

## ON THE RELATION BETWEEN WANTING AND WILLING\*

Tamar Schapiro  
Stanford University

### 1. The Problem

A traditional conception of human agency holds that action is a form of self-governance, an exercise of our natural authority over ourselves. Closely associated with this tradition is the thought that governing ourselves involves keeping a certain intrapersonal order intact, maintaining “justice” in the “soul.”<sup>1</sup> Different theories use different terms to specify the elements that we are to keep in order, but the common idea is that we have (at least) two distinct motivational capacities, or two distinct types of motives, such that the job of one is to govern and regulate the other. In Plato’s terms, “reason” is to govern “appetite.” In Kant’s terms, the rational “will” is to govern “inclination.” In more contemporary philosophy, “intentions” have the “functional role” of regulating “desires.” Obviously there are deep differences between these conceptual vocabularies. But for the purposes of this paper I am going to abstract as far as possible from those differences in order to describe and address a very general problem. Given the assumption that we ought, in fact, to maintain a certain hierarchy among our motives, what is the basis of this hierarchy? If there is a “constitution” in the soul, what justifies its distribution of authority?

I am going to approach this problem by exploring an analogy. The analogy is between adult authority over children and the authority of the higher motives over the lower. My main claim is that close attention to the details of this analogy helps us to see more clearly 1) the structure of the problem, 2) the space of possible solutions, and 3) the costs and benefits of adopting one solution over another. Although it will be clear that I favor one solution over the other, I will only be able to offer a limited defense of that solution here.

For the purposes of this exploration, I will need to make a few substantive but fairly abstract assumptions about our motivational structure. I will assume that whatever the governing motivational capacity or type of motive is, it is distinctively both rational and volitional.<sup>2</sup> That is, I assume that the contrast between higher and lower motivation is a contrast between rational and nonrational motivation, and that it is, at the same time, a contrast between active willing and passive inclining.<sup>3</sup> I will leave it open what rationality amounts to, but I will assume that to exercise rationality is to deliberate about the grounds for action so as to come to a practical conclusion. And I will think of this practical conclusion as volitional, as an exercise of the agent's natural authority over herself. To have arrived at a practical conclusion, I am assuming, is to have generated, through deliberation, a motive that has the authority to speak for you and that takes precedence over your lower motives in cases of conflict. For my purposes, we can use the terms "volition," "decision," "choice," and "commitment" interchangeably to refer to higher motives.<sup>4,5</sup>

As for the type of motivation that is governed, I will assume that it functions independently of the higher, and so is both distinctively nonrational and distinctively passive. What I mean by saying the lower form of motivation is "nonrational" is that our appetites and inclinations are not arrived at as conclusions of practical deliberation. They are intelligible responses to the way we see our environment, but they arise in us spontaneously and are not products of reasoning.<sup>6</sup> What I mean by saying they are "passive" forms of motivation is they are not under our direct volitional control. For example, you can't will yourself to want to eat a nutritious salad, even if you are convinced that eating a nutritious salad is the best thing to do. Nor can you rid yourself of a craving for ice cream just by deciding you ought not to crave ice cream. In both cases, you have to work on your appetites indirectly, by developing new habits and using your imagination to forge new associative links. The same is true with motives that are not physiological appetites. You know you ought to grade those papers, but you do not feel like grading them. And you cannot make yourself feel like grading them just by deciding you ought to grade them, or by deciding you ought to feel like grading them. Henceforth I will use the term "inclinations" to refer to these lower motives.

Roughly put, then, the authority relation I am interested in is between the part of us that wills and the part of us that wants.<sup>7</sup> Why should the part that deliberates and chooses have authority to regulate the part that simply wants, and not the other way around? That we do assume such an intrapersonal order is evident in very deep features of our practice. Consider the concept of "rationalization." We rationalize when we allow our wants to determine the course of deliberation. The thought that this counts as a corruption of reasoning arguably depends on the idea that there is a constitutional order in the soul. Our inclinations are not supposed to guide our deliberation. And when they do, we are corrupted. Similarly, consider

the concept of a “weak will.” A will is weak when it follows its inclinations in violation of its commitments.<sup>8</sup> And when our wills are weak, *we* are weak. The presupposition is that willing should check wanting, and not the other way around.

These are some rather obvious ways in which the intrapersonal hierarchy is written into our everyday practice. But there is another, more subtle way. So far the lower motives have appeared only as sources of corruption. But we also tacitly assume that they have a positive role to play. We assume that they provide “proposals” to be reviewed or ratified by our governing part.<sup>9</sup> Notice that inclinations differ from sensations, on the one hand, and from volitions, on the other. What inclinations and sensations have in common is a relation to feeling. In a broad sense of “feeling,” all sensations are feelings. And in an equally broad sense of “feeling,” all inclinations are “feelings-like”—as in, I “feel like” doing *x*. The difference is that inclinations are motives. They direct us to do something. Sensations as such do not—or at least not in the same way. You can have a sensation of the solidity of the chair you are sitting on without thereby having a motive to do anything. The sensation as such does not even purport to direct you.<sup>10</sup> But to be inclined, as I am understanding it, is to feel like doing something.

