

## Caring and Full Moral Standing\*

*Agnieszka Jaworska*

A being has moral standing if it or its interests matter intrinsically, to at least some degree, in the moral assessment of actions and events. For instance, animals can be said to have moral standing if, other things being equal, it is morally bad to intentionally cause their suffering. This essay focuses on a special kind of moral standing, what I will call “full moral standing” (FMS), associated with persons. In contrast to the various accounts of what ultimately grounds FMS in use in the philosophical literature, I will propose that the emotional capacity to care is a sufficient condition of an individual’s FMS as a person. In developing this account, I will appeal to a set of intuitions not previously mined for this purpose: those generated by conflicts of interests between different life phases of a single individual.

It is commonplace in ethical theory and standard moral practice to hold that ordinary adult human beings have an exalted moral standing—that we owe them special concern and respect. A core constituent of FMS is a kind of inviolability: roughly put, we are prohibited from destroying and from interfering in various other ways with a being with FMS for the sake of another being and its interests, or for any other value. For instance, we are prohibited from killing an innocent human being for the sake of saving one or several others, or even, perhaps, for the sake of justice and world peace; but we are allowed

\* Thanks to Jodi Halpern, Matthew Liao, Maggie Little, Debra Satz, Julie Tannenbaum, anonymous reviewers at *Ethics*, the members of the Fellows Seminar at the Center for Human Values at Princeton University, the audiences at the Bay Area Forum for Law and Ethics at the University of California at Berkeley, the Symposium on Evil and Psychopathy at the University of San Francisco, the Ethics Reading Group at the University of California at San Diego, and to my students in Philosophy 378 for helpful comments on earlier drafts. I am especially grateful to Jeff McMahan for extensive comments and for numerous very helpful conversations; to Melissa Fusco, Govind Persad, Caleb Perl, and Daniel Hallday for research assistance; and to the Center for Human Values at Princeton University for sabbatical support.

*Ethics* 117 (April 2007): 460–497

© 2007 by The University of Chicago. All rights reserved. 0014-1704/2007/11703-0005\$10.00

to kill a chicken for the sake of justice and world peace, or to save a human being, or even, perhaps, to save five other chickens.<sup>1</sup> Furthermore, all creatures with FMS are thought to possess such inviolability in equal degree.<sup>2</sup>

Full moral standing also encompasses weighty positive moral claims; admittedly, these are harder to isolate, since factors such as special relationships, the agent's own projects and commitments, and so forth, may override them. Still, other things being equal, beings with FMS are paramount in the calculus of positive duties—their interests matter much more than the interests of other types of beings. When faced with a choice of saving a human person or a chicken from a severe ordeal, barring further reasons that may complicate the moral picture (indirect consequences of saving the chicken for other people, etc.), one is required to pick the person—simply because of the kind of creature it is.<sup>3</sup> Further, and more important for my purposes, *ceteris paribus*, like interests of all creatures with FMS matter equally in moral decisions, giving rise to the requirements of fairness. For example,

1. The last example is borrowed from Judith Jarvis Thomson, *The Realm of Rights* (Cambridge, MA: Harvard University Press, 1990), 292. To put the point in the text differently, persons have special rights of noninterference, the content of which is specified by the correct theory of rights.

2. If A and B are both inviolable, how could they be unequally inviolable? It could be worse to violate A than to violate B in the same way. FMS rules out such comparative variations: if A and B have FMS, otherwise equivalent violations of each are equally bad.

3. Admittedly, privileging the person in such cases does not necessarily imply that the person as such matters more than the chicken: it is also possible that each creature matters morally just as much, but the person's interest at stake turns out to be more compelling. After all, the same level of pain may have more severe and lasting consequences for the person, given her ability to remember the pain, to attribute meaning to the pain in various ways, to fearfully anticipate further pain of this sort, etc. A similar imparity of interests could be behind the intuition that saving the life of a person is more important than saving an animal's life, since many feel that animals, due to their simple psychology, don't have a strong interest in continuing to live. Because, on this reading, the two interests could never (in the standard cases) be equal, I don't cite this intuition in the text. But for readers who don't accept that animals' interest in survival is weak, this intuition may be more telling (and all the more so since it persists even if the interests of the person are otherwise weaker: for example, even if the person is terminally ill, while the animal is young, with a long life ahead of it). In light of these complications, perhaps the clearest illustrations of the more stringent positive duties with respect to persons and their interests as a mark of FMS in commonsense morality are those involving simple inconveniences. Consider this scenario. There is only one very comfortable spot in the sun that both a cat and a person would greatly enjoy. You are in charge of the spot, and you don't have a special relation to either individual. Surely you would let the person have the spot. (Pet owners make analogous choices when they wave their pets off their favorite nooks to satisfy any wish of a person, however trivial, or when they lock pets in small enclosures to avoid even slight disturbances or inconveniences for themselves.) If you decided to flip a coin between the person and the cat, the person would rightly feel that you have treated her inappropriately by putting her on equal footing with an animal.

when distributing goods among creatures with FMS (who could all benefit similarly), barring special purposes, relationships, or independent claims on the goods, we ought to distribute equally.<sup>4</sup> By contrast, beings lacking FMS are outside the scope of the considerations of fairness. (A state policy that takes no account of the welfare of chickens living in the state is not deemed unfair.) Third, it is morally costly to forgo support of vital interests of a being with FMS, even when one is morally justified. For example, if one is distributing scarce resources among creatures with FMS (think of famine relief), and one lacks enough to protect vital interests of all affected, this is a moral misfortune of the gravest kind.

Commonsense morality seems to dictate that membership in the human species, regardless of capacities, guarantees FMS (at least for independently subsisting beings; the status of fetuses is indeed disputed). Infants and severely retarded humans are normally taken to have FMS, while nonhuman animals of the very same or even higher mental capacities are not. This commonsense approach, whose advocates I'll call "preservationists," is contested by theorists I'll call "revisionists."<sup>5</sup> Both preservationists and revisionists standardly back their positions by appealing—albeit in rather different ways—to valuable mental capacities that underlie FMS. However, neither side can defend, or even merely corroborate, its choice of the relevant capacities by consulting ordinary intuitions about who does and doesn't have the requisite inviolability and positive claims. The resulting stalemate on both sides of the debate motivates the project of this essay. Let me elaborate.

Revisionists see the commonsense view as unjustifiably inconsistent: moral standing must be a response to some valuable properties of the being in question, and similarly endowed beings should have similar standing. To fill in their theory, revisionists need to specify which capacities give rise to FMS, but, aiming to revise common sense, they cannot do so by relying on ordinary intuitions about who has FMS.

On the other side, preservationists do not simply claim that membership in the human species as such confers FMS. They typically keep

4. Ideas from John Broome ("Fairness," *Proceedings of the Aristotelian Society* 91 [1990–91]: 87–101) helped me in this paragraph, although not all I say is consistent with Broome.

5. For a prominent example of a preservationist view, see T. M. Scanlon, *What We Owe to Each Other* (Cambridge, MA: Harvard University Press, 1998), esp. 177–87. For a prominent example of a revisionist view, see Jeff McMahan, *The Ethics of Killing* (Oxford: Oxford University Press, 2002), esp. 203–32.

firmly in view the valuable traits of ordinary adult humans,<sup>6</sup> which account for FMS in these uncontroversial cases. They then explain the special standing of humans in the nonstandard cases in two ways. For infants and children, they appeal to the potential to acquire the relevant capacities in normal development. And they cover severely retarded human beings by appeals to the species nature, membership in a kind, the characteristic form of a species or a kind, and so forth,<sup>7</sup> that is, to some sort of complex (presumably biological) mode of organization that, barring defects, makes the emergence of the relevant properties possible. Clearly, then, preservationist theories also have to rely on an account of valuable capacities that ultimately underwrite FMS.<sup>8</sup>

Moreover, preservationists cannot use the commonsense intuitions that a particular individual possesses FMS as a base or source of reality checks for their account of which capacities ultimately underwrite FMS. This is because, as they take potential and/or species capacities to be sufficient for moral standing, in any given nonstandard case involving humans, it is unclear whether an occurrent or merely potential/species property is the source of FMS. For example, human infants have many potential capacities and some occurrent capacities; according to common sense, infants have FMS, but this tells us little about which capacity confers FMS.<sup>9</sup>

In short, both proponents and opponents of the commonsense view of FMS need an account of the mental properties that ultimately justify these special moral requirements. And they need to derive and test their accounts on the basis of something other than our everyday intuitions

6. There are important exceptions here—views that emphasize the role of relationships and of the possibility of the being in question sharing a form of life with us. See, e.g., Cora Diamond, “The Importance of Being Human,” in *Human Beings*, ed. David Cockburn (Cambridge: Cambridge University Press, 1991), 35–62; Peter Byrne, *Philosophical and Ethical Problems in Mental Handicap* (New York: St. Martin’s, 2000), chap. 3.

7. I list these different formulations since different proponents of the views in this category phrase their positions slightly differently. Sorting out whether these different formulations amount to substantive differences is beyond the scope of this essay.

8. As Peter Byrne perspicuously exhibits, even theological accounts, which ground FMS in humans’ unique relation to God, when fully worked out, are typically forced to appeal to the very same valuable properties of humans invoked by the secular accounts: if you don’t believe that God acts arbitrarily, something must explain why God cares about all human beings. (See Byrne, *Philosophical and Ethical Problems*, chap. 6.)

9. Note that the preservationists can use intuitions that a particular individual *lacks* FMS to confirm or disconfirm views about the ultimate source of FMS. These intuitions are not confounded by the possibility that potential or species capacities could ground FMS. For example, if, according to common sense, a great ape capable of using language lacks FMS, this rules out the view that language use is the ultimate basis of FMS and confirms views that require even more sophisticated capacities as the ultimate basis of FMS.

as to who is and isn't morally inviolable and paramount. This theoretical impasse calls for a new strategy. Hence I propose to consult our intuitions about one's moral standing within a life—that is, vis-à-vis one's earlier or later higher self—to test the plausibility of views about FMS across lives. To mark the contrast, I will call the former type of FMS “intra-agential,” and the latter, more familiar FMS “interpersonal.”<sup>10</sup> Since intra-agential FMS is unlikely to be based on membership in the human species or on potential abilities, examining our intuitions in the intra-agential case can be particularly helpful in uncovering what capabilities, in general, underlie FMS. The alternative intuitions may not be firm enough to settle the question, but when combined with further arguments, they can establish the capacity to care as a very plausible ground of FMS.

