological compulsion.

The quasi-mechanical picture is also inappropriate for a second, and for our purposes more important, reason. It is not the case that, throughout a given deliberative episode, a given deliberative boundary either is or is not continuously “in force.” Deliberation is a temporally extended activity. And, often, it consists in distinct or at least distinguishable phases. At one stage in her deliberation, an agent may decide that she has reason to allow her deliberation to be constrained by some boundary, or she may simply allow it to constrain her deliberation, without pausing to consider it. At another stage, she may take up as a question for deliberation whether she ought to continue to allow that boundary to constrain her thought – whether she ought to continue to exclude from deliberation those courses of action that her concern disposes her to exclude. And having reached a decision, she may at another stage in her deliberation begin to consider seriously possibilities beyond the boundary that she had until then allowed to constrain her thought. The boundaries that constrain her thought may have, in other words, a *provisional* character. This fact, I will argue, is an important key to understanding the temporal structure of reasonable practical deliberation. In order to show this, I will present a detailed example. The length of the example will be justified by the richness of the structure it will be shown to have in this section and in §§5-6.

Irene is a teacher in an under-funded, inner-city school. She has worked there since she finished her degree in education at Harvard, twelve years ago. She is that rare, talented, idealistic teacher who has never “burned out,” despite the terribly difficult circumstances of her work. She knows that she makes a difference in many of her students’ lives, and she finds this extremely rewarding. She is, however, newly widowed, with two children of her own, ages nine and eleven. When her husband (also a teacher) was alive, the family had been able to rent an apartment in an expensive suburb, with excellent public schools. She knows that she cannot afford to continue to live there on her own, very modest income. She also knows, first-hand, the state of the schools in the poorer areas of her city where she could more easily afford to move her family. She not only worries about her children’s educational prospects there; she fears that they would be simply unable to navigate the tough social scene outside the classroom.

Irene considers how she might keep her children in a school where they can thrive. She wonders whether other school districts in her area offer a significantly better pay scale for teachers than her own, but a quick round of inquiries confirms her suspicion that what differences there are would not be enough to help. A friend of hers in Iowa City, Iowa tells her that teachers can afford to buy homes there, and that the schools are good. But Irene rules this possibility out, without further investigation: she is not willing to consider moving away from her ailing mother, whom she and the boys see several times per week, and who is rooted in the area. A classmate from Harvard, now fairly high up the ladder at a management consulting firm, suggests she could help Irene get an entry-level

job, which would pay more than three times what she earns as a teacher. But Irene cannot imagine devoting the working hours of her life to an aim no higher than making money. As an educator, she makes a difference in the world that she can see and feel, and this sustains her. She knows that if push comes to shove, and she can find no other way to keep her children in a good school, she might have to consider whether doing work she finds meaningful is more important to her than her children's school. She does not know how she would answer that question. But she is not ready, yet, to concede that push has come to shove.

Irene determines to seek a better-paying job as a school assistant principal. She had been offered such a job a year earlier, when her husband was alive, but had turned it down because it would have taken her out of the classroom. Now, however, she is willing to consider this possibility. During several weeks of inquiries, however, she finds that the only jobs available in the area have been filled. Still unwilling to contemplate a job whose only rewards are financial, she looks for possibilities she has missed. She finally hits upon the idea of looking for a job with a nonprofit foundation working on inner-city issues. After a few weeks of searching, she discovers just such a job, with a salary that will allow her to keep her family where it is. She is sorry to be leaving the classroom, with the immediate, tangible sense of helping people it provides her; but she is consoled by the thought that the ground-level expertise she brings to the foundation will help to change the circumstances that shape students' lives outside the classroom.

A number of deliberative boundaries, reflecting a number of Irene's concerns, help to shape Irene's thought. Some of these boundaries remain in place during the entire course of her deliberation, and do not seem to be open to question. Because she cares about living near her mother, for instance, Irene excludes from consideration the possibility of moving to another city – and seems unwilling to revisit the issue. But her deliberation is also shaped by boundaries that prove to be merely provisional in the context of her current deliberation. Because she cares deeply about working in education, Irene is initially unwilling to consider other sorts of jobs. After she has convinced herself that she has exhausted the possibilities in education, however, she proves willing to step over this boundary and consider other options. And while, during the course of her deliberation, she never considers the possibility of a job whose only rewards are financial, she acknowledges that she would be willing at least to entertain this possibility, too, if “push came to shove” and she had exhausted all the alternatives. (As it happens, she does not have to do this.)

It is tempting to suppose that if the putative “boundaries” I have pointed to figure in Irene's thought only provisionally, then it is a mistake to think of them as *boundaries* for her thought at all.<sup>25</sup> But this supposition misses the fact that Irene *treats* these *as* boundaries for her thought, excluding from her deliberation possibilities which lie beyond them, and that doing so gives

---

<sup>25</sup> This discussion begins to discharge the promissory obligation I incurred in the parenthetical paragraph on pp. 12-13.

her thought a distinctive structure, visible in the *course* her thought takes over time. Only once she is sure that she has exhausted the possibilities of continuing to work in education does Irene consider jobs outside of education. Likewise, Irene will not consider jobs that she does not find meaningful (such as the job as a management consultant offered by her friend) until she is convinced she has exhausted the possibilities of finding meaningful work. When it at first appears that there are no meaningful jobs that would solve the problem she confronts, she does not immediately proceed to consider the sorts of jobs that she would not find meaningful; nor does she even proceed to open the question whether she *should* consider such jobs. Rather, the fact that she can see no obvious possibilities within the deliberative boundary imposed by her concern with meaningful work spurs her to search, creatively, for unobvious possibilities within this boundary – continuing to exclude from consideration courses of action that lie beyond it as she does this. It is only when her imagination is in this way spurred into action that she hits upon the idea, which was not at first apparent to her, of working for a non-profit foundation.

## 5. Provisional boundaries and the recursive structure of deliberation: The Boundary-Driven Model

So far, I have claimed that Irene’s thought is constrained by various deliberative boundaries; that (many of) these boundaries may be explained by reference to her diverse concerns, in the manner that *Exclusion* implies; that some of these boundaries constrain her deliberation only “provisionally”; and that these boundaries (including those which are merely provisional) give her deliberation a distinctive structure, evident in the course her deliberation takes over time. I have not, however, explained the order in which she confronts these boundaries: why is she willing to consider options outside of some of them before she is willing to look beyond others? In fact, §2 and §3 afford us the resources with which to answer this question: I claimed there that the differing stabilities of an agent’s various deliberative boundaries are explained by the differing depths of the concerns by which those boundaries are explained. This was the thought that *Stability* aimed to capture. To link this thought to the account we have so far, we need only interpret the “stability” of a boundary in temporal terms, making the plausible assumption that an agent will consider stepping over a less stable boundary *before* she will consider stepping over a boundary that is more stable.