On the other side, inclinations and volitions share the feature that they are both motives. They differ, of course, in that volition does not involve “feeling-like” in the sense that inclination does. But they also differ in their basic relation to practical deliberation. Volition is the conclusion of deliberation. Inclination is its starting point. To see this, consider an asymmetry in our intuitions about inclinations and volitions. Put in the abstract, the asymmetry is this. Having an inclination to  $\Phi$  provides the occasion for deliberating about whether to  $\Phi$ , whereas having formed a volition to  $\Phi$  does not. The point is that when I simply feel like  $\Phi$ ’ing, I am in a condition that calls for deliberation about whether to  $\Phi$ , whereas when I have formed the will to  $\Phi$  (decided to  $\Phi$ , chosen to  $\Phi$ , committed to  $\Phi$ ), I am in a condition that does not call for deliberation about whether to  $\Phi$ . Suppose you have decided to go to the movies this evening. And suppose that, after having made this decision, you ask yourself, “Should I go to the movies this evening?” From the perspective of someone observing you, it is natural to want an explanation of why you are asking yourself this. Have circumstances changed? Were you mistaken about which movies are playing? Were you especially tired or emotional when you first made that decision? Did you start to feel like doing something else? Now suppose nothing of potential relevance in your internal or external circumstances has changed. You are just reopening the question. Given that nothing has changed, there is a sense in which you are being paternalistic towards yourself by reopening the question. You are “second-guessing” yourself without any basis for thinking your deliberation was in any way faulty. What this shows is that having decided to  $\Phi$  is a condition such that arbitrarily reopening

the question constitutes a sort of paternalism towards yourself. To question your reason on any basis other than reason itself is to fail to accept its proper authority, and this is to fail to accept your own authority. Implicit here is the idea of the supremacy of reason within the constitutional order of your soul.

Now notice that the same is not true of the condition of having an inclination. Take the same example. Suppose that at lunchtime you simply had an inclination to go to the movies. Upon having this inclination, you asked yourself, "Should I go to the movies?" You then deliberated and decided on a plan to do so this evening. Now, was there anything paternalistic in the fact that upon having the inclination to go to the movies, you asked yourself, "Should I go to the movies?" Were you second-guessing yourself arbitrarily? Notice that we are not even tempted to ask whether anything changed between the time you had the inclination to go to the movies and the time you opened the question. We assume nothing changed. We assume that you had the inclination to go to the movies under circumstances *C*, and it was under the same circumstances that you asked the question, "Should I go to the movies?" We do not think of you as reopening a question that you had already settled. Rather, we recognize that to have an inclination is to be in a condition that calls for deliberation.<sup>11</sup> You were not paternalistic towards yourself for taking your inclination to be the starting point for deliberation. On the contrary, you were upholding your integrity by subjecting the claim of your inclination to review from the perspective of your higher capacity. You were maintaining the proper hierarchy in your soul.

Let me take stock of what I have said so far. Implicit in our everyday intuitions about rationalization, weakness of will, and second-guessing oneself is the idea that we maintain our integrity when we maintain a certain order among our motives or our motivational capacities. While one feature of that order is that our volitions should take precedence over our inclinations in cases of conflict, another is that our inclinations should provide proposals to be taken up in practical deliberation. By doing so, they provide starting points for the process that results in volition.

## **2. Analogous Problems**

All of this is just to describe the structure of our intrapersonal constitution, insofar as it can be read off of our ordinary practices. The task is to justify it, to show why it makes sense that our constitution has that structure. Some light might be shed on this problem by following Plato's lead and drawing an analogy between our intrapersonal constitution and a political constitution. But I think this strategy is limited. Political constitutions are grounded in positive conventions. There need be no deep reason why this particular group of people should play the role of legislators while others should carry out the function of the executive. By contrast, the order in

the soul must somehow be grounded in our nature. Something about the respective natures of our higher and lower forms of motivation must make it the case that each plays the role it does.<sup>12</sup> Therefore in order to gain a clearer picture of both the problem and the space of possible solutions, I want to lay out an analogy between a more familiar natural hierarchy and the hierarchy in the soul.

The analogy is between the adult/child relation and the relation between what I will call our “volitional selves” and our “inclining selves.” What I am calling the “volitional self” is the self insofar as it has generated a higher motive through deliberation. This is the self insofar as it has made a decision about what to do. What I am calling the “inclining self” is the self insofar as it simply feels like acting in some determinate way or other. Now for the purposes of bringing out the analogy (but somewhat artificially), we should think of these selves as non-overlapping. To be an inclining self is in some sense to anticipate being a volitional self, but it is not yet to be one. And to be a volitional self is no longer to be an inclining self.