Admittedly, a new account of the attributes that underlie FMS—and specifically my appeal to the capacity to care—will not resolve the debate between preservationists and revisionists. Preservationists can still insist, against their revisionist opponents, that all human beings have FMS because humans as a species possess the capacity to care—or indeed any attribute plausibly deemed relevant. On this issue, I will remain agnostic. But it still matters a great deal that both sides correctly identify the ultimate basis of FMS: that they zero in on the right explanation of why we owe special moral duties to persons. Such sound theoretical foundation will also allow the two approaches to delineate more accurately their disagreements and perhaps even discover previously unforeseen common ground.

## I. INTRA-AGENTIAL FMS

### A. *The Concept*

A human being has varied abilities throughout life. Abilities are limited in childhood and may diminish later in life due to brain disease or injury. Imagine yourself as a custodian for someone in a phase of life when abilities are underdeveloped or diminished. Suppose that your charge's vital contemporaneous interests come into conflict with vital interests—satisfiable now, at least to some extent—that she had or will have at a time when she possesses normal abilities. For example, the

10. The term ‘intra-agential FMS’ does not strictly capture the notion I mean. I have in mind FMS at issue in conflicts of interest within the same life, but on a standard understanding of the term ‘agent’ (as someone capable of acting) a human individual may not be an agent in all phases of her life. ‘Intra-individual’ would be a more apt term, but it is too cumbersome. I thus settled on ‘intra-agential’ with the caveat that ‘agent’ should be understood very broadly, to encompass at least all conscious beings.

current interests of a “pleasantly demented” individual can conflict with more sophisticated values from her past, or the current interests of a child can conflict with the needs for intellectual development serving her future self.<sup>11</sup> How should such conflicts be resolved? Is the contemporaneous individual morally on par with her “higher” self from a different time? Is it morally inappropriate—or, at least, does it involve a formidable moral cost—to override this individual’s contemporaneous interests by appeal to her interests as an (incontrovertibly) full-fledged person at a different time? If we decide that it is—by virtue of what the individual is like here and now—we can talk about the person she is now having FMS in the intra-agential sense. And we will thereby take a stand on a question apparently analogous to the one pressing on theorists of the more familiar interpersonal FMS: by virtue of which abilities does a person merit FMS in this sense?

For a very rough sketch of what I have in mind, consider different cases in which a demented individual’s current pressing interests conflict with more robust values from her past. If her dementia is only mild, it seems inappropriate to ignore her current interests for the sake of her former values. But in the case of severe dementia, privileging former interests does begin to make sense. The difference in our responses seems to depend on whether the current interests of the individual fully matter morally. In this way, these differing intuitions appear to reveal the moral standing of the current self: a mildly demented individual has intra-agential FMS, while the severely demented one does not.

Similar differences in intuitions can arise in conflicts involving a future “higher” self. But here, due to an asymmetry in how the future and the past matter to us, the full moral importance of the current interests is not revealed in how the conflict is resolved but, rather, in the moral cost of sacrificing the current self. For example, when the current salient interests of an infant conflict with the anticipated needs of her future adult self, we have no moral qualms in sacrificing the infant’s current interests. But if the current salient interests of a teenager generate a similar conflict, although we may still privilege the future self, sacrificing the teenager’s current interests does seem to be morally costly. And this suggests that a teenager has intra-agential FMS while an infant does not.

In both sets of cases the differences in moral standing seem to depend on the capacities of the current self. If so, explanations of the

11. As will become clearer in Sec. I.B, it is important that the conflicts of relevance here are generated by the interests of the “diminished” self, and not merely by her choices. An individual lacking decision-making capacity will likely make many inappropriate choices, and when custodians rightfully bypass such choices this may have no bearing on the individual’s FMS.

differing intuitions in such cases could help us uncover the capacities that underlie FMS.

To pursue this strategy, we need to be clear that interests stemming from different phases of a life could be in genuine conflict. First, how can a *past* self be a party to such a conflict? How can the interests of a past self be meaningfully satisfied in the present? Indeed, since many interests depend on the interest bearer having specific kinds of experiences (such as enjoyment), they cannot be satisfied once the bearer no longer exists. However, not all interests are experiential: satisfaction of a desire can advance a person's interests even if she is unaware of the desire being satisfied, even if she is dead at the time, or unconscious, or too demented to notice. And further, the realization of one's long-standing projects and the fulfillment of one's values can advance one's interests in a similar way, provided one has not reassessed the values in the interim. So a past self can have interests, and these could potentially conflict with the interests of the present self.<sup>12</sup>

Second, more fundamentally, we need to be clear that life-phase interests can indeed be individuated, so that there could really be two separate parties to an intra-agential conflict. Some readers will assume, in opposition to my picture, that so long as a creature retains her identity over time, ultimately, what's in her best interest is what will lead to her having the best life overall. On this view, her interests cannot be broken down into the competing interests of her time-slice or life-stage selves: since she is one and the same creature, her past and future interests are still (or already) her interests here and now. For example, the past values of an Alzheimer's patient are still defining of her present interests (provided she has not reassessed them in the meantime), and the interests of the adult into whom the current child will develop are already part of the interests of this child as she is now.

The intuitive appeal of this view stands or falls with a specific interpretation of prudential concern, so let me briefly introduce this concept. Prudential concern is familiar to all of us as the special kind of concern we each have with our own person: concern for our own discomforts, our own experiences, our own conduct, the satisfaction of our own desires, and the shape of our own futures or pasts. (To see that prudential concern extends also to the past, consider how you would care about your past wrongful or embarrassing conduct.) Some people care more strongly about others than they care about themselves, but even for them, concern with themselves has a unique quality. Think of the difference between dreading your own boredom and anticipating

12. For a further discussion, see, e.g., Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1986), app. I, esp. 495.

an analogous episode of boredom for a loved one. You would care in either case, but you would care very differently.

Since we know the phenomenon of prudential concern from contemplating our own futures (or pasts) in ordinary circumstances, it seems natural to assume that we inevitably have prudential concern for our own lives and experiences and only by virtue of their being our own. The assumed inevitable prudential concern for our own lives is what matters here, as it leads to the picture of interests at odds with my approach: so long as one retains one's identity over time, one ought to have prudential concern for one's past and future selves, thereby generating a unified sense of one's overall interest and so ruling out the individuation of separate interests of one's time-slice or life-stage selves.

However, the connection between prudential concern and identity is not this close. Jeff McMahan has argued persuasively that one should have little prudential concern for oneself in the future (or the past) if one's later (or earlier) psychological makeup is not sufficiently connected and continuous with one's current psychology. In such cases one would rationally lack the ordinary high level of prudential concern for one's own future (or past) self.<sup>13</sup> I take for granted here McMahan's view that prudential concern must be proportional to the strength of one's psychological ties with one's future (or past) self. Thus, in cases in which the psychological ties are loose—such as the more severe cases of dementia—it definitely makes sense to distinguish potentially competing interests associated with those psychologically detached phases of life.<sup>14</sup>

But are these the only kinds of scenarios in which interests stemming from different phases of one's life could be separated out? On the view I have sketched so far, a creature with sufficient psychological continuity and cohesion over time ought to have robust prudential concern for her past and future selves, and this suggests that such a creature would generate a unified perspective on her lifelong interests and sep-

13. McMahan, *The Ethics of Killing*.

14. One may worry that, whenever a severe enough psychological discontinuity is involved, discontinuity alone would generate a prohibition against overriding the individual's current interests for the sake of her interests from her psychologically remote future or past. If this were the case, such a prohibition would not be indicative of FMS—it would not be a function of the current capacities of the individual, but only of the weakness of the relationship between her life stages. This worry is exacerbated once we realize that, realistically, loss of intra-agential FMS could occur only in cases involving psychological discontinuity: after all, the very capacities that make strong psychological cohesion possible would likely provide the basis for FMS. However, the worry can be easily allayed. As we shall see shortly, if an individual lacks intra-agential FMS, he is not intra-agentially inviolable, so there is no prohibition against sacrificing his interests for important enough interests stemming from a more sophisticated stage of his life. Lack of psychological continuity with that life stage is insufficient to introduce such a prohibition.

arate interests of her time-slice or life-stage selves could not arise. According to this suggestion, it is only severe forms of psychological discontinuity—those occurring between infancy and adulthood, or those wrought by extensive dementia—that would give rise to distinct sets of interests stemming from different life stages. However, this view is too narrow: differentiations of life-stage interests can emerge without such radical psychological discontinuities. Even ordinary changes of values, so long as the new values are stable enough, could be thought to lead to the individuation of vital interests associated with one phase of life (e.g., old age) that don't apply at other phases. (Derek Parfit's case of the Russian nobleman is an apt illustration.)<sup>15</sup> And, more important for our purposes, a phase of life such as childhood, due to the child's limited abilities, may be associated with its own set of interests, even if the psychological integration between an older child and the subsequent adult is rather robust.<sup>16</sup>

All this leads to an intricate picture of possible intra-agential conflicts. In cases in which there is no radical psychological discontinuity, in one sense—the sense that treats life-as-a-whole as a locus of interests—the past or future interests are part of the individual's current interests.<sup>17</sup> But in another sense—the sense that treats the individual's current nature as the locus of interests—past and/or future interests may be severed from current interests. The former view of interests allows weighing current interests against past or future interests in a prudential way without a cost to anyone, while the latter does not. Accordingly, we get the complicated result that, in some cases (as it turns out, only in cases of conflict with a future self—for reasons I will elaborate shortly), the custodian may be permitted to override a person's current pressing interests—on the basis of the first, life-as-a-whole sorts of considerations—and yet such a choice leaves a moral remainder, a severe moral cost, because the individual's current nature is not properly respected. (For example, we may be permitted to override a teenager's current pressing interests to make her life better overall, but this choice incurs the moral cost of failing to respect the teenager here and now.) This moral complexity is due to the individual's dual nature: as someone who is such-and-such now—a child, a teenager, and so forth—but also as someone who is becoming an adult. And because of this complexity, intra-agential FMS is often revealed not through intuitions as to whether a given set of interests ought to prevail in cases of conflict, but rather through intuitions concerning the moral remainder if the interests are

15. Parfit, *Reasons and Persons*, 327.

16. See McMahan, *The Ethics of Killing*, 499–500.

17. Life-as-a-whole considerations may still be somewhat relevant even in cases of radical discontinuity. I don't settle this question, since it doesn't affect my analysis here.

in fact overridden for good reasons. When, as in mild dementia, a custodian is not morally permitted to override the current interests of an individual—even though the individual is in a “diminished” phase—this straightforwardly demonstrates this individual’s current intra-agential FMS. But if the permission does seem reasonable (as it may be for infants and teenagers), we need to consult our intuitions about the moral cost of such a choice in order to determine the intra-agential moral standing of the “diminished” individual.