When we make this assumption, we have in *Exclusion* and *Stability* the resources for an idealized model of the temporal structure of deliberation, which I will call the Boundary-Driven Model. (I will postpone discussion of one of the respects in which the model is idealized until the end of this section.) I will set out the model in indicative terms, as a set of three claims about how an agent whom we would ordinarily regard as reasonable, A, will, in fact, deliberate. I set it out in indicative terms because, in establishing the

plausibility of the model, I mean to appeal to its descriptive adequacy: agents do in fact typically deliberate in (rough) accordance with its claims. I will try to show this by illustrating the model by reference to the example of Irene as I go, showing that the Boundary-Driven Model helps to render the structure of her deliberation intelligible. Because the model makes claims about how an agent whom we regard as *reasonable* will deliberate, however, it does not merely describe the deliberative dispositions that agents typically exhibit: it also articulates one part of our ordinary, pre-theoretical conception of reasonableness in practical thought. In §§7-8, I will ask what *justifies* the conception of reasonableness that the Boundary-Driven Model elaborates. First, however, I want to try to establish that it elaborates *our* conception of reasonableness.

Suppose that an agent, A, is prompted to deliberate because she cares about some end,  $e$ , (as Irene cares about keeping her children in a school in which they can thrive), and she recognizes that her circumstances pose a difficulty for realizing or sustaining  $e$ . (Of course, not all deliberation begins in this way; I will attempt to generalize this account to deliberation prompted in other ways in §6.) A normal agent will also bring to her deliberation a very large number of concerns with ends other than  $e$ . Some of those ends will not appear to A to have any relevance to her goal of realizing or sustaining  $e$ . Irene may care, for instance, about the health of her pet cat. Insofar as nothing suggests to Irene that she might find a solution to her problem if she were willing to compromise the health of her cat, her concern for her cat is unlikely to play any role in her deliberation. Let us denote those ends other than  $e$  that A cares about and that appear, from A's point of view, to be *relevant* to her deliberative problem,  $f_1, f_2, f_3, \dots$ . Her concerns with these ends will have varying depths. For simplicity of exposition, assume that A can easily rank all of these concerns in terms of their depths. This is certainly a fiction, but it is a fiction that relates to reality as an idealization: an agent can typically order *some* of her concerns in terms of their depths. She may, for instance, be sure that she cares more deeply about her own survival than about continuing to live in her current home, and that she cares more deeply about continuing to live in her current home than about what sort of car she drives. In keeping with this idealization, suppose that  $f_1$  is the end about which she cares least deeply, that she cares more deeply about  $f_2$ , more deeply still about  $f_3$ , and so on.

According to *Exclusion*, insofar as A cares about  $f_i$ , she will be disposed to exclude courses of action that she believes to be incompatible with  $f_i$  from the range of courses of action which she will entertain as options in practical deliberation; insofar as she cares about  $f_2$ , she will be disposed to exclude courses of action that she believes to be incompatible with  $f_2$ . And so on. If we make the plausible assumption that these dispositions agglomerate, we arrive at the first plank of the Boundary-Driven Model:

**B1** When *A* *begins* deliberating, her thought will be constrained by all of the deliberative boundaries that her diverse concerns establish. That is, as she looks for a way to bring about or sustain *e*, she will exclude courses of action that she believes to be incompatible with any of  $f_1, f_2, f_3, \dots$  from the range of courses of action that she will entertain as options.

The second plank of the Boundary-Driven Model is exegesis: it explains what is intended by the claim that an agent's thought is constrained by a deliberative boundary:

**B2** If there is no obvious course of action within the space that *A*'s deliberative boundaries define, this fact will serve as a goad to her imagination, and set her on the hunt for novel, unobvious possibilities within that limited space.

**B2** amounts to the claim that the dispositions to exclude that are established by *A*'s concerns are at least somewhat stable: *A* *continues* to exclude courses of action incompatible with the ends she cares about from her deliberation, even when she can see no solution to her problem that is compatible with those ends.

Suppose that *A* begins her deliberation in conformity to **B1** and **B2**, but that it becomes apparent to *A* – either immediately, or after an imaginative search for possibilities – that there really is no course of action available to her which will accomplish *e* without compromising any of her other concerns. If, as I have suggested, the “stability” of a deliberative boundary can be interpreted temporally, *A*'s deliberative attention will turn first to that deliberative boundary which is *least stable* – that is, according to *Stability*, to the deliberative boundary explained by the least deeply held of her relevant concerns. (According to the nomenclature I have adopted, this is  $f_i$ .) Thus, although Irene cares about continuing to work in the classroom as a teacher (recall that when her husband was alive, she had turned down the opportunity to work as an assistant principal), she cares less deeply about this than she cares about continuing to work in education, or about doing some kind of work she finds meaningful, or about continuing to live near her mother. And so this is the first deliberative boundary (of those I have mentioned) that she is willing to transgress.

As I described Irene when I laid out the example, she did not have to deliberate about whether it was more important, in her current circumstances, to keep her children in a good school or to continue working as a teacher. When it became evident to her that there was no teaching job that would solve her problem (subject to the constraints imposed by her other concerns), she proceeded immediately to look for work as an assistant principal, without deliberating about this question. Nonetheless, it is clear that her decision to step over the deliberative boundary that her concern with teaching imposed gave expression to an implicit decision that teaching was less important, in her current circumstances, than keeping her children in a

good school. Moreover, Irene recognizes in advance that a structurally analogous decision to step over the deliberative boundary imposed by her concern with doing work she finds meaningful would be harder to reach, and would require deliberation: if push came to shove, she would have to ask herself whether, in the circumstances she faces, it was more important to send her children to a good school than to work in a job she finds meaningful. A decision to take such a job would reflect an affirmative answer to this question.

In both of the cases mentioned above, we can summarize the result of Irene's deliberative or non-deliberative decision process, *after* the fact, by saying that she has concluded that, in the circumstances she faces, the former end is "more important," or perhaps "better," than the latter. (The phrase "in these circumstances" is not dispensable: the conclusion she reaches, in both cases, is likely to be wholly context-dependent.) These ways of describing Irene's conclusions are natural and illuminating because they fasten on the two ends that have focused her deliberation in each case. But for just this reason, they risk misleading, in at least two ways.