As we have already seen, our constitution is such that as volitional selves, we have asymmetric authority to check and regulate our inclining selves. But our inclining selves are not just potential sources of interference in our lives. The positive role of our inclining selves is to provide input to the deliberative process that results in our being volitional selves. Analogously, adults have asymmetric authority to check and regulate the conduct of children. But children are (of course!) not just potential sources of interference in the lives of adults. Their positive role is to be agents in early stages of a developmental process through which adults are formed.

In previous work on the moral status of children, I argued that children’s status in the moral community has a half-full/half-empty character.<sup>13</sup> Children are not mere things, but neither are they full members of the moral community. What counts as a violation of the rights of an adult does not always count as the violation of the rights of a child. This is evident in certain very basic elements of our practice. Children have to be raised, whether they like it or not. Children have to live under the supervision of adults, whether they like it or not. And in general, children’s choices, even self-regarding ones, are subject to adult approval in a way that would be objectionable if they were adults. Children’s intermediate status is also evident in our practices regarding responsibility. Children are not properly held accountable for what they do in the same, direct way that adults are. Instead of being objects of blame and punishment, they are objects of disapproval and discipline. These attitudes issue from the standpoint of a superior rather than an equal. They reflect a paternalistic attitude that would be inappropriate if taken towards an adult under the same circumstances.

Now notice that the status of our inclining selves likewise has a half-full/half-empty character relative to that of our volitional selves. To see this, consider the point I made earlier about arbitrarily second-guessing ourselves. Once you have made a decision to  $\Phi$ , if you ask yourself whether

to  $\Phi$  even though no internal or external circumstance of possible relevance has changed, then you are arbitrarily second-guessing yourself. The intuition that this counts as a kind of paternalism towards yourself suggests that your volitional self has something like a right not to be supervised by anything other than reason, which (I have assumed) is its own nature. The same is not true of your inclining self. Once you have an inclination to  $\Phi$ , you can certainly ask yourself whether to  $\Phi$  under exactly the same conditions without thereby being paternalistic towards yourself. This suggests that the claims of your inclining self are properly subject to review from a rational perspective.

The analogy holds as well on the side of responsibility. From the perspective of our volitional selves, we do not hold our inclining selves accountable in the same, direct way that we hold our volitional selves accountable. Suppose you are tempted to violate your marriage vows. You meet someone attractive and you feel like having an affair. But you decide not to act on that feeling. Reflecting on this situation, you might disapprove of your inclining self, insofar as it was susceptible to such temptation. You might also resolve to take indirect actions to make it less likely that you will be so tempted in the future. But it would be inappropriate for you to blame yourself simply for being tempted in the same way that you would blame yourself for having decided to cheat. And it would be inappropriate for you to punish rather than discipline yourself simply for being tempted. This suggests that, from the perspective of our volitional selves, our inclining selves are only responsible for their (inclining) activity in an attenuated sense.

The adult/child analogy helps us to see that implicit in our ordinary practice is the idea that our inclining selves have attenuated “rights” and “responsibilities” in relation to our volitional selves. The question is why. What is it about our inclining selves that makes them less than full participants in the intrapersonal order?

### **3. Analogous Spaces of Solutions: The Proficiency View**

Extending the analogy farther helps us see in broad outline the space of possible answers to this question. Elsewhere I have argued that there are two main ways of thinking about the basis of children’s subordination to adults.<sup>14</sup> The first I call “the proficiency view” and the second I call “the attributability view.” The proficiency view holds that children are subject to adult authority because they are likely to act badly when it comes to leading their own lives. The attributability view holds that children are subject to adult authority because their actions, whether good or bad, are not fully their own. Theories of the adult/child relation are generally not detailed enough to distinguish between these two approaches. But the costs and benefits of adopting one over the other are very different. Analogously, I will argue, the

proficiency and attributability views define the two main ways of thinking about the basis of the inclining self's subordination to the volitional self. On the proficiency view, our inclining selves have subordinate status because they are likely to lead us astray. On the attributability view, our inclining selves have subordinate status because their activity, whether reliable or not, is not fully our own. As in the case of children, theories of agency are generally not detailed enough to distinguish between these two approaches. And yet the costs and benefits of adopting one over the other are very different.

Let me start with the proficiency view as applied to the adult/child relation. This view is rather intuitive. It holds that the reason why children's claims are subject to review by adults is that children are relatively unskilled when it comes to leading their own lives. They are likely to make mistakes, including very costly ones. Adults are entitled to intervene in children's lives, even against children's wills, for the sake of making sure they avoid such mistakes.