*B. Parallels between Intra-agential and Interpersonal FMS*

My strategy will be to uncover what capabilities, in general, underlie FMS by examining our intuitions in the intra-agential case. But to import these results into theories of the more familiar interpersonal FMS, we need to develop sufficient confidence that, morally speaking, the two senses of FMS are truly aspects of the same concept—that they comprise morally equivalent requirements with a common rationale—so that they would appropriately be based on the same set of capacities. I will begin by considering reasons for denying the parallel.

One may think that the prohibition against (or the moral remainder generated by) overriding the interests of a person with intra-agential FMS in the name of more sophisticated interests stemming from another part of her life is a variant of a straightforward, antipaternalistic prohibition. Just as we are not allowed to impose on a person with sufficient decision-making capacity our own—even if clearly superior—view of her interests, it is likewise morally problematic to impose on a person with sufficient decision-making capacity a conception of her interests, no matter how superior, that stems from a different phase of her life. If this analogy were defining of intra-agential FMS, interpersonal and intra-agential FMS would be associated with somewhat different requirements, and we wouldn’t have grounds for assuming that they are based on the same set of capacities. This is because antipaternalism is not a core requirement of interpersonal FMS. Granted, it often accompanies interpersonal FMS, in the sense that many individuals with interpersonal FMS also have the requisite decision-making capacity and so aren’t supposed to be treated paternalistically. But antipaternalism is plainly not necessary for interpersonal FMS, since while most of us think that young children have interpersonal FMS, antipaternalism certainly does not apply to them.

However, a closer analysis makes clear that intra-agential FMS is not merely or even necessarily about protecting a person against paternalism—individuals who evidently lack decision-making capacity can have intra-agential FMS. To see this, consider Ms. P, a happily married woman in her forties, in very early stages of early onset Alzheimer’s disease. She is somewhat disoriented and forgetful and cannot keep

track of the details of her emerging additional medical problem—breast cancer. She is overwhelmed by the complex set of treatment options, each with its own advantages and drawbacks, and so forth. This patient clearly lacks decision-making capacity with respect to the medical decision at hand, and it would be appropriate, even required, to have a surrogate make decisions on her behalf. Suppose also that, with the onset of dementia, Ms. P's values have shifted as follows. She is accustomed to being deemed beautiful, but, having devoted her adult life to intellectual pursuits, she used to derive her sense of self-worth from professional success; she loathed treating looks as a source of pride. But now that her intellect is fading, her appearance has become a central locus of concern and the one remaining solid source of self-esteem. Imagine that she now faces two options for cancer treatment: either a lumpectomy, in her case somewhat more risky to health and life, yet designed to largely preserve the shape of her breast, or a radical mastectomy, much safer, but also certain to leave her disfigured.<sup>18</sup> There is no doubt that had Ms. P made the choice before the onset of dementia, she would have picked the safer option—her relatives even recall her explicitly asserting this in an earlier hypothetical discussion of such a choice. Ms. P's present lack of decision-making capacity, being simply due to mild confusion and forgetfulness, in no way implies that current decisions on her behalf ought to be based on her past interests (in which health and life take priority while appearance is irrelevant) rather than on her current interests (in which appearance is paramount). In fact, given that her dementia is only in very early stages, and that her mental capacities have deteriorated only slightly and remain basically intact, it seems quite certain that she retains FMS in the intra-agential sense. So this case demonstrates that respect for intra-agential FMS is not an analogue of an antipaternalistic prohibition and that intra-agential FMS and decision-making capacity can diverge.<sup>19</sup>

We thus need not worry that the requirements of intra-agential FMS match the antipaternalistic requirements of respect for autonomy (which, recall, often accompany interpersonal FMS without being a necessary ingredient of it) rather than the core requirements associated with interpersonal FMS and the inviolability of persons. But the parallel may be doubted for a different reason. Granted, the prohibitions against overriding an individual's contemporaneous interests (or the moral costs of doing so) do indicate some moral standing of that contem-

18. These were typical treatment options at the time, in the late 1990s.

19. Note also that if the requirements of intra-agential FMS amounted merely to antipaternalism, these requirements would be stricter than what I have outlined: paternalistic treatment of a competent time-slice self is presumably morally prohibited and not merely morally costly.

poraneous self, but are they a mark of FMS? After all, I have alluded to other—weaker—types of standing: animals might have some moral standing while lacking FMS. Thus, we need to carefully lay out what the requirements of intra-agential FMS must be to truly match the requirements of interpersonal FMS specific to persons rather than the requirements of a moral standing of a weaker sort.

Just as an essential part of interpersonal FMS is the prohibition against treating vital interests of a person as available to be interfered with and sacrificed for the sake of the interests of someone else or even of a large number of others, so the intra-agential FMS of a person requires that her vital contemporaneous interests not be treated as straightforwardly available—without a formidable moral cost—to be interfered with and sacrificed for the sake of more weighty interests of her own past or future selves.<sup>20</sup> And just as it is part of interpersonal FMS that, other things being equal, interests of a person must be treated fairly, on par with otherwise equivalent interests of other persons, so the intra-agential FMS stipulates that a person's contemporaneous interests be treated on par with the interests of her more sophisticated past or future self.<sup>21</sup> And similarly, by analogy to interpersonal FMS, recognition of intra-agential FMS implies a moral misfortune of the gravest sort when vital interests associated with a phase of life marked by personhood go unmet.

On this picture, both interpersonal and intra-agential FMS have as

20. None of the cases I discuss below involves interference, but there are plenty of situations in which intra-agential FMS manifests itself by making interference morally problematic. Suppose parents face a decision whether to subject their child to a painful operation with lasting effects that will diminish the child's quality of life throughout childhood but enhance the child's prospects later in life (bone operation for short stature, cochlear implants, etc.). It may be permissible to impose the surgery on the child, but the choice involves the severe moral cost of considerably disrupting childhood, a whole phase of a life of a person. The moral cost remains significant even when the negative effects of the surgery during childhood are less severe, perhaps just a short period of convalescence. By forcing the child to undergo surgery, one uses the child against its present interests for the sake of the interests of its later, more advanced self, and this in itself generates a high moral cost if done to a person. (Thanks to Jeff McMahan and Jodi Halpern for inspiring this comment.)

21. It may appear odd that the requirements of fairness apply even intra-agentially. Concerns with fairness are typically motivated by the recognition of the separateness of persons. (For example, a Rawlsian would oppose utilitarianism on the grounds that while maximizing well-being seems appropriate within a life, separateness of persons makes it inappropriate—unfair—to maximize well-being across lives.) So applying fairness to person stages may seem to overextend this concept beyond its proper home in interpersonal morality. However, if the framework I introduced in Sec. I.A is correct, the interests of person stages can be sufficiently independent to require separate moral attention, and it is precisely this separateness of interests that makes the requirements of fairness applicable. Thanks to Winston Chiong for prompting this clarification.

their point the preservation of different aspects of essentially the same inviolability and moral paramountcy of a person. They both encode the same moral importance of persons, and the only difference has to do with whom the interests of the person are traded off against—against other persons or against other phases of the same life. And thus there is good reason to believe that both kinds of moral standing are (ultimately) merited by the same properties of the beings involved.<sup>22</sup>

Let me spell out the comparison in more detail, showing how the parallel requirements are triggered in intra-agential conflicts. When contemporaneous interests of an individual are sacrificed for the sake of her temporally absent better self, in many cases this will be structurally quite different from one person's interference in the life trajectory of another—it will not qualify as one life-stage self's interference in the affairs of another. Yes, the interests of the contemporaneous self are sacrificed, but perhaps only in the sense in which a bystander faced with two people in need, when he is physically able to attend to only one, ends up “sacrificing” the interests of the other. There is no interference here, and no breach of inviolability of a person. And this will often be the position of a caregiver who forgoes (and in this sense “sacrifices”) the contemporaneous interests of his charge for the sake of the interests of her better self, because he can't promote both. But, crucially, as we have seen, such choices can still reveal assessments of intra-agential FMS: individuals with intra-agential FMS must be treated fairly in such decisions; interests of each must be given their due weight.

When a conflict of this nature is with a past self (as in the dementia cases), it is analogous to a conflict between the interests of a now existing person and the otherwise equivalent interests of a deceased person. Other things being equal, the interests of the dead count for less than the interests of the present or future people: it matters to the strength of interests whether the live experience and/or the actively anticipated prospects of the subject will be affected. When death blocks all possibility of such effects, the importance of the interests diminishes considerably.

22. Some readers may still think that, despite all these structural parallels, the bar for meriting interpersonal FMS is lower than the bar for meriting intra-agential FMS—because protection against others seems more basic and vital than protection against one's more advanced self. Even if this were correct, we could still infer interpersonal FMS from intra-agential FMS, but lack of intra-agential FMS would not imply lack of interpersonal FMS. However, the detailed requirements of intra-agential FMS, which I go on to discuss presently, militate against this approach. It is true that protection against one's more advanced self is not as stringent as protection against others: in some cases, it is morally appropriate to sacrifice the interests of the current self to promote the best life as a whole for the individual. But there is a special moral cost to such a sacrifice if the current self is a person, and the bar for incurring such moral cost is not higher than the bar for meriting interpersonal moral standing of a person.

For instance, while a deceased person has a legitimate interest in having her last will carried out—say, in being able to influence her projects and relationships after her death—this is not nearly as vital as her interest in the success of her projects and relationships when she was alive. Analogously, other things being equal, the interests of a past self count for less than the interests of the present self. That is, in a conflict between a present and past self, it matters that the proposed intervention affects the live experience and/or the actively anticipated prospects of the present self but has no chance of ever similarly affecting the past self. If the contemporaneous individual has intra-agential FMS (recall the mildly demented Ms. P), this possibility of a live benefit from the satisfaction of her interests weighs sufficiently to break the deadlock—just as it would in an interpersonal case—so that fair treatment requires privileging such live and active interests of the contemporaneous individual over otherwise comparable interests of her past self with intra-agential FMS. Only if the contemporaneous individual lacks intra-agential FMS (as presumably happens in later stages of Alzheimer’s disease) can the interests of her past self with intra-agential FMS trump: although the past interests are not live and active, they are interests of a more morally weighty being.