First, these natural ways of describing her deliberative conclusion might suggest that these are the only two ends she took into account in reaching her decision. That is unlikely to be the case. In deciding that teaching was less important, in her current circumstances, than keeping her children in a good school, for instance, Irene might have considered the fact that a job that did not involve nights and weekends spent grading papers and devising lesson-plans might allow her to spend more time with her children and her mother, and might allow her to devote more time to other causes and projects she cares about, such as writing poetry or volunteering with an organization advocating for environmental causes. If the non-teaching job available to her would indeed have these advantages, they would be among the deliberative "circumstances" in which she decided that teaching was less important, *in her current circumstances*, than keeping her children in a good school.<sup>26</sup>

Second, describing her decision as the conclusion that one end is (in her current circumstances) "more important" than another might seem to suggest that Irene must have reached her decision by introspecting, in order to determine which of these ends she cares about more deeply, or perhaps by scrutinizing the two ends in question, in order to determine which bears a greater quantity of "importance" or "value" – where these qualities attach to these ends in accordance with some antecedently determinate schedule or function. But the statement that she has found one end to be "more important" (or "better") than the other, in the circumstances she faces, may be entirely *verdictive*: it may describe the conclusion of her deliberation, without telling us anything at all about how she has reached it, or about what she takes to justify it.<sup>27</sup>

---

<sup>26</sup> Thanks to Jennifer Church for pressing me to clarify this point.

<sup>27</sup> Compare: A Kantian may decide that she must help someone who is injured by the side of the road, because every other maxim then available to her fails the test of the Categorical

How, then, does Irene reach the decision to pursue one end at the expense of a second, competing end? The answer to this question hinges on the answers to basic questions in the theory of practical reason – questions with regard to which the Boundary-Driven Model is agnostic. The Boundary-Driven Model is a model of deliberative attention. It explains the way in which an agent frames questions for deliberation, and it explains the answers to them that she will not entertain. To the extent that the model is correct, it comprises one part of an adequate theory of practical reason. But it is only a part. Indeed, I will argue in §§7-8 that the model is compatible with a range of competing theories of practical reason. All that we need now is the thought that some such more general theory is correct, that an agent may, at least on some occasions, reach a non-arbitrary decision to side with one end over another, and that, when this is possible, we may summarize her decision by saying that she has decided that one end is more important or better than the other (in her current circumstances).

With these considerations in place, we can make a first stab at a third plank of our model of the temporal course of a reasonable agent's deliberation. I said above that A will begin looking for a way to bring about  $e$  within the deliberative boundaries established by her concerns with  $f_1, f_2, f_3, \dots$  (Recall that  $f_1, f_2, f_3, \dots$  include only those ends that seem to A to be relevant, and that they are numbered in order of the increasing depths with which A holds them.) I said that if she cannot find a possibility within these confines, this will goad her imagination, and set her on the search for unobvious possibilities. Suppose, however, that she becomes convinced that there really is no solution to her problem within this limited space of possibilities. We can now add the following, provisional claim:

**B3\*** If A is convinced that there is no way to bring about  $e$  within the deliberative boundaries established by her concerns with  $f_1, f_2, f_3, \dots$ , she will turn her deliberative attention to the boundary imposed by  $f_i$ , the least deeply held of her relevant concerns. In order to decide whether to transgress this boundary, she will then ask herself (perhaps merely implicitly, or perhaps explicitly and deliberatively) whether  $e$  is more important, in her current circumstances, than  $f_i$ .

In answering this question, A will likely consider, either in general or more specifically, what sorts of possibilities lie on the far side of the boundary imposed by her concern with  $f_i$ . (The answer to this question is one of the circumstances in which she makes her circumstance-dependent decision.) In turning her attention to this deliberative boundary, then, she is already “stepping over” it in thought. Her question is whether she should transgress it in deed, by taking a course of action that would be incompatible with  $f_i$ .

---

Imperative. We may say, and she may say, that helping the person is the best course of action available to her in the circumstances. Clearly this verdictive use of “best” tells us nothing about how she reached her decision, nor about what she takes to justify it. (It certainly does not show that she is, after all, a closet maximizer.)

Suppose that A decides to step over the boundary established by her concern with  $f_1$ . The deliberative boundaries imposed by  $f_2, f_3, f_4...$  are still in place. A will then search for a way to bring about  $e$  within the confines of this newly enlarged, but still limited, space of possibilities. If she can find none, this should spur her imagination and set her on the search for unobvious possibilities, in accordance with **B2**. But eventually, if she concludes that she has exhausted the possibilities within this space, her deliberative situation will be structurally analogous to the one she was in a moment ago. Given that “stability” is to be interpreted temporally, she will now turn her attention to  $f_2$ , and ask herself whether  $e$  is more important, in this context, than  $f_2$ . And so on. We can represent this recursive structure in our model by replacing **B3\*** with:

**B3\*\*** For any  $n$ , if A’s deliberation is constrained by boundaries established by her concerns with  $f_n, f_{n+1}, f_{n+2}...$ , and A is convinced that there is no way to bring about  $e$  within these deliberative boundaries, she will turn her deliberative attention to the boundary imposed by  $f_n$ , the least deeply held of the relevant concerns by which her thought is currently constrained. In order to decide whether to transgress this boundary, she will then ask herself (perhaps merely implicitly, or perhaps explicitly and deliberately) whether  $e$  is more important, in her current circumstances, than  $f_n$ .

Eventually, as A’s thought progresses recursively through boundaries posed by ever more deeply held concerns, one of two things will happen. She may discover an acceptable solution to her problem. Or, she may reach a point at which the only deliberative boundaries that remain in place are boundaries that she is not willing to cross for the sake of a solution to the problem she faces. That is, she may decide that each of the remaining objects of concern,  $f_n, f_{n+1}, f_{n+2}...$ , is more important, in her current circumstances, than  $e$ . She may be forced to conclude, at this point, that there is no *acceptable* solution to the problem she faces – that is, no acceptable way to bring about  $e$ .

In the example of §4, Irene’s deliberation is constrained by her concerns with the following ends, listed in order of increasing depth: working as a teacher, working in the field of education, doing work that she finds meaningful, and continuing to live near her mother. No doubt her thought is constrained by boundaries imposed by many further, still deeper concerns – many of them so deep, in fact, that they never come into the foreground of her deliberation. Irene may care very deeply, for instance, about living a life that is compatible with her own moral ideals and her self-respect. Because of this concern, she may fail to consider (and, indeed, even to notice) possibilities, such as engaging in some form of criminal enterprise, which might otherwise provide very convenient solutions to her problem.<sup>28</sup> As I described

---

<sup>28</sup> In “Caring and Incapacity,” draft ms., I argue that understanding the way in which an agent’s concerns structure her deliberation helps us to make sense of the fact, to which Ber-

Irene's case, she found a solution to her problem after transgressing the first two of the deliberative boundaries that I noted. But suppose she had not found this solution, and so had proceeded to turn her attention to deliberative boundaries established by more deeply held concerns. She would then have had to ask herself whether it was more important to keep her children in a good school than to do work she finds meaningful. As I noted when I described her case, she does not know, in advance, how she will answer this question: it is a hard, deliberative problem. But suppose that she is able, somehow, to decide, that she is willing to take a job whose only rewards are financial. And suppose that, contrary to her expectations, she is not, in the end, able to obtain a job of this kind that will meet her financial requirements. She may then go on to decide – either deliberatively, or without having to deliberate – that she is *not* willing to move away from her mother for the sake of her children's school, and that she cannot see any other way (subject to the constraints imposed by other concerns on which she is not willing to compromise) to keep her children in a good school. At this point, she will give up the hunt for a way to realize this end and turn her attention to other deliberative problems raised by her situation – for instance, how to mitigate the effects on her children of the bad schools they will have to attend.