One problem with this argument is that it provides no reason in principle why adults who are more skilled at living their lives should not intervene paternalistically in the lives of other adults who are less skilled at living their lives. And the root of this problem is that the view takes for granted that being likely to make a mistake is sufficient to deprive one of the right to make that mistake.<sup>15</sup> Adults are, by and large, entitled to make their own mistakes when it comes to leading their own lives. What is it about the condition of childhood that deprives children of the same entitlement? The proficiency view assimilates the condition of childhood to that of an adult whose functioning is limited by educational or intellectual handicap. But handicapped adults are entitled to make their own mistakes and to waive the assistance offered to them. Children are not. Suppose we provided children with teams of advisors, so that they would have access to expert assistance when making decisions. The proficiency view gives us no reason to deprive children of the right to waive that assistance, or to act against the advice they receive. But in this respect, it fails to capture the depth of paternalism implicit in our ordinary ideas about the adult/child relation.

The proficiency view has trouble on the side of accountability as well. What we need to explain is why children are not directly accountable for what they do. The view that children are less skilled than adults does not show why they are not proper objects of blame and punishment when they do make mistakes. Perhaps the explanation is that we should take the attitude that we are training them, so when they make mistakes, we should just correct them rather than punish them. But this does not explain the specifically disciplinary attitude that we actually take towards children. Adults whose job it is to train other adults in particular skills appropriately correct them when they make mistakes. But adults who discipline other adults treat them like children. The proficiency view might explain why we do not punish children, but it does not explain why we discipline them instead.

How does the proficiency view fare when applied to the relation between our inclining selves and our volitional selves? The claim here is that our choosing selves are entitled to supervise and regulate our wanting selves because our wanting selves are likely to lead us astray. Again, this is an intuitive view. But it suffers weaknesses analogous to the weaknesses we saw in the children case.

Remember what the theory has to explain. It has to explain why our volitional selves have a “right” not to be arbitrarily second-guessed, whereas our inclining selves do not. It has to explain, in other words, why having an inclination to  $\Phi$  provides the occasion for deliberating about whether to  $\Phi$ , whereas having formed the will to  $\Phi$  does not. The proficiency view explains this by saying that the inclining self is more likely to make mistakes than the volitional self. But even if this is true, how does it show that the inclining self does not have the same “right” to make mistakes that the volitional self has?

You might think there is no question of unequal “rights” here, because the inclining self and the volitional self are ultimately constituents of one self. If both constituents are equally us, if both are equally in a position to “speak for us,” then there is no hierarchy to maintain, and the decision which to identify with on any given occasion is purely pragmatic. Alternatively, if the idea of constitutive “selves” is just metaphorical, and all we have are mental states of different kinds, then the question we face is again pragmatic: whether to allow one or another type of mental state to serve as a reliable indicator of what we should do. The core of the proficiency view is that it treats inclination and volition in one of these two ways. It assumes that each has the same relation to the self (whether constitutive or external), and that the question of how to put them in relation one another is purely pragmatic.

So it is in principle possible, on the proficiency view, to hold that your decisions provide the occasion for deliberation, rather than your inclinations. For example, you might be a person who has learned from experience to trust your “gut” more than your “head.” You have found that your gut tends to lead you in the right direction, and your head tends to lead you astray. So you adopt a policy of second-guessing your decisions, while going along with your inclinations.

But even in this case, I would argue, you cannot escape tacit commitment to the traditional idea that your volitional self has an authoritative status that your inclining self lacks. Given that you have decided that you have good reason to second-guess each of your first-order decisions, your second-guessing of any given first-order decision is not arbitrary. It is as if, in a moment of sobriety, you discovered that the rest of the time you are drunk. Assuming that your drunk volitional selves have undermined their proper authority, you have good reason to act paternalistically with regard to them. Moreover, insofar as you do adopt this policy, you have to see yourself as upholding your integrity by adhering to it. But then you cannot maintain that your volitional selves as such are drunk. You have to take



the volitional self who has committed to this paternalistic policy to be in its proper, authoritative condition.

The point is that you cannot coherently adopt a policy of dissociating yourself from your decisions as such. At least some of your decisions are your will, and you identify your integrity with them. By contrast, you can coherently adopt a policy of not identifying any of your inclinations with your will, and of not defining your integrity in terms of them. Although there is a sense in which your inclinations can “speak for you” independently of whether you endorse them or not—as when they lead you to behave in ways that are predictable and familiar from the perspective of someone who knows you intimately—they cannot define the standard of your integrity unless you actively adopt a policy of acting on them. Whatever it is that makes this true also makes it the case that you are not being paternalistic when you arbitrarily second-guess your inclining self. The point of reflecting on your inclinations is not essentially to double-check their content, but to give them a kind of authority they would otherwise lack. The proficiency view misses this point. It sees volition and inclination as essentially two different methods of arriving at the right practical conclusions, without recognizing that they bear different relations to the self.