By contrast, when the conflict involves a future self (as it would for an infant or a teenager), the live and active nature of the contemporaneous interests is no longer a distinguishing factor: the intervention affects the active life and experience of the present self, but it will (in due course) also presumably affect the active life and experience of the future self. Just as the otherwise equivalent interests of present and future separate people are on par, so are the interests of a present and a future self if both possess intra-agential FMS. But now, given this parity, there is a further complication: the promotion of the best trajectory for the individual’s life as a whole matters too and will typically favor the more sophisticated interests of the distant self over the interests of the present self. In such cases, the best course of action, overall, may be to privilege the future interests, but if the contemporaneous individual has intra-agential FMS, treating his current interests as subordinate to otherwise equivalent interests of his distant self will incur a weighty moral cost, the cost of treating a person unfairly, and, in most cases, also the cost of forgoing the support of vital interests of a person. The presence or absence of such moral costs reveals the intra-agential moral standing of the contemporaneous self. For example, the cost is absent for an infant but present for a teenager.<sup>23</sup> In these ways, our intuitive judgments

23. The moral cost will be incurred only if the sacrificed interests are vital, or at least on par with the interests of the distant self that are given priority. Consider a ten-year-old who wants a facial tattoo. “She has a present interest in getting the tattoo: it’ll definitely

about how intra-agential conflicts ought to be resolved, and about the moral remainders generated by such resolutions, do encode the requirements of the morally appropriate treatment of persons, parallel to the familiar requirements of interpersonal FMS.

*C. The Crucial Difference between Intra-agential and Interpersonal FMS*

One may worry that the parallel between interpersonal and intra-agential FMS I have now outlined is too close. That is, if both senses of FMS express aspects of the same inviolability and moral paramountcy, then they should apply to exactly the same individuals, *on exactly the same grounds*. If potential persons or members of species that typically become persons have interpersonal FMS, they should also have intra-agential FMS—if they are morally paramount in the one way, they should also be morally paramount in the other. And this would mean that examining intra-agential FMS would not have the advantage I have anticipated—the intuitive assessments of intra-agential FMS in particular cases would not be any less confounded.

Fortunately, though, the parallel is not exact. While it may make sense to attribute interpersonal FMS by virtue of an individual's potential or species abilities (thereby recognizing her inviolability and moral paramountcy vis-à-vis other beings), it does not make sense to assign intra-agential FMS (and thereby recognize an individual's inviolability and paramountcy vis-à-vis more sophisticated stages of her own life) on the same grounds. That is, the views I am dismissing here would attempt to justify an individual's intra-agential FMS in a less advanced stage of her life by appeal to her potential and/or species properties, which the individual presumably approximates more closely in the more sophisticated stage of her life. But it does not make sense to privilege stage A over stage A+ (to promote the interests of stage A as opposed to the interests of stage A+, let alone to make stage A inviolable against the interests of stage A+) <sup>24</sup> on the ground that at stage A the individual has the potential to develop into stage A+, or on the ground that at

---

establish her as cool among her friends. But her parents forbid it because they correctly realize that the tattoo would imperil her later prospects for marriage, employment, etc." (McMahan, unpublished comment). Not only is the parents' action permissible, but it also doesn't seem to carry the moral cost of failing to respect the girl, the person she is now. This is presumably because the interest in being cool, as a type of interest, is not terribly important and not nearly as weighty as the girl's later interest in employment, etc. If we sacrificed one person's interest in being cool for the sake of another's interest in marriage and employment, this similarly wouldn't involve a failure to respect the first person or a grave moral cost. Things would be different, though, if forbidding the tattoo were to somehow ruin the girl's childhood—a weighty interest of the child would generate a moral cost.

24. And also, in the moral remainder cases, it does not make sense to worry that the interests of stage A are sacrificed to the interests of stage A+.

stage A the individual has the species capacity for *c*, which he instantiates more fully at stage A+ rather than at A. For example, it is incoherent to protect a fetus against the interests of the mature human being it will become on the ground that the fetus is now precious because it will become that very mature human being. And it is incoherent to protect a severely cognitively impaired individual against the interests of the full-fledged person he used to be on the ground that the impaired individual belongs to a kind whose members normally develop into such full-fledged persons.

The point may also be put this way. When we respect a creature by virtue of its potential, by virtue of what we expect it will become, the current interests of that creature are understood in terms of this potential. Many conceptualize the interests of infants in this way. Similarly, if we respect a creature by virtue of its membership in a species or a kind, its interests are understood in terms of the species interests or the interests characteristic for that kind. (This is, arguably, why we normally consider a cognitively impaired human being to be afflicted, why we think his lot calls for a remedy, why we may even think it tragic, although we take an ordinary dog, or any other animal whose cognitive abilities are significantly more limited, to be just fine.) If interests of a creature are so framed, it doesn't make sense to protect her contemporaneous interests against the interests stemming from the more advanced stages of her life. And this is to say that merely potential or species properties cannot ground intra-agential FMS.

So it turns out that the relationship between interpersonal and intra-agential FMS is likely to be exactly what is called for if the study of intra-agential FMS is to have the theoretical advantage I anticipated. Both variants of FMS are ways of recognizing aspects of the same kind of inviolability and moral paramountcy, so both, I argued, should ultimately be grounded in the same capacities of the beings involved. I have allowed that *potential capacities* or membership in a *species possessing the relevant capacities* may ground interpersonal FMS in particular cases—still, it is crucial to remember that even in those cases the moral importance of those *capacities themselves* (as they are when actualized) must be the ultimate foundation of this interpersonal FMS. In the intra-agential case, by contrast, the inviolability and moral paramountcy at issue can only be a response to the individual's *occurrent* capabilities, so ascriptions of intra-agential FMS cannot be confounded by the complication that inviolability and moral paramountcy might also be due to potential or species capacities. Individuals who enjoy intra-agential FMS must concurrently possess the capacities that ground their intra-agential FMS—and consequently (ultimately) ground any other form of FMS that issues analogous moral requirements. If my analysis is right, by

examining which individuals these are and what capacities are distinctive of them, we can learn a great deal about the ultimate basis of FMS.<sup>25</sup>

More specifically, we can usefully test the prevalent ideas about the basis of FMS against our intuitions about the presence and absence of intra-agential FMS in particular cases in which the capacities of the individuals in question are well understood.<sup>26</sup> This is what I propose to do next. As we have seen, intra-agential FMS is a complex concept, so our intuitions about it are not straightforward and sharp-cut, but they can be a source of a valuable new perspective as to which attributes ultimately justify FMS and which do not.

## II. TESTING THE RECEIVED WISDOM ABOUT THE BASIS OF FMS

Full moral standing is usually thought to be grounded in the ability to reason. But there is quite a bit of variation in how this ability is understood and which aspects of it are taken as important for FMS. I will first show that when the ability to reason is interpreted in standard ways, it is either insufficient or unnecessary for intra-agential FMS. Our task is to find attributes that are both necessary and sufficient for intra-agential FMS and thus likely to be sufficient for interpersonal FMS.<sup>27</sup>

### *A. Are Minimal Rational Capacities the Basis of FMS?*

On many views, very minimal rational capacities, such as the ability to use language, manipulate abstract concepts, or form a conception of oneself, are sufficient to ground FMS. But are these capacities indeed sufficient for intra-agential FMS?

25. Given this analysis, from now on I will speak of the basis (ground, foundation, etc.) of FMS, and I will mean by this the capacities that constitute the ultimate basis of interpersonal FMS and, by extension, the basis of intra-agential FMS (or, what amounts to the same thing, the capacities that constitute the basis of intra-agential FMS and, by extension, the ultimate basis of interpersonal FMS).

26. Since I chose not to take a stand on whether potential or species membership can be the source of interpersonal FMS, this limits somewhat the power of my arguments: I am not always able to draw inferences about interpersonal FMS of a particular (type of) being from intuitions about intra-agential FMS of the same (type of) being. (To be exact, I can draw such inferences from the presence of intra-agential FMS, but not from its absence.) But this leaves my strategy intact. If I am right that the ultimate foundations of interpersonal and intra-agential FMS are the same, the ultimate foundation of interpersonal FMS can be discovered by investigating the basis of intra-agential FMS. A particular being who lacks intra-agential FMS may still possess interpersonal FMS due to her species or potential capacities. If so, the moral importance of those capacities themselves is the ultimate foundation of her interpersonal FMS, and we can discover which capacities these are by examining which beings occurrently possess (or lack) intra-agential FMS.

27. This difference in what we are looking for in the two cases arises, recall, from the possibility that potential or species capacities could also be sufficient for interpersonal FMS, so the presence of specific attributes might not be necessary for interpersonal FMS.

Consider an Alzheimer's patient who still possesses minimal rational abilities—he can talk sensibly, say, about the weather or food, and he competently employs self-reflexive pronouns—but whose desires and corresponding interests are consistently confined to sensual and bodily appetites. His life revolves around eating, smoking cigars, sunbathing, and so forth. Given the rudimentary level of this individual's current interests, one would be tempted to sacrifice them, in cases of conflict, for the sake of the more sophisticated interests from his past. For instance, if the man's current choices of simple pleasures—the remaining sources of joy in his life—go against the religious values of his past self, we would be tempted to sacrifice the simple pleasures to honor the former values. But those who consider such treatment permissible may well be simply inconsistent if they also maintain that this man retains the abilities that ground FMS. An individual endowed with abilities that ground FMS is entitled to fair treatment, and this includes fair treatment vis-à-vis the claims of his past self, no matter how refined.<sup>28</sup> Sacrificing the present self's active life and experience for the sake of inactive interests of the nonextant past self would not be fair to the present self. The sacrifice is permissible only if the present self lacks intra-agential FMS and the requirements of fairness no longer apply to him.

The weakness of the view that minimal rational capacities underlie FMS can easily go unnoticed if one focuses solely on interpersonal FMS. Many people assume that being a person is the basis of FMS, where 'person' is a metaphysical category delineated independently of moral concerns. And one could, of course, reasonably think that the ability to use language, the ability to manipulate abstract concepts, and/or the capacity for reflective self-awareness are marks of personhood in the metaphysical sense. One could then, understandably, come to believe that these capacities ground FMS. Since this claim is hard to verify by examining one's intuitions about interpersonal FMS (because potential/species capacities as possible sources of interpersonal FMS confound the picture), it may appear robust. Only by examining intuitions about intra-agential FMS (with the background realization that fairness to persons must also be maintained in intra-personal conflicts) does the weakness of this view come to light.

Thus, at least those prepared to sanction subordinating current interests of an individual who retains only minimal reasoning abilities for the sake of his more sophisticated interests from the past may not

28. Note that, given my assumptions, not every individual who has interpersonal FMS needs to be currently endowed with the abilities that ground FMS. So I am not claiming here that an individual who has interpersonal FMS must be treated fairly vis-à-vis his past self. That is, what I say here is consistent with my earlier view that interpersonal FMS does not imply intra-agential FMS.

be able to also maintain that these abilities are a mark of (moral) personhood. Since I concur with this intuitive judgment of intra-agential FMS, I am inclined to see it as damaging to the view that minimal rational capacities are the ultimate ground of FMS.