## 6. Extending the model to more open-ended deliberation

As I have so far set it out, the model assumes that A is prompted to deliberate by her felt need to bring about or sustain some particular end, *e* – as Irene is prompted to deliberate by her felt need to keep her children in a good school. It is time to jettison this assumption and so further generalize the model beyond purely instrumental or specificatory deliberation. Surely agents sometimes deliberate, not with an eye to bringing about already-determined ends, but rather with an eye to determining what their ends should be.<sup>29</sup> This may happen on a small or a large scale. On a Sunday morning, with no obligations in front of me, I may ask myself what I should do with my day. As she approaches graduation from college, Susan may ask herself what she should do with her life. Both of us may seek to answer our questions *deliberatively* – for instance, by considering the pros and cons of various alternatives.

In fact, such cases of open-ended deliberation appear to depart further from the model than they actually do. In both cases, we can specify, without

---

nard Williams and Harry Frankfurt have both drawn attention, that an agent may find certain courses of action “unthinkable.” And I argue in “Two Sides of Silencing,” *Philosophical Quarterly*, 55: 218 (2004), pp. 68-77, that this understanding can help us to untangle the true from the false claims that John McDowell makes about the way in which reasons may be “silenced,” and to explain the truth of those that are true.

<sup>29</sup> Someone might argue that even in such apparently open-ended deliberation, an agent is in fact always attempting to specify the end of a good or *eudaimon* life. I doubt that this is true. But its truth would not be incompatible with the strategy I adopt in the main text for generalizing B3\*\*.

much ingenuity or much insight into the minds of the agents concerned, formal ends that the agents seek, deliberatively, to find a way to specify and accomplish, and which can therefore play the role of  $e$  in the model. I seek something to do with my day. Susan seeks something to do with her life. Of course, attributing formal ends of this sort to an agent does not tell us very much about the agent or explain much about her deliberation. In order to understand the course of Susan's search for something to do with her life, you would need to know much more about the background of concerns and prior plans against which she deliberates. This background will not only help to explain the deliberative boundaries by which her thought is constrained; it will also help to explain the positive attractions she will see in various courses of action. It is because she cares about making enough money to continue the very comfortable lifestyle she grew up with, for instance, that a career in corporate law will strike Susan as a possibility especially worth considering.<sup>30</sup>

The unspecificity of the  $e$  that we ascribe to Susan does not threaten the basic structure of the model, but it does force us to add an important detail. Suppose that Susan begins deliberating, as the model implies, by searching for a course of action that is compatible with all of her concerns. And suppose she concludes that there is no such course of action. (Perhaps she cares not only about earning a large income, but also about having a job that will utilize her considerable artistic talents, and she can find no job that will accommodate both these concerns.) She will, the model claims, then turn her deliberative attention to one of the deliberative boundaries which has until now constrained her deliberation. Let us suppose that she turns her attention to the boundary imposed by her concern with using her artistic talent. (Let this be our  $f_1$ .) As the model describes it, she will ask herself whether  $f_1$  is more important, in this context, than  $e$ . But we have said that  $e$ , in this case, is simply "finding something to do with her life." *That* cannot plausibly be the end she compares with  $f_1$ . What she compares with  $f_1$ , in this case, must be some (or more than one) *specification* of  $e$  that acting against  $f_1$  would make possible – such as having a career in law, or having a career in investment banking.

In fact, this addition of detail to the model simply makes explicit what was already implicit in our claim that the comparison of  $e$  and  $f_n$  is *context-dependent*. When Irene sought to ask, with respect to various objects of concern which imposed constraints on her deliberation, whether they were more important than "keeping her children in a good school," this latter end needed to be understood, not abstractly, but as specified in light of the particular circumstances of her situation. In some other circumstances (where the good school in which she sought to keep her children was, for instance, so highly competitive as to impose its own stresses on her children), "keeping her children in a good school" would compare differently to the other

---

<sup>30</sup> This is an instance of the way in which an agent's concerns dispose her to regard certain considerations as positive *pro tanto* reasons. (See note 9 above.)

objects of her concerns. We can make this sort of context-dependency explicit by reformulating the third claim of the model one final time. We can replace **B3\*\*** with:

**B3** For any  $n$ , if A's deliberation is constrained by boundaries established by her concerns with  $f_n, f_{n+1}, f_{n+2}, \dots$ , and A is convinced that there is no way to bring about  $e$  within these deliberative boundaries, she will turn her deliberative attention to the boundary imposed by  $f_n$ , the least deeply held of the relevant concerns by which her thought is currently constrained. In order to decide whether to transgress this boundary, she will then ask herself (perhaps merely implicitly, or perhaps explicitly and deliberatively) whether  $e$  (as she could realize it by transgressing the boundary established by  $f_n$ ) is more important, in her current circumstances, than  $f_n$ .

**B1, B2 and B3** together constitute the Boundary-Driven Model of deliberative attention. I have tried to motivate the model's claims to descriptive accuracy in two ways. I have tried to render plausible the twin claims, *Exclusion* and *Stability*, which the model elaborates and modestly extends. (However, I noted when I presented them that one source of support for those claims lies in the plausibility of the larger model of deliberative attention in which they figure. So, those claims and the Boundary-Driven Model must in the end stand or fall together.) And have I tried to show that the model illuminates the temporal structure of an instance of deliberation (Irene's) that is detailed enough to be lifelike, and that I think we can recognize as a paradigm of reasonableness as we pre-theoretically understand it.<sup>31</sup>

---

<sup>31</sup>The descriptive accuracy of the model also admits of a less direct sort of confirmation in the sense that it makes of two (related) facts about ordinary moral psychology. The first of these is the fact, to which Bernard Williams and Harry Frankfurt have both drawn attention, that an agent may find certain courses of action "unthinkable." In my "Caring and Incapacity," I argue that the phenomenon of unthinkability is a consequence of the boundary-driven structure of deliberation. As I have described Irene, the possibilities of moving away from her mother in order to move her children to a better school, or of engaging in some form of criminal enterprise, are "unthinkable" for her. These possibilities are likely, in fact, simply never to occur to her. More importantly, if they do occur to her (or if someone proposes them to her), she will immediately rule them out. The Boundary-Driven Model explains why she does so and why this is revelatory of her character: these possibilities are unthinkable for her because they are incompatible with ends about which she cares deeply; and they are revelatory of her character because her character is reflected in her concerns. That she cares about living near her mother, and about living in a manner compatible with her various moral ideals, are deep facts about the sort of person she is.