The proficiency view has parallel problems on the side of accountability. What needs to be explained here is why we don’t hold our inclining selves accountable in the same, direct way that we hold our volitional selves accountable. In the case where you feel inclined to have an affair but you decide not to, you don’t hold yourself directly accountable simply for having the inclination, at least not in the same way that you would hold yourself accountable for actually deciding to have the affair. How would the proficiency view explain this? The mere fact that our inclinations are likely to lead us astray does not in itself show that they are any less representative of us than our volitions. The proficiency view tacitly assumes that our inclinations can speak with our own authority just as much as our choices can. Or at least it is silent on the question whether this is so. As such, it cannot explain why we should not hold ourselves directly responsible simply for having the inclinations we have, even when we have refused to allow them to determine our wills.

Overall, the proficiency view has an undeniable simplicity and is intuitively very attractive. But a closer look shows that it underestimates the depth of the status distinctions that we routinely presuppose in both our adult-child relations and our intrapersonal relations. Indeed in effect, the proficiency view characterizes both hierarchies as conventional rather than natural. If the point of having the hierarchy is simply to avoid making mistakes, then the decision to establish it must be essentially pragmatic. We adopt a convention of putting adults in charge of children, and a convention of putting our wills in charge of our inclinations, because doing so tends to leave everyone better off. The weaknesses of the proficiency view might be avoided, then,

on a view that characterizes the hierarchy as grounded in nature rather than convention. The attributability view does just that.

#### **4. Analogous Spaces of Solutions: The Attributability View**

Whereas the proficiency view holds that the hierarchies are based on differences in the agents' degrees of skill or competence, the attributability view holds that it is based on differences in the agents' relations to their own activity. Applied to the adult/child relation, the attributability view holds that children are subject to adult authority because their actions, whether good or bad, are not fully their own. Applied to the will/inclination relation it holds that our inclining selves are subject to the authority of our volitional selves because their activity, whether good or bad, is not fully our own. I want to show now that in both cases, the attributability view avoids the weaknesses of the proficiency view by offering a deeper account of the respective natures of the agents who are hierarchically related.

In the adult/child case, the basic claim is that children's actions are not attributable to children in the same, direct way that adult's actions are attributable to adults. Why might this be so? The attributability view explains this by putting forth a conception of the condition of childhood as it relates to human agency. That is, it puts forth a view about what it is to be in the condition of childhood, given that one faces the task of exercising authority over oneself. The version of this view that I have laid out elsewhere holds that the condition of childhood is one in which the agent has not yet constituted her authority over herself.<sup>16</sup> She is indeed a human agent, but her distinctively human structure has not yet been established. She is on the way to constituting herself as a rational animal. At the risk of confusing the analogy I have been drawing, we can describe the condition of childhood in terms of its intrapersonal structure. As I see it, an agent in that condition does not yet admit of a clear distinction between her inclining self and her volitional self. She does not first have inclinations and then act on them. She is more like a creature of instinct, a creature whose wants just are its will. But she is not a creature of instinct. She is a rational animal who is working her way out of her animality, carving out a reflective distance between herself and her inclinations, and thereby constructing her identity. Or we could say she is a rational animal whose rational capacity is in the process of developing. Is maturation an activity she undertakes or a process she undergoes? The availability of both active and passive descriptions of it is entirely appropriate. Insofar as she is not yet herself, the distinction between what she does and what happens to her is not yet established. So maturation is no more and no less an activity than it is a process, even though the structure of our language is such that we have to describe it as one or the other.

Given that she has not yet constituted her authority over herself, the child's actions are not attributable to her directly as full or straightforward exercises of that authority. They are more like "play" actions, where "play" has the sense of rehearsal or experiment. But of course the child is not in a position to choose between play action and genuine action. Playing at being a full self is all she can do.

Does this sort of account explain the features of children's subordination that the proficiency account failed to explain? I believe it does. First, it explains why children, even if they are likely to make mistakes, do not have the same right to make their own mistakes that adults do. This is because children are not yet in a position to make mistakes that count straightforwardly as their own. As such, adult intervention in children's lives does not count as a violation of *them* in the same way that it would if they were adults. Paternalism, in other words, is only objectionable on the presupposition that there is a well-constituted self to interfere with. In the case of children, this is not straightforwardly the case.

Second, the attributability view explains why children are not held directly accountable for their mistakes. Resentment and punishment are attitudes that presuppose the wrongdoer's full authorship of his wrongful actions. Because the actions are attributable to him, he is accountable for them. But to the extent that this relation of attributability is attenuated by the condition of childhood, accountability should be likewise attenuated.

The attributability view locates the basis of the adult/child hierarchy in nature rather than convention. This is both its strength and its weakness. It is a strength because it allows the view to account for the depth of the status distinction between adults and children. But it might be considered a weakness because it forces the view to commit to a substantive conception of the condition of childhood that, because of its depth and detail, is bound to be philosophically controversial. I believe the gains outweigh the costs, but I will not try to give a full argument for that here. Instead I want to show how further extension of the analogy leads to a substantive philosophical conception of the condition of having an inclination.