*B. Is the Faculty of Practical Reason the Basis of FMS?*

Next, consider the influential view associated with Immanuel Kant, according to which the distinctive mark of persons, and the consequent ground of FMS, is the ability to employ reason in the practical domain, in the domain of action. Persons use reason in order to figure out what ends to pursue, what ends are good to have,<sup>29</sup> and also to figure out how to realize those ends in the concrete circumstances in which they live. In this sense, persons are able to live their lives by their own lights: through the use of reason, they can set their own standards, their own values, and then lead their lives according to those self-imposed standards. Persons can live by laws they impose on themselves—they can be autonomous.

This may be interpreted as a very demanding set of requirements. In order to figure out what ends are good to pursue, one needs, at a minimum, sophisticated conceptual abilities to grasp various ends; one needs to be able to compare and contrast, weigh pros and cons, and understand the consequences of various options. One may also need to be able to grasp (at least implicitly) and to apply the very complex concept of justification.<sup>30</sup>

Are these sophisticated abilities necessary for FMS in the intra-agential sense? The somewhat demented Ms. P, recall, lacks crucial aspects of this sort of autonomy, and yet we are inclined to attribute intra-agential FMS to her. Also, children do not develop the requisite abilities probably until age eight or nine, and some think that even adolescents don't adequately command them.<sup>31</sup> But it seems implausible to insist that younger children lack intra-agential FMS and that their contemporaneous interests can be sacrificed for the sake of their future autonomy without a grave moral cost.<sup>32</sup> These results cast doubt on the

29. For a Kantian, these two tasks amount to the same thing, but I leave room for the possibility that they might be distinct.

30. Note that even the most minimal requirements of practical reason are much more cognitively demanding than the "minimal rational capacities" (language use, self-awareness, etc.) discussed in the previous section.

31. Tamar Schapiro, "What Is a Child?" *Ethics* 109 (1999): 715–38.

32. Recall that we cannot test our views about interpersonal FMS in analogous ways: Ms. P, children, etc., surely possess interpersonal FMS, but this doesn't show that the capacity for autonomy cannot ultimately ground FMS: in those special cases, interpersonal FMS could be due to past, potential, or species capacity for autonomy.

view that the capacity for autonomy, as traditionally understood, grounds FMS.

### III.THE CAPACITY TO CARE AS AN ALTERNATIVE BASIS OF FMS

So far, I have suggested that minimal reasoning capacity is insufficient for intra-agential FMS, while robust Kantian autonomy is unnecessary. We are looking for mental attributes both necessary and sufficient for intra-agential FMS and thus presumed to be sufficient for interpersonal FMS. As a first step, I will develop the observations that young children have the capacity for caring and that they also seem to have intra-agential FMS into the hypothesis that the capacity to care is sufficient for intra-agential FMS.<sup>33</sup>

I focus on the case of young children because here the ability to care stands out as crucial in the assessment of intra-agential FMS: young children have not yet developed a lot of other abilities traditionally thought to underlie FMS, including most elements of Kantian autonomy, and yet we attribute intra-agential FMS to them. (By contrast, the case of Ms. P—while effective in showing that full Kantian autonomy is unnecessary for intra-agential FMS—cannot be used to showcase the ability to care as the most likely explanation of FMS, since patients with mild dementia retain many abilities that could be thought to explain their FMS, including some individual elements of Kantian autonomy, such as the ability to make evaluative judgments.)

Consider this description of the concerns of a young child, from the earliest memories of Sergei Aksakov:

My little sister I loved at first more than all my toys, more than my mother; and this love took the form of a constant desire to see her, and a feeling of pity for her: I always fancied that she was cold or hungry and in want of food, and I wished constantly to give her my food and dress her in my clothes; of course I was not allowed to do this and that made me cry. . . . I could not bear to see her tears or hear her cry without beginning at once to cry myself. . . . I lay whole days in my crib with my sister beside me, amusing her with different toys or by showing her pictures.<sup>34</sup>

We readily interpret little Sergei's concern for his sister as a form of caring. And we marvel that a young child may be capable of such a

33. My discussion of the nature of caring and the internality of caring, here and in Secs. III.A and III.C, including several examples of caring in children, is borrowed, sometimes verbatim, from Agnieszka Jaworska, "Caring and Internality," *Philosophy and Phenomenological Research* 74 (2007), forthcoming.

34. Sergei Aksakov, *Years of Childhood* (Oxford: Oxford University Press), 1–6. According to the recollections, Aksakov would have been around three at the time.

sophisticated attitude. Intuitively, something notable, something quintessentially human, is manifest here.

There is ample evidence that children as young as two and three are capable of caring. Developmental psychology research shows that at around one-and-a-half years children begin to “act constructively to relieve another’s distress,” both at home and in laboratory playrooms.<sup>35</sup> In a study of early sibling relationships, by the time the older siblings were thirty months, at least 30 percent were observed expressing empathic concern for the younger child, typically through helping gestures, such as “offering toys or food when the sibling was crying.”<sup>36</sup> There is also anecdotal evidence that two- and three-year-old children exhibit other forms of caring, caring about objects or ideas.<sup>37</sup> Consider a prototypical toddler who consistently insists on doing many routine tasks—from eating to tying her shoes—by herself. The child can usually convey just how important being in charge is to her: she attempts to take charge in novel ways and in novel circumstances, she communicates joy at her successes, she gets angry at a parent who doesn’t let her try, and so on. Or take the story of a colleague’s son who took very seriously the task of blowing out all the candles on his third birthday: he prepared for the event, got very frustrated with his initial not-so-successful attempts, triumphantly announced his eventual success, and reminded his parents about the feat for days.

There is something compellingly human in these displays of caring, a special form of motivation expressed in action. These children do not engage in the complex reasoning required for Kantian autonomy, and they surely lack the necessary capabilities. And yet we would be very tempted to attribute FMS to them, even in the intra-agential sense.

Imagine, for instance, this continuation of Aksakov’s story. Sergei and his parents live in a small rural community in Russia. The girl Sergei calls “sister” is not in fact a blood relative: she is a daughter of their local servant. The parents have determined that for the sake of Sergei’s future they need to move to the big city, where Sergei will have better prospects of healthy intellectual development. Yet Sergei loves his “sister”; he absolutely doesn’t want to move and lose close contact with her. His parents may ultimately decide that, all things considered, it is still

35. Ross A. Thompson, “Empathy and Emotional Understanding: The Early Development of Empathy,” in *Empathy and Its Development*, ed. Nancy Eisenberg and Janet Strayer (Cambridge: Cambridge University Press, 1987), 131. For more recent data, see Carolyn Zahn-Waxler, Marian Radke-Yarrow, Elizabeth Wagner, and Michael Chapman, “Development of Concern for Others,” *Developmental Psychology* 28 (1992): 126–36.

36. Judy Dunn and Carol Kendrick, *Siblings: Love, Envy, and Understanding* (Cambridge, MA: Harvard University Press, 1982), 114.

37. I am assuming that caring about things and ideas and caring about people have core features in common. The basis of this assumption will become clear shortly.

best to move, but Sergei's separation from the girl he cares about carries a serious moral cost, the cost of failing to respect little Sergei, the person he is here and now. No such cost would have been at issue had Sergei been two years younger—still an appetite-driven creature—and his parents realized that he would be denied the satisfaction of one of his strongest desires because his favorite local baby food would be unavailable in the city of their destination.

If such intuitions are to be trusted, the Kantian account of FMS must be reconfigured in order to accommodate the suggestion that caring attitudes of young children can serve as its source. At first, it may appear that only a modest modification is called for. After all, caring attitudes may be thought of as the starting points of the ability to reason about goals and actions—as inputs into decision making—and this perhaps would qualify them as appropriate bases for FMS on the Kantian picture. Developing this idea, one could think that caring about something must involve positively evaluating the object of care, so that evaluative judgment or evaluative conviction of some sort is part and parcel of caring.<sup>38</sup> Such an evaluative judgment need not draw on the ability to form a conviction on the basis of reasoning or justification. A rudimentary judgment that the object of care is good could consist merely in properly applying the normative concept “good” to the object. This, at a minimum, requires only a commitment (albeit implicit) to the correctness or appropriateness of the judgment—a conviction that one is right. For example, the judgment holder must understand and (at least provisionally) accept that if she were not to hold the judgment—say, if she were to change her mind in the future—she would be making a mistake or that other people whose situation is relevantly similar would be mistaken if they didn't hold the judgment.<sup>39</sup>

It might be reasonably presumed that two-year-olds are capable of evaluative judgments in this thin sense. However, developmental psy-

38. What follows is a quick sketch of my argument against this interpretation. For details, see my “Moral Psychology in Practice: Lessons from Alzheimer's Disease and the ‘Terrible Twos’” (unpublished manuscript, Department of Philosophy, Stanford University).

39. This is a requirement for evaluative beliefs, not meant to apply to other beliefs (about tables and chairs, etc.). For a belief to have evaluative content, the belief holder must understand the evaluative concept evoked in the belief. And one does not really understand the evaluative concept one purportedly employs in one's belief unless one would recognize the lack of such a belief—especially one's own former, future, or counterfactual lack—as a mistake. Believing that my chair is blue does not immediately involve me in imputing error to anybody (including a counterfactual me) who lacks this belief—my chair-related belief is not inherently tied to considerations about believing, my own or others'. But if I believe one ought to help others in need, or that it's good to have fulfilling personal relationships in one's life, I do implicitly hold that those who lack this belief (esp. my past or projected self who lacks it) would be importantly mistaken.

chology provides a convincing body of evidence to the contrary: two- and even three-year-olds don't understand the idea of correctness as it applies to beliefs, including their own beliefs. Hundreds of studies corroborate this conclusion. These are variations of the so-called false-belief tasks, which test children's predictions about how a deceptive object will appear to others. In one scenario, children are presented with a familiar candy box, which, they soon find out, is filled with pencils. "They are asked what someone else will think when they first see the box. Three-year-old children consistently say that the other person will think there are pencils in the box. They apparently fail to understand that the other person's belief may be false."<sup>40</sup> In other experiments, they make the same errors even when they are asked about *their own* immediately past false beliefs.<sup>41</sup> They simply don't understand that beliefs are representational states that can be correct or incorrect.

Two- and three-year-olds are thus incapable of being committed to the correctness of their own convictions, and so they cannot hold evaluative judgments even of the most rudimentary sort. If carings of these children are starting points of reasoning about goals and actions, this cannot be because they incorporate evaluative judgments even in the weakest form. A more radical rethinking of the Kantian view is needed to accommodate the intuitions suggesting that carings of young children can be the source of FMS.