The second fact about ordinary moral psychology that the Boundary-Driven Model renders accessible to theory is this: As they figure in the practical thought of a morally decent agent, some moral considerations seem to have each of two properties that moral theorists have often supposed cannot fit together – they seem to have a genuinely *deontic* import, on the one hand, but to be *non-absolute* or *overridable* in their import, on the other. It has seemed to many moral theorists that if some moral consideration can be overridden in the thought of a morally decent agent, then the agent must implicitly be operating a consequentialist calculus of value and disvalue, so that the consideration does not play a genuinely deontic role in her thought. The Boundary-Driven Model, by contrast, locates a place in the practical

Like any model, the Boundary-Driven Model is idealized. The model suggests, for instance, that deliberation consists in a series of temporally discrete phases. But real deliberation is unlikely to be so tidily structured. It is plausible to suppose that even as Irene is searching for work as a school assistant principal, for example, she is also considering, however provisionally, what possibilities might be available to her if this search turns out to be fruitless. (In fact, given the length of time that the search requires, this would be a prudent use of her cognitive resources.) In this and no doubt in other respects, the Boundary-Driven Model falls well short of descriptive accuracy. However, so long as the model remains *sufficiently* lifelike, this need not be a blow to its aspirations. The model aims to provide a *sketch* of deliberation that we regard as reasonable, which brings some structurally important lineaments into special prominence. It will serve its purpose if, when we hold this sketch up to real cases of deliberation which we regard as reasonable, we can make out structural features of the real cases which would otherwise be obscured by the details which surround them. A sketch can serve this purpose even if, when the sketch is superimposed on the reality it portrays, no line of the sketch perfectly matches the lineament it represents.

The aspirations I have announced for the model, however, go beyond describing how agents whom we regard as reasonable in fact deliberate: the model aims to articulate one part of our ordinary, pre-theoretical conception of *reasonableness* in practical thought. It will be descriptively adequate, then, only insofar as we not only expect reasonable agents to deliberate in (rough) accordance with its claims, as I have tried to show that we do, but *also* insofar as we regard agents who do *not* deliberate in rough accordance with its claims as, at least *ceteris paribus*, open to rational criticism. The most obvious place to turn, then, in order to make a more complete case for the descriptive adequacy of the Boundary-Driven Model, would be toward instances of deliberation which fail to conform to it – and which we ought, according to the

---

thought of a morally decent agent for considerations whose import is genuinely deontic but is not absolute. Insofar as a morally decent agent cares about acting honestly, for instance, merely attaching a quantum of disvalue to a dishonest course of action, to be weighed against whatever value the action might achieve, will not seem to her to be *an adequate deliberative response* to the consideration that it is dishonest. Rather, as the Boundary-Driven Model predicts, she will take the consideration that the action is dishonest to give her reason to rule out this course of action. Moreover, when she can find no honest course of action that will allow her to achieve whatever she is trying to achieve, her response will still not be to weigh the disvalue of acting dishonestly against the value that she might achieve by doing so. Rather, according to **B2**, this will goad her imaginative and creative capacities, and prompt her to search for an unobvious possibility that does not involve acting dishonestly. On the other hand, as **B3** allows, when a morally decent agent becomes convinced that there really is no honest course of action available to her, she may decide to override the consideration that some course of action is dishonest, if the end she is trying to accomplish (preventing some great injustice, for instance) seems to her more important in the circumstances she faces. (Insofar as this decision is context-specific, it need not reflect any *general*, systematic ordering of the importance of acting honestly and combating injustice, of the sort a consequentialist might hope to discover.)

model, to regard as instances of unreasonableness. However, while I relied on the reader's intuitions in response to examples in order to argue for the truth of *Exclusion* and *Stability*, this sort of argument is less likely to convince here, where the question at issue is whether a given pattern of deliberation is reasonable or unreasonable. Discussions of standard models of rationality reveal the considerable extent to which our intuitions are malleable and inevitably shaped by theory, or just by ground-level argument. What we need, then, is to show *why* it is unreasonable of an agent to deliberate in a way that is not boundary-driven. In order to confirm the descriptive adequacy of the model, in other words, we need to consider its normative plausibility.

### 7. Justifying boundary-driven deliberation in the context of maximizing theories of practical reason

In §3, I accepted *Reason*, the minimal normative assumption that if an agent cares about some end and it is not unreasonable of her to care about that end, then she has a *pro tanto* reason not to take a course of action incompatible with that end. And I accepted *Strength*, the assumption that the strength of this reason increases as the depth of her concern increases. I left pending the question of how to bridge the gap between the reasons that *Reason* and *Strength* identify and the dispositions that *Exclusion* and *Stability* describe. The Boundary-Driven Model extends *Exclusion* and *Stability*. So the question before us is: how do we get from *Reason* and *Strength* to the claim that an agent has reason to deliberate in accordance with the Boundary-Driven Model?

Suppose that some sort of maximizing conception of practical rationality is correct. Suppose, that is, that there is some index of value that successful practical choice maximizes, or with respect to which it satisfies. Perhaps a rational agent brings to her deliberative situation a complete set of preferences, conforming to the axioms of utility theory, and the best course of action is the one that maximizes the agent's expected utility. Or perhaps the best course of action maximizes some index of objective value, and this index is a function of multiple, distinct objective values realized by each course of action.

In the context of a maximizing conception of practical reason, *Reason* and *Strength* amount to the claim that a course of action that is incompatible with some end that an agent cares about and is not unreasonable to care about involves a *disvalue* that will register on whatever index of value it is that the agent ought to maximize, and the claim that the deeper the concern at issue (assuming the depth of the concern is not unreasonable), the greater the disvalue involved. In this theoretical context, it is clear that a creature with unlimited cognitive capacities, for whom cognition consumed no time or energy, would have *no* reason to deliberate in accordance with the Boundary-Driven Model. Such a creature could survey the whole landscape of possibili-

ties at once, without missing any options, and then choose that option which was “best” according to the relevant index of value.<sup>32</sup>

But *we* might have a reason to deliberate in accordance with the model. For as theorists of “bounded rationality” have long emphasized, our cognitive capacities are finite.<sup>33</sup> If we were to attempt to choose among a “wide-open” field of options, we would be overwhelmed by the unlimited possibilities. Not only would the task of considering the many advantages and disadvantages of the options that were apparent to us be beyond our powers, we would be liable to miss unobvious solutions to our practical problems. Recursively structured deliberation in accordance with the Boundary-Driven Model is useful to us, then, because it goads our imagination and focuses our attention as we search for options, directing us toward possibilities which have the advantage, at least, of being incompatible with few and relatively shallow concerns and so – if we interpret *Reason* and *Strength* in the light of a maximizing conception – involving relatively little disvalue.