## 5. Having An Inclination

How does the attributability view apply to the relation between will and inclination? The claim is that our inclining selves have subordinate status because their activity, whether reliable or not, is not fully our own. Now it might seem odd to attribute activity to the inclining self at all. For recall the notion of the inclining self is that of the self, insofar as it is in the condition of having an inclination to do something. And I stipulated that the condition of having an inclination to  $\Phi$  is not identifiable with the condition of having formed the will to  $\Phi$ . Inclining is not willing in the full sense, and the

inclining self is not the volitional self. But that is not to deny that inclining is a form of activity. In fact it has to be. For to be inclined is to be motivated, and motivation is self-movement. One who is inclined to  $\Phi$  is not externally moved in relation to  $\Phi$ -ing. She is in some sense moving herself in relation to  $\Phi$ -ing. But she is not yet  $\Phi$ -ing. So when I talk of the inclining self's activity, what I am referring to is the self-determination in virtue of which she counts as being motivated with respect to  $\Phi$ -ing.

The attributability view holds that while the activity of my volitional self is directly attributable to me, the activity of my inclining self is only attributable to me in an attenuated sense. So when I am in the condition of having an inclination to, say, eat that chocolate cake, the active responsiveness that constitutes my motivation is not something I stand behind in the way that I stand behind my actions. If I am in any sense its source, I am so indirectly.

But what independent reason do we have to think this is so? On analogy with the case of childhood, what we need here is a substantive conception of the nature of the inclining self. We need a conception of what it is for an agent to be in the condition of having an inclination, a conception that relates this condition to the task of exercising one's authority over oneself.

Before offering such a conception, let me make two clarificatory points about the concept of 'having an inclination' as I am using it. First, at the risk of confusing the analogy, keep in mind that given my conception of childhood, the condition of having an inclination is distinctive of us as adults. For as I mentioned, childhood is a condition in which the distinction between wanting and willing is not yet fully established. Correspondingly, it would be misleading to talk about children as having distinct inclining and volitional selves. Artificial as this way of speaking is, it is illuminating only if we use it to characterize an adult.

Second, in order to get at the fundamental nature of the condition, we should assume that when I have an inclination to  $\Phi$ , my will is as yet truly unformed with respect to  $\Phi$ 'ing. I have not made a prior commitment to  $\Phi$ , nor have I made a prior commitment to not  $\Phi$ . There is as yet no volitional self with respect to this particular action. Having an inclination is, in this case, the initial condition of the process by which my will is formed with respect to  $\Phi$ 'ing.<sup>17</sup>

My claim is this: when I am in the condition of having an inclination to  $\Phi$ , I am on the way to exercising my authority over myself so as to  $\Phi$ . I am not already  $\Phi$ 'ing, but neither am I doing anything other than  $\Phi$ 'ing. I am in between. Now I am not 'in between' in the sense that I could be in between the starting line and the finish line in a roadrace. That picture presupposes that I am fully constituted throughout the process. The sense in which I am in between is that my self is not yet constituted as the source of this specific action. I am not yet determinately the source of this exercise of my authority, but it is in some sense my authority, and I am in some sense exercising it.

Let me try to make this more intelligible. On the view I have been developing elsewhere, we are composed of two motivational capacities, one rational and one nonrational.<sup>18</sup> Our nonrational capacity is the source of our inclinations, and its structure is something like that of a nonhuman animal. It sees and responds to the world in a way that is intelligible but based on associative connections rather than reasoned justification. Our rational capacity, by contrast, demands and offers grounds for what it does.

Now given this dual nature, what is it to be in the condition of having an inclination to  $\Phi$ ? First of all, when I am in the condition of having an inclination to  $\Phi$ , I am both inside and outside my nonrational capacity. I am seeing the world through its eyes and responding with its heart, but I am also aware of a part of myself as not being the source of that way of seeing and responding. So when I have an inclination to eat that chocolate cake, I am aware of and responsive to the cake as delicious-sweet-thing-to-be-eaten, and I am also aware of myself as not being wholly identifiable with the part of me that is going for cake-eating.<sup>19</sup> To have an inclination is to be divided.

Moreover, I am divided without being conflicted. The case is not one in which I am tempted to violate a prior commitment. As I mentioned earlier, we are assuming the case is one in which my will is as yet unformed with respect to cake-eating. I have neither resolved to eat the cake, nor have I resolved not to.

The idea of division suggests a dualistic picture, according to which an independent agent that happens to be located inside me has determined itself to go for cake-eating, while the agent I am has not. But if this were the case, there would be no such thing as my inclining self. There would be no whole self that ‘has’ the inclination to eat the cake. There would just be a nonrational capacity that determines itself to eat the cake, and a rational capacity that does not relate to that action in one way or another. I could not even say that my inclination is my motive, because it would be the activity of that other agent inside me. It might move me but it could not motivate me.