Of course, interpersonal FMS of young children is easy to explain by appeal to their (or their species') potential to develop sophisticated mental capacities, including the capacity for Kantian autonomy. But these explanations, as we have seen, do not apply to the intra-agential FMS. So, whether or not one accepts my more contentious claims, one must face the difficulty of explaining the intra-agential FMS of young children. And the idea that caring is the source of intra-agential FMS offers a plausible way out of this difficulty, even for those who doubt my more ambitious thesis that intra-agential and interpersonal FMS have analogous requirements and so must have a common source.

#### A. *A Closer Look at Caring*

To appreciate the relevant features of carings, we need to see them first and foremost as emotional attitudes. An emotional episode is a particular pattern of thoughts, feelings, bodily states, involuntary facial expressions, predispositions to act, perceptual selectivity, direction of attention and imagination, and so forth, which amounts to a person being, for example, angry, or fearful, or ecstatic at a particular time. By contrast,

40. Alison Gopnik, "How We Know Our Minds: The Illusion of First-Person Knowledge of Intentionality," *Behavioral and Brain Sciences* 16 (1993): 4.

41. *Ibid.*

what I mean by an emotion is a more enduring state, an ongoing psychic orientation, comprised of various interrelated emotional episodes and dispositions to experience subsequent concurring emotional episodes, unfolding intermittently over time, waxing and waning, at least partly in response to the context at hand.<sup>42</sup> For example, the emotion of grief typically involves episodes of recurring painful thoughts concerning the lost person or object, especially the recollection of the circumstances of the loss; a tendency to imagine how things could have gone differently; a predisposition to notice and dwell in one's thoughts on objects, events, or locations that remind the bereaved of the lost figure; a predisposition to hold onto the mementos of the lost figure; and so on. Other combinations of gradually and intermittently evolving elements, structurally linked in a similar fashion, constitute anger, fear, joy, disgust, jealousy, pity, and so forth.

In this way, emotions are constituted by conceptual connections between various elements of a person's psychology occurring at different points in the history of her mental life. Emotional episodes have conceptual content, they are about something, and the contents of episodes of the same emotion are tied to one another. Sometimes one episode of an emotion directly refers to another: as when a bereaved person recalls the experience of putting away the mementos of the lost figure. But otherwise, most of the episodes are interconnected referentially by virtue of referring in a consistent way to the same figure or to the same circumstances; in our example, the connected episodes are those of recalling, imagining, wishing something about, and associating one's current experience with the lost figure and the circumstances of the loss. That is, the episodes are various ways of being mentally agitated by the thought that the lost figure is missing.

Now, caring has an even more complex structure than most ordinary emotions—it is best understood as a structured compound of various less complex emotions, emotional predispositions, and also desires, unfolding over time in response to relevant circumstances. Typical components of caring include joy and satisfaction when the object of one's care is doing well and advancing and frustration over its misfortunes or setbacks, anger at agents who heedlessly cause such misfortunes or setbacks, pride in the successes for the object and disappointment over its defeats or failures, the desire to help ensure those successes and to help avoid the setbacks, fear when the object is in jeopardy and relief when it escapes untouched, and grief at the loss

42. See Peter Goldie, *The Emotions: A Philosophical Exploration* (Oxford: Oxford University Press, 2000), esp. chap. 2. What I say here is indebted to Goldie's work, but it is not meant to be a full or fully faithful representation of his views.

of the object and the subsequent nostalgia.<sup>43</sup> Not all of these elements must always occur in a given case of caring, but if enough of them are missing in the relevant circumstances, talk of caring is not warranted.

Such component emotions, emotional predispositions, and desires all construe the same object—a person, an animal, an ideal—as a source of importance commanding emotional vulnerability. This object, as a steady focus of emotional attunement, conceptually connects all the components to one another and gives the ensuing complex emotion its structure. The complex emotion is, as a result, *about* the object, for example, a caring about Mom.<sup>44</sup>

The connections between the components of caring are not limited to a shared reference to a common object. A complex pattern of rational requirements is involved, so only certain emotions concerning the object and only certain sequences of such emotions qualify as part of caring about that object.<sup>45</sup> Joy at the successes of the object of care rationally requires sadness at the object's failures; if things go well for the object, fear or hope are rationally required to turn into an emotion such as relief; and so forth. Thanks to their rational and referential interconnection, the individual caring emotions are intelligible as mental states of one agent.<sup>46</sup> But, more important, by virtue of his steadfast emotional attunement to the ebb and flow of the fortunes of the object, the caring subject imbues the object with importance. A subject cognitively sophisticated enough to employ, at least implicitly, the concept of importance would *comprehend* the object's importance, and this can inspire further cognitive activity, for example, further inquisitiveness about the object, or the formation of stable intentions, plans, and policies concerning the object. This cognitive sophistication is necessary for genuine caring.<sup>47</sup>

43. This aspect of my understanding of caring is indebted to Bennett Helm's work. See, e.g., "Freedom of the Heart," *Pacific Philosophical Quarterly* 77 (1996): 76–77.

44. For a very helpful elaboration of this idea, see Helen Nissenbaum, *Emotion and Focus* (Stanford, CA: Center for the Study of Language and Information, 1985), esp. chaps. 5–7.

45. For a more detailed discussion, see Bennett W. Helm, *Emotional Reason: Deliberation, Motivation, and the Nature of Value* (Cambridge: Cambridge University Press, 2001), 67–69.

46. The connections don't guarantee that the individual caring emotions belong to the same agent. But they contribute to the overall psychological unity and so, on the neo-Lockean picture, to the unity of agency over time.

47. There may be a weaker notion of caring for which such conceptual sophistication is not necessary: the subject could imbue the object with importance by virtue of his emotional reactions, without having a conceptual grasp of the object's importance. In fact, animals such as cats and dogs seem capable of "caring" in this weaker sense. However, I focus on caring of the more sophisticated kind described here because, on my view, only this kind of caring can be the basis of FMS. As I will explain in Sec. III.C, by virtue of

Findings in developmental psychology suggest that complex emotional structures of the sort I have just described do indeed lie behind the caring behaviors of young children. A longitudinal study of “concern for others” in the second year of children’s lives reports “consistently observed linkages between prosocial actions, expressions of concern, and verbal attempts to comprehend the nature of distress events that children cause and witness.”<sup>48</sup> The studied children exhibit consistent caring behaviors toward the same person (usually their mothers) over a period of time. When the person is in distress, they provide verbal and physical comfort, offer advice, try to help in various ways, or at least to distract the person to make her feel better; they also try to protect her from injury or distress. These actions are accompanied by “emotional arousal that appears to reflect sympathetic concern for the victim,” as evidenced by “facial or vocal expressions (e.g. sad looks, sympathetic statements . . . in a soothing or reassuring tone of voice, or gestures such as rushing to the victim while looking worried).” The children also focus their attention in a manner characteristic of caring: they attempt “to label or understand the problem,” asking “What happened?” or trying to diagnose the problem on their own, sometimes making “complex inferences.”<sup>49</sup> As I read it, they use their emotional sense of the mother as important to guide their reasoning and behavior.<sup>50</sup> While the study did not record the full gamut of caring emotions, and while perhaps some relevant emotions may have not yet had a chance to arise in the short time span of early childhood, it is hard to interpret the documented combination of reactions as anything other than the beginnings of a pattern of caring.

My anecdotes about individual children make better sense in light of these more systematic findings. Knowing that two-year-olds do begin to develop the complex emotional structures characteristic of caring, we can more justifiably assume them as background to the individual episodes of caring behavior. Besides, the anecdotes do contain enough interrelated elements evolving over time so that the emotional structure of caring is somewhat explicit. In the story about the birthday candles, witness the boy’s frustration at his original failures to accomplish the task he cared about, elation at his eventual success, and pride in the accomplishment, as evidenced by his repeated reminders about the in-

---

inspiring further cognitive activity (including the formation of stable intentions, plans, and policies), this kind of caring helps to support the unity of agency over time and thereby helps to fashion a self worthy of special respect.

48. Zahn-Waxler et al., “Development of Concern for Others,” 134.

49. *Ibid.*, 129. I omitted a confusing use of brackets from the original.

50. It is important to note that we can be at least confident, based on young children’s linguistic capabilities, that they could grasp the concept of importance. This point will be relevant again in my discussion of great apes below.

cident to his parents. The structure is even more conspicuous in Aksakov's recollections about his sister.

*B. Is the Capacity to Care Necessary for Intra-agential FMS?*

With this detailed picture of caring in the background, the importance of the ability to care for intra-agential FMS can be further corroborated by examining cases in which this ability is lost.

Consider Elliot, a patient described by Antonio Damasio. Elliot, who sustained brain damage in the ventromedial prefrontal cortices, performed normally or even superiorly on a full battery of psychological tests (tests of intelligence, knowledge base, memory, language, attention, basic reasoning, etc.), and yet he was an extremely poor decision maker in everyday life. He was unable to work, maintain personal relationships, or stick to a plan or task: he was incapable of sustained goal-directed behavior in any aspect of his life. And yet he was remarkably emotionally unaffected by his tragic life history. He would recount the story of his failures "with a detachment that was out of step with the magnitude of the events. . . . Nowhere was there a sense of his own suffering, even though he was the protagonist."<sup>51</sup> Elliot did not seem to care about himself. He did not experience joy or satisfaction when he was flourishing or frustration over his misfortunes; no anger at agents who caused his misfortunes; no pride in his successes or disappointments over his failures. He had no sustained desire to help ensure those successes or to help avoid the failures. He did sometimes experience fear when he was in danger, especially in response to stereotypical danger clues, such as a loud noise.<sup>52</sup> But this is not enough to constitute a pattern of caring about himself. Caring about oneself is such a basic form of caring that we wouldn't expect someone incapable of caring about himself to be able to care about anything else. And so it was with Elliot.

Elliot showed no abnormalities in means-ends reasoning and problem solving; he was perfectly able to come up with a full array of options for action in a particular situation as well as to work out the consequences of each option. Yet his ability to choose was impaired. After a full analysis of all the options he would comment, "I still wouldn't know what to do!"<sup>53</sup> His emotional responses and feelings were severely blunted, and this "prevented him from assigning different values to different options, and made his decision-making landscape hopelessly

51. Antonio R. Damasio, *Descartes' Error: Emotion, Reason, and the Human Brain* (New York: Putnam's, 1994), 44.

52. *Ibid.*, 138–39.

53. *Ibid.*, 49.

flat.”<sup>54</sup> He was no longer sufficiently invested in anything; he simply ceased to care.

This is not to say that Elliot had no contemporaneous interests stemming from the part of his life after the brain injury. He was still, of course, a sentient creature and had interest in avoiding physical pain and in mental and physical comfort. He had likes and dislikes and habitual pastimes such as collecting. Most important, he did find things captivating “in the moment.” What he lacked was sustained and caring interest in anything, a pattern of emotional investment retained over time.<sup>55</sup> His fleeting impulses, motivations, and attractions constantly took him off track, so he couldn’t maintain job commitments or personal relationships. Consequently, his interests did not revolve around relationships and commitments, but for that, they were not any less defined.