(Once our deliberative attention has been guided to some particular deliberative boundary, and we seek to answer the narrowly framed question whether we ought to cross this boundary for the sake of a solution to our deliberative problem, we can employ straightforward maximizing rationality: we can then ask whether the positive value that any of the solutions which crossing this boundary would make available is greater than the disvalue it involves, and we can ask which of these solutions has the greatest net value.)

On this picture, there is no guarantee that the option we settle upon in boundary-driven deliberation will be the best option – the option that a cognitively unlimited deliberator surveying an unbounded landscape of possibilities would choose. It is possible that there is some option, beyond one of the deliberative boundaries by which our thought was still constrained when we settled on a solution, which would in fact score higher on the relevant index of value than would the option we have settled upon. But boundary-driven deliberation may nonetheless help us to find an option that is good enough – and perhaps the best that we, as cognitively finite creatures, can find. Boundary-driven deliberation, on this picture, is a *heuristic device*, of the sort that theorists of “bounded rationality” have long emphasized.

I do not want to claim that boundary-driven deliberation is the only heuristic device we could employ to cope with the indefinitely large number of options that are typically available to us; and I certainly have no argument that it is the best device we could possibly employ. But if the claims I made in §§4-6 about the descriptive adequacy of the Boundary-Driven Model are

---

<sup>32</sup> A caveat: in order to avoid the difficulties for coordination that Buridan cases present, the deliberation of even this non-limited deliberator would need to be constrained by her prior *plans*.

<sup>33</sup> For pioneering work in the theory of bounded rationality, see Herbert Simon, *Reason in Human Affairs* (Stanford: Stanford University Press, 1983). For fascinating recent work in this tradition, see Gerd Gigerenzer, Peter Todd and the ABC Research Group, *Simple Heuristics That Make Us Smart* (Oxford: Oxford University Press, 1999).

correct, boundary-driven deliberation is one of the devices we actually use. Insofar as it offers a way for finite creatures like us to navigate an indefinitely large field of possible courses of action, and insofar as it often offers a route to a satisficing solution to our practical problem, this is surely a *pro tanto* reason to employ it. That this reason is sufficient to *justify* us in employing it awaits disconfirmation by anybody who can propose a *better* device that we might use in its place.<sup>34</sup>

### **8. Justifying boundary-driven deliberation in an incommensurabilist framework**

Understanding and justifying boundary-driven deliberation as a heuristic device makes sense in a theoretical framework in which there is a determinate standard, conceptually independent from the Boundary-Driven Model, by which to measure the value or overall choice-worthiness of every potential course of action. But this understanding of boundary-driven deliberation is not available to a philosopher who believes that the many values which the diverse courses of action open to an agent realize are incommensurable, and so that there is no single index of value that successful practical choice maximizes (or with respect to which it satisfices).<sup>35</sup> Absent such an index of value, we simply lack a determinate, independent standard: a standard in the light of which a cognitively unlimited agent who did not deliberate in accordance with the Boundary-Driven Model could survey a wide-open field of possibilities and judge that some course of action, from among the indefinitely large number of courses of action open to her (realizing large numbers of incommensurable values), is “best.”<sup>36</sup>

---

<sup>34</sup> Or who can make the difficult case that we’d be better off using no device at all and simply letting our deliberative attention flow in some arbitrary way.

<sup>35</sup> My use of “value” here is meant to be maximally vague, and so to import a minimum of theoretical commitment. I use the term to refer to whatever properties of persons, works of art, states of affairs, etc. an agent ought to regard as sources of practical reasons. This use is compatible with the possibility that there are objective or intersubjective values to which a deliberating agent ought to respond. But it is also compatible with the possibility that some form of internalism about reasons is true, and all value is based, for instance, in an agent’s antecedent desire.

<sup>36</sup> Kantians or Aristotelians might complain that I assume, unwarrantedly, that maximizing conceptions of practical rationality are the only ones going. I do not assume this. Rather, I assume that, however true their respective claims about practical rationality, neither a Kantian nor an Aristotelian theory provides us with a standard for the evaluation of potential courses of action which is both generally applicable and conceptually independent from the Boundary-Driven Model.

Kant offers us a standard, conformity to the Categorical Imperative, by which to divide actions into the impermissible, the permissible and (when every alternative to a given course of action is impermissible) the obligatory. (Perhaps the doctrine of imperfect duties allows us add a category of the “praiseworthy” within the larger category of the permissible.) But in those (frequent) cases in which no course of action is obligatory and a very great number are permissible (and even “praiseworthy”), a Kantian conception of practical reason provides us

I do not mean to claim that, if we suppose that values are plural and incommensurable, we must also suppose that an agent can never make a justified, non-arbitrary decision that one course of action is more choice-worthy than another, all things considered. In fact, various philosophers have suggested deliberative devices by which agents might make justified choices in the face of plural, incommensurable values. Elijah Millgram has argued, for instance, that, when an agent specifies the apparently conflicting values she faces carefully enough, she might find that the demands that they make on her *in the situation at hand* are not, after all, in conflict, or that they are in conflict but that their importance can, *in the situation at hand* (but not more generally), be weighed on a single scale.<sup>37</sup> Elizabeth Anderson has suggested that in some cases an agent might be able to appeal to some relevant principle of moral obligation that obviates the need for “weighing” competing goods in that context.<sup>38</sup> Aurel Kolnai has suggested that an agent might be able to accommodate each of several conflicting values by adopting complex, temporally extended plans: I can attend my child’s piano recital today, and still get done the work I need to do, if I work late at the office tomorrow night.<sup>39</sup> And Charles Taylor has argued that an agent might be able to reach a justified choice between courses of action involving incommensurable goods by recourse to considerations about the effects which the choices open to her would have upon her own integrity.<sup>40</sup> Even in the absence of deliberative resources such as these, moreover, an agent may be able to make a non-arbitrary decision in the face of incommensurable values. After doing her best to imagine herself taking each of the courses of action open to her, for instance, she might simply reach the decision – perhaps backed by further

---

no standard by which to evaluate the many permissible (or praiseworthy) courses of action open to an agent.