The alternative way of thinking about it is this. To have an inclination is already to be determining oneself, to be exercising one’s authority over oneself. But it is to be doing so in a way that is incomplete. The exercise is primitive; it requires cultivation in order to be what it purports to be. What we identify as “the inclination” just is this primitive exercise. It is a primitive act of self-determination attributable to the nonrational motivational capacity. Its primitiveness, its nonrationality, consists in the fact that it can neither demand nor offer grounds for its activity. Insofar as I have an inclination to eat the cake, I am determined with respect to cake-eating in the spontaneous, associative way that a creature of instinct determines itself. I am responsive to the delicious-sweet-cake-to-be-eaten, but not to the appropriateness of taking delicious sweetness as grounds for eating the cake.

If I were a creature of instinct in the sense that nonhuman animals are, this wanting would constitute my will. There would be no distinction

between me-wanting and me-willing. But because I am a rational animal, I am aware of this primitive exercise of my authority, and I am aware of it as being insufficient to determine me to act. My rational nature makes it the case that I need to base my responses on something more than associations of ideas; I need a justification for what I do.

This creates a condition of normative instability. I already purport to be doing something, namely determining myself with respect to cake-eating, but I am aware of myself as not yet being in a condition to back up that purport. Hence the pressure to deliberate is pressure to make it the case that I really am doing what I already purport to be doing, namely determining myself with respect to cake-eating. And to do this, I have to exercise my natural authority in a way that takes account of my rational capacity. I have to make this primitive dictate into a law of my will. By doing so, I make my inclining self into my volitional self.

Again, is this development best conceived as an activity or a process? It is no more and no less one or the other. Because my self is not yet constituted with respect to cake-eating, there is as yet no bright line between what I do and what I undergo with respect to cake-eating. It is not as if I am a well-constituted agent who, upon having an inclination to eat the cake, then chooses either to eat the cake or to not eat the cake. Rather, my having an inclination to eat the cake sets me the task of constituting myself in relation to cake-eating. Unlike a child, however, I do this against the background of a self who is determined in relation to a wide range of other actions. As such, there are constraints on the process that are expressive of the person I already am.

It may seem as though we have wandered far from our main line of argument, which was on the brink of assessing the merits of the attributability view in the intrapersonal case. We needed a substantive account of the nature of the inclining self in order to get an idea of why the inclining self's activity is not directly attributable to it. With that on the table, how does the attributability view avoid the weaknesses of the proficiency view?

The first weakness of the proficiency view is that it cannot explain why the inclining self does not have the right to make its own mistakes—why it is subject to arbitrary second-guessing whereas the volitional self is not. The answer is that the inclining self is not yet a well-constituted self with respect to its object, whereas the volitional self is. When I respond to the condition of feeling-like  $\Phi$ 'ing by asking myself, "Should I  $\Phi$ ?" I am relating to a self that is not yet constituted in relation to  $\Phi$ -ing. By contrast, when I respond to the condition of having decided to  $\Phi$  by asking myself, "Should I  $\Phi$ ?" I am relating to a self that is fully constituted in relation to  $\Phi$ -ing. This is why only the latter counts as a kind of paternalism towards myself.

The second weakness of the proficiency view is that it cannot explain why we do not hold our inclining selves accountable for the ways in which they incline, whereas we hold our volitional selves accountable for the ways

in which they act. Again, the attributability view explains this by appeal to the idea that the inclining self is not a well-constituted source of its own inclining activity. Since the self that simply feels like  $\Phi$ -ing does not speak for me in the direct way that the self that decides to  $\Phi$  does, it is appropriate for me to refrain from holding the my inclining self directly accountable for its inclining activity.

## 6. Conclusion

I started this paper with a very general question: what is the basis of the alleged hierarchy in the soul? Why should we, as volitional selves, have asymmetric authority over our inclining selves? It is not because volition is intrinsically better than inclination, or because rationality is intrinsically better than nonrationality. Nor is it because as a pragmatic matter, it is better to put the rational part in charge. It is because the nature of our inclining selves is to be incompletely self-determined. As inclining selves, we are already on the way to being volitional selves, in something like the way that children are already on the way to being adults. We are in an internally unstable normative condition. Deliberation is our natural response to this instability. But this shows that deliberation is not fundamentally a skill among others. Rather, deliberation is more like maturation. It is the activity, and the process, through which we humanize ourselves, over and over again.

## Notes

- \* I owe thanks to the Radcliffe Institute for Advanced Study, which supported me while writing this essay. I am also grateful to all those who discussed earlier versions of this essay with me at the MIT Philosophy Department Work in Progress Seminar, the Legal Theory Workshop at Yale Law School, the Tufts University Department of Philosophy Colloquium, and the Ethicists Network Workshop at the Princeton Department of Philosophy.
1. Some of the clearest exemplars of this tradition are Plato, Aristotle, Butler, and Kant.
  2. The deeper issue here is whether rationality fundamentally consists in self-government, or whether it consists in achieving some kind of agreement with an external order. I am assuming that these somehow coincide, and I am ignoring cases that would put pressure on that assumption. A much deeper discussion of the relation of rationality and will would be needed to answer the obvious worries that arise when self-government and rationality appear to come apart.
  3. Traditionally, the concept of "desire" has been used both in a broad sense, to refer to motivation as such, and in a narrow sense, to refer to what I am here calling the "lower" form of motivation. I am deliberately avoiding the term "desire" so as not to invite ambiguity. For a helpful account of how contemporary action theory has suffered from a failure to disambiguate broad