A condition such as Elliot’s has been called “acquired psychopathy” because the psychiatric profile of patients with those brain injuries is strikingly similar to the well-studied profile of a psychopath. While the popular image of psychopaths emphasizes their brazen lack of concern for other people, the psychiatric diagnostic criteria have recognized the much broader scope of their deficit: their underlying problem is a lack of deep emotional attachment to anything, including other human beings and even their own long-term welfare. Like Elliot, they merely get caught up in the desires of the moment and switch pursuits according to what grabs their attention.<sup>56</sup> To get a better flavor of the motivational patterns of a psychopath, consider this vignette, reported by a psychiatrist:

Milt’s mother couldn’t walk after surgery, and he volunteered to drive her on an errand. Their car broke down on a bridge, leaving them stranded in a dangerous spot as darkness fell. Milt offered to walk to a nearby garage to replace a blown-out fuse, the cause of the mishap. He assured his mother that he would return within fifteen minutes. The mother waited for about an hour, worrying about her son and her own

54. *Ibid.*, 51.

55. Note that, given my analysis of caring as a complex of emotions, “momentary caring” is not a possibility. Caring is a structured compound of various emotions and emotional predispositions, unfolding over time in response to relevant circumstances (see Sec. III.A).

56. If the desire of the moment happens to persist, a psychopath may pursue something for an extended period of time and give the impression of being devoted to a goal. This, I think, is what happens with psychopaths we tend to hear about in popular culture: they have strong and persistent desires and, since no deep concern about anything puts a check on their activities, they can end up doing terrible things to others in the process of pursuing what captivates their current attention. But monomaniacal attention to a goal is not the same as caring.

safety. Finally, a passerby drove her home. When Milt eventually returned, he

showed vexation at his mother for not having waited until he so belatedly got back and a bland immunity to any recognition that he had behaved irresponsibly or inconsiderately.

Milt had begun his trip to the garage with commendable haste. Shortly after leaving the bridge, he passed a cigar store. Noting . . . the afternoon's football scores . . . posted on a blackboard, he lingered for ten or fifteen minutes to check results.

During this interval he recalled that a girl he knew lived . . . in this neighborhood and decided to drop in on her. . . . He spent approximately an hour in her company. There is no evidence that any sudden sexual urge or any other strongly tempting impulse diverted our patient. He had no special liking for the girl, and no attempt was made to gain even the mildest erotic favor.

Milt chatted with the girl . . . desultorily about trifling matters. His departure followed the arrival of her date for the evening, whose rights to her company he acknowledged after a pleasant exchange of courtesies. . . .

. . . This conduct did not result from absent-mindedness, from specific amnesia or confusion, or from some attraction so enthralling or distracting as to delay or divert a person from even a mildly serious mission. He was quite aware . . . of his mother waiting on the bridge and seems to have been free from any grudge or other impulse that would influence him deliberately to offend her or cause her hardship. Missing from his realization, apparently, was the evaluation of her emotional reactions that would in another have outweighed a whim so petty as that which in Milt gained easy ascendancy.<sup>57</sup>

As hard as it is for us to enter into the frame of mind that generates such a motivational sequence, we see here a creature with interests, only interests very different in kind from our own. However fleeting and shifty, the motivations, desires, and impulses of individuals like Elliot and Milt determine the content of their contemporaneous interests, at least for the purposes of assessing their intra-agential FMS.

Since Elliot was unable to run his own affairs—even to the extent of being able to assure in the long term a stimulating environment for his motivational meanderings—it would have been appropriate for a surrogate to make decisions on his behalf. But should the surrogate be guided by the values Elliot professed before his brain injury (e.g., his devotion to his family), or should she focus on satisfying Elliot's current desires (which give no attention to the family)? Since nothing seems to matter deeply to Elliot now and his choices simply express

57. Harvey Cleckley, *The Mask of Sanity* (Augusta, GA: Cleckley, 1988), 160–62.

whims of the moment, we can reasonably surmise that Elliot doesn't have contemporaneous interests that would command full moral respect. It does appear morally appropriate to override Elliot's current motivations in the name of the value-based interests of his former self. On this view, the requirements of intra-agential fairness no longer apply.<sup>58</sup>

Note that this permission to override Elliot's preferences is not merely a matter of deeming Elliot incompetent to make his own decisions, releasing us from the obligation to respect his actual choices in the way we normally respect the choices of competent persons. After all, it is possible to run the affairs of an incompetent person according to her contemporaneous interests, as we recommended for Ms. P. What I am endorsing for Elliot is a permission to run his affairs not only against his choices but also against his contemporaneous interests. The difference is crucial, since only the latter permission would be a breach of intra-agential fairness if Elliot were still a person.<sup>59</sup>

In brief, I am suggesting that contemporary Elliot lacks FMS in the intra-agential sense, due to his loss of the capacity to care. So the capacity for caring appears necessary for intra-agential FMS. This doesn't, of course, imply that Elliot lacks interpersonal FMS—for instance, that he could be killed, subjected to medical experiments, and so on, for the

58. In cases such as the conflict between Elliot's current interest in spending his money willy-nilly and his former commitment to his family, the intuitions that Elliot's earlier interests ought to trump seem robust, and they establish the notion that Elliot lacks intra-agential FMS. However, there may be other cases in which the intuitions are more uncertain. Consider, for instance, what should be done if healthy Elliot wrote an advance directive requesting that, in case he develops ventromedial prefrontal brain damage, he should be allowed to die and not be given life-saving medical care. In this matter of life or death for the brain-damaged self, it may not seem so obvious that the earlier self ought to prevail. I suspect that this is because, in such dire circumstances, Elliot's lack of intra-agential FMS interacts with his interpersonal FMS to complicate the picture. (The somewhat different grounds for the two senses of FMS I discussed in Sec. I.C give rise to the possibility of a conflict between them: even if Elliot currently lacks intra-agential FMS, he may well have interpersonal FMS by virtue of his past or species capacities. Even if, intra-agentially, Elliot's earlier interests ought to trump, lives of beings with interpersonal FMS are morally paramount and should be supported, other things being equal.) The interaction between the two senses of FMS is a large and interesting topic, but I cannot address it here. The permissions to disregard Elliot's current interests in less dire cases establish Elliot's lack of intra-agential FMS, and this is all that's necessary for the purposes of this article.

59. Of course, Elliot's contemporaneous interests are largely determined by his desires, which get expressed in his choices, so it would often be practically impossible to override his choices without violating his contemporaneous interests. Still, it is crucial to note that we are sanctioning the latter and not merely the former. (Besides, this practical coincidence is not absolute, since an effort to assure that Elliot does not, in the long run, end up in a confining environment, such as a jail or a mental institution, would support his contemporaneous interests, but may well go against many of his explicit choices.)

sake of sufficiently weighty benefits for ordinary human beings. Elliot's interpersonal FMS might still be based on his species potential to be a full-fledged person or, more plausibly, on his past status as a full-fledged person.<sup>60</sup> However, it does imply, if my earlier analysis is on the right track, that Elliot currently lacks the capacities the moral importance of which would (on a species-capacity type of view) ultimately ground his FMS, even in the interpersonal sense.<sup>61</sup> This may be surprising to some, since Elliot's intellectual abilities, and much of what we think of as "the ability to reason," are intact.

We have now seen two important cases pointing to a crucial role of the ability to care in intra-agential FMS: young children seem to have intra-agential FMS due to their ability to care, and (some) patients with ventromedial prefrontal brain damage seem to lack intra-agential FMS because they have lost this ability. In each case, we can be reasonably confident that the ability to care makes the crucial difference in the assessment of intra-agential FMS because the ability stands out in the relevant way: in the case of young children, a lot of other abilities traditionally thought to underlie FMS are missing, and yet we grant intra-agential FMS; conversely, in ventromedial prefrontal brain damage, many of the same abilities are present, and yet we deny intra-agential FMS. If I am right that the ability to care makes this crucial difference, it must make the difference even in cases in which other plausible explanations are also available and the actual basis of intra-agential FMS is harder to pinpoint. Thus, presumably, only the retention of the ability to care explains why the interests of the slightly demented Ms. P trump those of her earlier self, and only the loss of the ability to care explains why the current interests of a person in advanced dementia can be overridden in analogous circumstances; but neither claim can be easily confirmed in a direct way.<sup>62</sup>

Most important, based on the parallels between intra-agential and

60. Note that if Elliot now has interpersonal FMS only by virtue of his former capacities, it makes sense that he lacks intra-agential FMS vis-à-vis his past, fully capacitated self.

61. I chose not to take a stand on whether potential or species membership can be the source of interpersonal FMS. But suppose for a moment that species capacities cannot confer interpersonal FMS. My arguments would now seem to imply that individuals like Milt lack interpersonal FMS because of their inability to care. Is this a plausible result? I cannot embrace it. Indeed, cases like this are part of the reason why I am unwilling to rule out approaches appealing to potential and species capacities, despite my misgivings about the available defenses of them.

62. I used the case of Ms. P to show that full Kantian autonomy is unnecessary for intra-agential FMS, and I used the case of the more severely demented patient capable only of appetitive desires to show that minimal rational capacities are insufficient. I picked these cases because they involve just the right set of capacities to help us determine what is not the basis of intra-agential FMS, but they are not well calibrated to positively pinpoint what such basis in fact is.

interpersonal FMS, we can also anticipate a crucial role of the ability to care in interpersonal FMS. That is, if the ability to care is necessary and sufficient for intra-agential FMS, this establishes the presumption that it ultimately grounds FMS, in both senses.<sup>63</sup> Yet, so far, these are somewhat tentative results, since, given the intricacies of the very concept of intra-agential FMS, our intuitions about it are not clear-cut, and the parallel between the two senses of FMS is likewise unstraightforward. To boost these intuitive results, I'll now argue more directly that the ability to care is a plausible candidate for a foundation of FMS, including intra-agential FMS.