Aristotelians will claim that the *phronimos* provides a standard by which to evaluate actions. The best course of action is just the one that the practically wise person would choose, bringing to bear as she does a full array of ethical and intellectual virtues. I do not disagree. But if the arguments of this paper have been persuasive, and the norms involved in Boundary-Driven Model are constitutively bound up in our conception of practical reasonableness, then the *phronimos* is someone who, *inter many alia*, engages in boundary-driven deliberation. (In fact, I argue in “Two Sides of Silencing” that the fact that the deliberation of the *phronimos* is boundary-driven *explains* the (partial) truth in John McDowell’s claims that the virtue of a *phronimos* “silences” certain considerations in his thought.) And so “what the *phronimos* would choose” is not a standard for evaluating acts that is conceptually independent from the Boundary-Driven Model.

<sup>37</sup> Elijah Millgram, “Incommensurability and Practical Reasoning,” in his *Ethics Done Right: Practical Reasoning as a Foundation for Moral Theory* (Cambridge: Cambridge University Press, 2005), pp. 273-294.

<sup>38</sup> Elizabeth Anderson, “Practical Reason and Incommensurable Goods” in R. Chang (ed.), *Incommensurability, Incomparability, and Practical Reason* (Cambridge, Mass: Harvard University Press, 1997), pp. 90-109.

<sup>39</sup> Aurel Kolnai, “Deliberation is of Ends” in his *Ethics, Value, and Reality* (London: Athlone Press, 1977), pp. 44-62. (The example is my own.)

<sup>40</sup> Charles Taylor, “Leading a Life,” in Chang, *op. cit.*, pp. 170-183.

reasons, perhaps not – that she can live with one of the options open to her, but that she could not live with the others.

Two things are important to notice about these and other, similar deliberative devices. The first is that none represents a *general* strategy for navigating a landscape of incommensurable values. Rather, these devices together comprise a toolbox of deliberative resources for a skilled deliberator, one of which might work in one deliberative context, another in another.<sup>41</sup> In any given deliberative situation, moreover, there is no guarantee that any of these deliberative devices (or any other) will allow the agent to reach a non-arbitrary solution to her deliberative problem. Whether a device is available that will allow an agent to do this will be, at least in part, a matter of luck.

The second thing to notice about all of these deliberative devices is the sort of deliberative problem to which they in fact provide solutions. None of these devices appears to offer an agent much help in answering a practical question of the form “which of the indefinitely many courses of action now available to me (realizing many diverse, incommensurable values) would be best?” Rather, each of these devices has *application* to some more narrowly framed deliberative problem – a problem in which an agent seeks to choose between two, or perhaps a few, obviously salient alternatives which realize apparently incommensurable values. So far as I am aware, the examples with which these devices are illustrated by their authors all, in fact, involve practical problems of this narrower form. This is, in fact, a testament to the verisimilitude of the examples: that real agents do not frame their practical problems as choices among wide-open fields of possibilities is precisely what the Boundary-Driven Model explains and predicts.

These considerations point to one part of a rationale for the Boundary-Driven Model in the theoretical framework provided by an incommensuralist, non-maximizing conception of practical rationality: boundary-driven deliberation plays a crucial role in making a justified, non-arbitrary choice in the face of plural and incommensurable values possible, because an agent who deliberates in accordance with the Boundary-Driven Model will frame the deliberative questions she asks herself narrowly enough that deliberative devices of the sort mentioned above might have application.

Suppose, for instance, that an agent is unable to find a course of action compatible with all of her concerns. Insofar as she deliberates in accordance with **B3**, her response to this fact is not simply to turn to the wide-open field of incommensurable possibilities, asking, “What is the best thing I could do?” Rather, she turns her deliberative attention to some *one* of the deliberative boundaries by which her thought has until then been constrained (or, if we suppose that this is over-idealized, to some manageable few), and asks of

---

<sup>41</sup> Some of these authors appear to view the device he or she offers as a general strategy, but it is dubious whether any of these devices could succeed in this role. (The existence of several different devices, each of which might help an agent to choose in the face of incommensurables, is itself evidence that none of these approaches can be relied upon to work in every circumstance.)

it a very limited question: in these circumstances, is the end I am seeking to accomplish by solving my practical problem more important than the end that I would “act against” in stepping over this boundary?

In order to reach a well-grounded answer to this question, she will need to find a way to make a choice involving incommensurable values. The Boundary-Driven Model does not itself provide her with a way to answer the question it has framed; and there is no guarantee, in advance, that she will be able to hit upon any device that will allow her to make a principled choice. She might, in the end, simply decide that the resources of reason have run out, that neither choice would be more reasonable than the other, and so that she must simply make an existential choice by plumping for one or the other. But although she might end up in this situation, she might not. If we are willing to accept an incommensurabilist theoretical framework, it seems that we should also be willing to accept that, by employing deliberative devices of the sorts I mentioned above, agents do, sometimes, find ways to make well-grounded, non-arbitrary choices among a few salient options involving a few salient, incommensurable values. And what we have seen is that boundary-driven deliberation makes success at least *possible*. It does so by narrowing the agent’s problem from “what, of the indefinitely many courses of action available to me, involving many incommensurable values, would it be most reasonable of me to choose?” (a question to which there may be no determinate answer, even for a cognitively unlimited agent) to “would it be reasonable of me to choose one of *these* courses of action (compromising, as they do, an end that I care about) as a way to accomplish my goal?”

If an agent manages in this way to arrive at a reasoned, non-arbitrary choice of a single course of action, we may say (and she might say) that boundary-driven deliberation (together with whatever deliberative devices have allowed her to make her final, narrowly framed decision and any preliminary, narrowly framed decisions along the way) has enabled her to discover “the best course of action available to her.” In saying this, however, we appeal to no *independent* standard to which a cognitively unlimited agent whose deliberation was not boundary-driven might have looked. Rather, our use of “best” is verdictive: it simply registers, after the fact, that this is the unique course of action on which an agent deliberating in a manner that is responsive to the diverse norms that jointly constitute our conception of reasonableness has been able to settle. If the claims I made in §§4-6 about the descriptive adequacy of the Boundary-Driven Model are correct, then the requirements of the model are central among the norms to which the deliberation of such an agent will have been responsive.