- and narrow uses of “desire,” see G.F. Schueler, *Desire: Its Role in Practical Reason and the Explanation of Action* (Cambridge, MA: MIT Press, 1995), 29–38.
4. I am deliberately avoiding the term “intention,” because it has acquired a very specific technical meaning in contemporary action theory (see the work of Michael Bratman).
  5. I leave open the question whether practical conclusions are identifiable with actions.
  6. This is indeed a substantive view. I defend it in “The Nature of Inclination,” *Ethics* 119 (2009): 229–256. Some might argue that this assumption, while compatible with Plato’s theory of the nonrational part of the soul, is incompatible with Aristotle’s, since Aristotle says that the nonrational part “listens to” and is “in some sense persuaded by” reason (Aristotle, *Nicomachean Ethics* [Indianapolis: Hackett, 1985], 1102b13–1103a). Though I cannot argue the point here, my sense is that Aristotle does not develop this claim sufficiently to distinguish his conception from those that share this assumption.
  7. My ideal reader is someone who already assumes there is some important distinction between willing and wanting. This reader need not be committed to the idea that that distinction can be cashed out in terms of distinct motivational capacities rather than distinct types of motives. But a reader who does not believe there is an important distinction between willing and wanting will probably not be moved by my arguments here.
  8. I am assuming that inclinations are the source of weakness in the will, and that the essential failing consists in the self-betrayal, not the choosing of the perceived lesser good as such.
  9. See Christine Korsgaard’s description of “the constitution model” in Plato and Kant in her *Self-Constitution: Agency, Identity, and Integrity* (Oxford: Oxford University Press, 2009), 133–176, especially 141–42. Korsgaard accepts that the role of “appetite” is to propose, but she does not explicitly take on the question of how this assumption is embedded in our ordinary practice, or of why this is appetite’s proper role. I try to address the first of these questions here, and the second in later sections of this paper.
  10. This is a bit oversimplified. Sensations of pleasure and pain are more like inclinations, and require further discussion.
  11. Many contemporary theories of ‘desire’ tacitly acknowledge this intuition without explaining it. I have in mind particularly those according to which to have a desire is to apprehend certain features of the world as “appearances of the good” or as “seemings” that something would be good, or as “perceptions” of reasons or goodness. The language of “appearance,” “seeming,” and “perception” suggests that the condition of having a desire provides the occasion for further reflection. See, for example, Dennis Stampe, “The Authority of Desire,” *The Philosophical Review* 96 (1987): 359–62; T.M. Scanlon, *What We Owe To Each Other* (Cambridge, MA: Harvard University Press, 1998), 39–41; R. Jay Wallace, “Addiction as Defect of the Will: Some Philosophical Reflections,” *Law and Philosophy* 18 (1999): 641–43; Cheshire Calhoun, “Cognitive Emotions?” in *What is an Emotion?* ed. Robert Solomon (Oxford: Oxford University Press, 2003), 245–47; Christine Tappolet, “Emotions and the Intelligibility of



Akratic Action,” in *Weakness of Will and Practical Irrationality*, ed. Sarah Stroud and Christine Tappolet (Oxford: Clarendon Press, 2003), 108–111; Sergio Tennenbaum, *Appearances of the Good* (Cambridge: Cambridge University Press, 2007), 21–27. Elsewhere I argue that appeal to these notions does not explain anything unless accompanied by a substantive theory of perception. “The Nature of Inclination,” 243–44.

12. See note 9.
13. For more detail on this point, see my “What is a Child?” *Ethics* 109 (1999): 716–17 and especially “Childhood and Personhood,” *Arizona Law Review* 45 (2003): 575–78.
14. “Childhood and Personhood,” 579ff.
15. For more detailed versions of my arguments against the proficiency view, see “Childhood and Personhood,” 580–85.
16. For more detail, see “What is a Child?” 728–734, and “Childhood and Personhood,” 586–591.
17. This is related to a point I make in “Foregrounding Desire: A Defense of Kant’s Incorporation Thesis,” *The Journal of Ethics* 15 (2011): 160–161. I believe the primary role of inclination is its positive role as a source of proposals to reason. The inclinations that count as corrupting influences are proposals we have already implicitly considered but have ruled out by making a commitment. They have the status of rejected proposals.
18. See “The Nature of Inclination,” 246–53 and “Foregrounding Desire,” 154–56.
19. I believe my conception acknowledges what is true in two seemingly incompatible conceptions of desire: the view that to have a desire is to take a perspective on the world, and the view that to have a desire is to experience something exerting an influence on one’s will, as if from without.