In the context of an incommensurabilist view of value, then, Boundary-Driven deliberation has at least this much to be said for it: together with whatever further deliberative devices complement it, it allows us, sometimes, to reason our way to a well-grounded, non-arbitrary choice in the face of incommensurable value – and so to give a determinate, verdictive, meaning to the phrase “the best course of action available to me.” Boundary-driven

deliberation frees us, then, from viewing all of our choices in the face of incommensurables as ungrounded, existential assertions of the will – a view that it would be hard to escape from, if we had always to choose in the face of a wide-open field of possibilities realizing diverse, incommensurable values.<sup>42</sup> Insofar as this would be a profoundly uncomfortable view of our own choices, from the trivial to the most significant, attended by all of the risks of angst and dread that writers in the existentialist tradition have chronicled, the possibility that boundary-driven deliberation affords us of reasoning our way to non-arbitrary choices in the face of incommensurables is itself a *positive good* for us.

This good would be illusory, however, if the requirements of the Boundary-Driven Model, and so the standard for well-grounded choice that the Boundary-Driven Model partly defines, were *themselves* arbitrary. If this were the case, boundary-driven deliberation would seem to be little more than a device for dodging the heavy burden of our existential freedom. But the Boundary-Driven Model and the pre-theoretical conception of reasonableness it embodies are not arbitrary.

To see that the Boundary-Driven Model is not arbitrary in an incommensurabilist framework, we need to revisit the twin, minimal normative assumptions, *Reason* and *Strength*. In the context of a maximizing theory of practical reason, I said that *Reason* amounts to the claim that a course of action that is incompatible with some end an agent cares about involves a disvalue that will register on whatever index of value it is that the agent ought to maximize. While an incommensurabilist recognizes no such index of overall value, an incommensurabilist surely *can* nonetheless acknowledge that a course of action incompatible with an end that an agent cares about (and is not unreasonable to care about) involves a *loss* from the perspective of the deliberating agent – a loss that she ought to take into account in her deliberation. The incommensurabilist, moreover, will note that most of an agent's concerns with the ends she cares about – ends such as the preservation of a beloved old-growth forest near her childhood home, the success of her career, the continuance of affirmative action in American university admissions, and the moral development of her teenage son – are mutually irreducible. And so the incommensurabilist will insist that the loss of one of these ends – even when it is a loss that the agent has most reason, all things considered, to accept – is, typically, a loss for which the agent cannot be compensated *in kind*.<sup>43</sup>

---

<sup>42</sup> This picture of an agent, facing a wide-open field of “eligible” possibilities that realize incommensurable values, and choosing freely among them by an a-rational exercise of will, is very close to the picture that Joseph Raz endorses in the essays collected in his *Engaging Reason* (Oxford: Oxford University Press, 1999).

<sup>43</sup> For the notion of “compensation in kind,” see David Wiggins, “Weakness of Will, Commensurability, and the Objects of Deliberation and Desire,” in *Needs, Value, Truth*, 3<sup>rd</sup> edition (Oxford: Oxford University Press, 1998), pp. 239-267, at pp. 259-60.

This claim by the incommensurabilist motivates a simple but powerful rationale for the deliberative disposition that *Exclusion* describes: if an agent cannot respond in deliberation to the prospective loss of an end she cares about simply by subtracting its disvalue from the value of the potential gains by which that loss might be compensated, in order to arrive at an estimation of the overall value of the course of action in question, then a reasonable deliberative response to such a prospective loss will simply be *to aim to avoid it* – that is, to treat the end in question as a negative end, which imposes deliberative boundaries on her thought. This is just the deliberative response that *Exclusion* and its elaboration in **B1** and **B2** describe.

What about *Stability* (and **B3**)? Often, of course, no course of action (or inaction) is available to an agent that does not involve the loss of something she cares about. A commensurabilist supposes that, where this is the case, the agent ought to minimize her net loss, by taking the course of action that scores best on the relevant index of value (so far as her cognitive limitations permit her to discover it). For the incommensurabilist, no such index of the overall choice-worthiness of courses of action is available. But that does not prevent an incommensurabilist from recognizing that a course of action that is incompatible with an end an agent cares more deeply about involves, as such, a *greater loss* than a course of action that is incompatible with an end that the agent cares less deeply about. Even in an incommensurabilist framework, *Strength* implies at least this much. The fact that course of action A involves a lesser loss than B *in this respect* constitutes at least one determinate advantage of A over B, whatever further advantages or disadvantages A and B might be found, upon deliberation, to have.

The fact that A is incompatible with a less deep concern than B does not entail that A is better than B all things considered. Indeed, it does not entail that it is possible to make an all-things-considered comparison of the choice-worthiness of A and B at all. Nonetheless, we can see that the requirement that our deliberative attention be guided by sensitivity to *this* particular advantage of A over B is not arbitrary. For in contrast to most other advantages A may have, this is an advantage of A over B that typically requires no deliberative attention to A and B to discover. An agent's concerns can direct her deliberative attention "from the background," in the manner that *Stability* and **B3** describe, toward courses of action which have this advantage. And such a non-deliberative way of answering the question of where to direct our deliberative attention is just what cognitively limited creatures like us need, if we are to avoid directing it arbitrarily, on the one hand, and are to avoid the regress of deliberating about where to direct our deliberative attention, on the other.

Understood within an incommensurabilist framework, boundary-driven deliberation cannot be justified instrumentally: there is no conceptually independent standard of success in practical deliberation, to which boundary-driven deliberation helps us to approximate. But I have claimed that sometimes, with the help of deliberative devices beyond the scope of the Bound-

ary-Driven Model, boundary-driven deliberation makes it possible for an agent to reach a justified, non-arbitrary decision about what to do – and so to give a determinate, verdictive meaning to the phrase “the best course of action available to me.” And I have argued that the requirements of the Boundary-Driven Model are not themselves arbitrary. Insofar as these claims are correct, I think that they make it intelligible why the norms articulated in the Boundary-Driven Model have a place among the norms which jointly constitute our conception of reasonableness in practical thought; and I think they give us sufficient grounds to endorse those norms.

\*

When an agent deliberates about what to do, her attention cannot be everywhere at once. Nor, if she is reasonable, is her deliberative attention directed arbitrarily. The Boundary-Driven Model aims to explain the path that a reasonable agent’s deliberative attention takes, as her deliberation unfolds over time, and to illuminate the temporal structure of ordinary instances of practical deliberation that we can recognize as reasonable. I have begun to argue at length elsewhere that the model is theoretically fecund in other ways: it helps to explain features of ordinary moral thought which have proved puzzling for moral psychologists and moral theorists, and to reveal ways in which moral thought is continuous with other practical thought. And boundary-driven deliberation of the sort the model describes can be justified within the theoretical frameworks provided by both maximizing and non-maximizing conceptions of practical rationality. These are reasons to believe that the limits that our concerns establish for our practical thought are reasonable bounds.

Jeffrey Seidman  
Vassar College  
Department of Philosophy  
[JeSeidman@vassar.edu](mailto:JeSeidman@vassar.edu)