*C. Theoretical Considerations That Favor Caring as the Basis for FMS*

It seems difficult and even paradoxical to distance oneself from one's carings, to view them as foreign or external, to feel oneself taken over by them. It is common enough to experience a strong desire that one fights off like a foreign intrusion or to have an outburst of anger with respect to which one is a "mere passive bystander."<sup>64</sup> But being a "passive bystander to one's caring attitude" is an oxymoron. Granted, it is not so hard to evaluatively distance oneself from one's carings. We do, not so mysteriously or infrequently, consider our own carings mistaken or misplaced, but even in those cases we don't normally view them as alien forces or as attitudes that we simply "find occurring within us."<sup>65</sup> If our ordinary subjective identifications are to be trusted, we cannot but be identified with what we care about. Carings always represent one's point of view as an agent: in Harry Frankfurt's terminology, they are internal to the agent. (One can, of course, give up on one's caring—decide to cease to care and follow through. What is ruled out is caring and being dissociated from what one cares about at the same time.)<sup>66</sup>

Consider a person fully and explicitly convinced that a particular caring is bad for her. She really doesn't want to care and even begins to take steps to bring herself to stop caring—because, let us suppose, the individual she cares about routinely harms her. Even in her case, so long as she has not yet succeeded in ceasing to care, she would be making a mistake if she viewed the caring as a mere happening in her psychology, not integral to who she is. This is why the predicament of a woman who wants to leave an abusive husband whom she still loves

63. In other words, if the ability to care is necessary and sufficient for intra-agential FMS, it seems to be a ground of intra-agential FMS. And given that intra-agential and interpersonal FMS are different applications of fundamentally the same moral requirements, we can presume that this ability also ultimately grounds interpersonal FMS.

64. Harry Frankfurt, "Identification and Externality," in his *The Importance of What We Care About* (Cambridge: Cambridge University Press, 1988), 59.

65. Phrase borrowed (*ibid.*).

66. Thanks to Michael Bratman for prompting this clarification.

is especially tragic. The conflict she experiences runs deep—it is a conflict within her identity. (Indeed, her own identification with the caring is likely to play a large role in why she judges the caring inappropriate or pernicious.)<sup>67</sup> By contrast, the conflict experienced by an unwilling addict intent on overcoming his addiction may be harder to resolve in practice, due to the sheer strength of the desire at issue, but because the addict doesn't care about the drug, the conflict does not "tear him apart."<sup>68</sup>

The built-in internality of our carings is particularly conspicuous when we find ourselves caring about things we never suspected we cared about, things that don't fit our prior self-image. One experiences such revelations as finding out something important about oneself, as discovering a stance toward the world that is truly one's own, and not as observing a mere occurrence in one's psychic life.

The detailed account of caring I gave earlier can help elucidate why carings are the sort of attitudes with which agents are always identified. By combining various individual emotions into a complex rational and reference-based structure, carings forge a network of rational and referential connections that support the agent's identity and cohesion over time.<sup>69</sup> Further, as we have seen, by being steadfastly emotionally attuned to the changing fortunes of the object of care, the caring subject imbues the object with importance. And once the subject cognitively grasps the object's importance, this can inspire further psychological alignments: most important, the formation of stable intentions, plans, and policies concerning the object, which keep the subject on track and thus weave the web of unified agency. (It is therefore no accident that individuals unable to care, like Elliot, break down as unified agents over time.) Because they connect various aspects of our psychology together and support our psychological unity and continuity over time, carings

67. I thank Gary Watson for this way of putting the point.

68. What if a person is both overcome or swept away by a caring, so that she has no control over it, and also judges it bad to care? Is she still identified with the caring? The abused woman in our example may be precisely in this predicament: she is caught up in her love for her husband, and yet she thinks she ought to cease to love him. This doesn't make the conflict she experiences any less deep; we would trivialize her problem if we invited her to view her love for her husband as a mere happening in her psychic life. (I am grateful to David Hills, Sibyl Schwarzenbach, and Allen Wood for pressing me to elaborate on these examples.)

69. Carings are constituted by complex rational and referential connections, which synthesize and organize disparate elements of one's psychic life, allowing for convergence of several psychological elements into a coherent cluster. In this sense, they support the agent's identity and cohesion over time. Note that, to play this role, carings need not be consistent with each other. Love and hatred of the same person, so long as each is constituted by an entire network of its own emotions and emotional episodes, will each function to support the agent's ongoing identity.

are tied to our sense of self more closely than other attitudes—they are more strongly our own.<sup>70</sup>

Once we recognize that carings are invariably internal attitudes, our intuitions about their role as grounds for FMS gain theoretical backing. It makes sense that a creature who possesses a sense of self—enough of a self that the distinction between motivations that merely occur within him and those that are truly his own applies to him—would have the underpinnings of a higher moral standing than a creature who lacks such self-delineation. The latter creature, who is simply a substratum of events occurring within the history of his mental life, and who doesn't have enough mental organization for any of these events to be truly his own, would be appropriately considered a wanton rather than a person.<sup>71</sup>

(Note that, for the relevant sense of self-delineation, the creature need not explicitly recognize attitudes as his own. Internality is not a matter of how the agent feels about aspects of his psychology and how his attitudes appear to him. Specifically, carings are invariably internal by virtue of the kind of attitudes they are, due to their structure and the consequent role they play in forging the agent's Lockean identity and cohesion over time, and not by virtue of some special additional act or attitude of identification that all agents invariably apply to their carings. The claim that carings are invariably internal is not meant as a prediction that all agents will recognize their carings as internal; it simply identifies a class of attitudes that always represent one's point of view as an agent. As such, it can apply to creatures incapable of taking their attitudes as objects of reflection, let alone of recognizing attitudes as their own or alien.)<sup>72</sup>

70. For a fuller exposition of this point, see Jaworska, "Caring and Internality."

71. My use of the term 'wanton' here is somewhat different from that familiar from Frankfurt's work. According to Frankfurt, a wanton is a creature who does not assess his own desires in any way, who is "not concerned with the desirability of his desires themselves" (Harry Frankfurt, "Freedom of the Will and the Concept of a Person," in Frankfurt, *The Importance of What We Care About*, 17).

72. Healthy agents, not alienated from themselves, will usually explicitly identify with aspects of their psychology that are truly their own and will not identify with the external aspects. Hence, we can usually tell that an attitude is internal to the agent if the agent experiences the attitude as his own, or at least if he doesn't dissociate himself from the attitude and doesn't see it as alien. Our patterns of associating ourselves with some attitudes and dissociating ourselves from others are vital inputs into any workable theory of internality, and I relied on them at the beginning of this section. Such a theory is meant to delineate which attitudes are truly the agent's own, and not merely when an agent would feel that this is the case. Once we construct the theory based on paradigm cases of agents identifying themselves with some attitudes and distancing themselves from others, we can apply it to instances in which the agent is not explicitly aware of her fully internal attitudes, and even to agents outright incapable of such awareness, as is likely the case with young children.

Because carings are very deeply our own, thanks to their role in structuring the unity and continuity of our psychology over time, they are appropriate sources of principles of action that are authentically our own, principles according to which we can govern ourselves as truly autonomous decision makers. In this way, carings function as the most elemental building blocks of the capacity for autonomy: a caring agent at least has the beginnings of a self with which to engage in self-legislation. And due to their role as starting points of autonomous decisions, caring attitudes once again emerge as plausible bases for FMS.

The intuition that the capacity to care may be sufficient for FMS was at first, recall, a source of difficulties for the Kantian approach. But now we see how the claim that the capacity to care is a foundation of FMS can be incorporated into a view that preserves the fundamental connection between autonomy and FMS, and so remains, in this sense, Kantian in spirit. If we amend the Kantian picture to the (sizable!) extent of accepting carings rather than evaluative judgments as the more fundamental starting points of autonomy, it will accord perfectly well with our intuitive assessments of intra-agential FMS in young children and patients with ventromedial prefrontal brain damage.<sup>73</sup>

73. A few points to clarify the relation of my proposal to Kantian concerns. On my proposal, the capacity to care is the foundation of FMS, where caring is a type of attitude, not necessarily associated with a specific object or content. Similarly, as I read it, the standard Kantian view treats the capacity to value as the foundation of FMS, without requiring the agent to hold specific values. In particular, just as on my proposal the capacity for altruistic caring is not necessary for FMS, so the Kantian view doesn't require the specific ability to recognize moral reasons as a precondition for FMS—the more general ability to engage in evaluative reasoning is sufficient. Thus, amoral agents and other agents whose concerns may be misguided or mistaken are handled similarly by both approaches. But what about the misanthropic, yet morally upright man whom Kant describes in *Groundwork* I? It might be thought that he fails to meet my criterion of FMS, since he is unable to bring himself to care. However, first, while this man doesn't care about humanity, there is no indication that he doesn't care about anything at all. And further, even if we imagine him as someone who really altogether doesn't care, so long as he is a standardly endowed human being, there is no reason to suppose that he lacks the capacity to care. So here, too, the two approaches yield parallel results. The reader may wonder how my view would assess the moral standing (esp. the intra-agential FMS) of a depressed person, since severe depression may appear to rob one of the very capacity to care. When a disease leads a person to lose all motivation to pursue the projects and activities she has ordinarily engaged in, this begins to look like a systematic loss of the capacity to care. However, severe depression would not be so excruciatingly difficult to bear if it left the agent altogether uncaring. Depression may skew the person's perspective, so she responds more acutely to the negative aspects of life, but depression by no means removes a person's susceptibility to emotional arousal. For one, a severely depressed person typically cares a great deal about her condition, and this is why she finds it unbearable. (Thanks to the anonymous associate editors of *Ethics* for prompting these clarifications.)

## IV. FURTHER IMPLICATIONS

The view I have developed here is meant to elucidate, in the first instance, by virtue of what capacities, be they actual, species, or potential, we can take human beings to have interpersonal FMS. It is also meant to clarify how various ethical dilemmas involving intra-agential FMS ought to be approached, especially in cases in which the task of assessing the individual's contemporaneous capacities is difficult (due, for instance, to his diminished facility with language), and some guidance is necessary as to what might be relevant for his intra-agential FMS and what the assessor should look for. But the view is also of consequence for our judgments about interpersonal FMS of creatures other than human beings.

Those who have worked with or extensively observed great apes would readily testify that these are animals capable of the cognitively sophisticated kind of caring described in this essay. The minute details of chimpanzees' and gorillas' long-term intimate bonds with one another probably provide the best evidence of this, but even less comprehensive accounts are very telling:

Consider Koko, a gorilla conversant in American Sign Language, made especially famous by a *National Geographic* cover, in which she, a 230-pound intimidating-looking animal with gigantic hands, is gently cradling a tiny kitten. The sheer fact that Koko could even go near the kitten without trampling him is remarkable enough, but there is much more to the story. Koko developed a long-term bond with the kitten, whom she named All Ball. She was able to play with Ball without frightening or injuring him, caressing him gently even though he habitually bit her following his kitten ways. She dressed him in linen napkins and hats, and signed phrases like "Koko love visit Ball" and "Soft good cat cat."<sup>74</sup> Most remarkably, she went into what could only be understood as deep mourning after Ball escaped one day and got run over by a car.<sup>75</sup> She "cried shortly after she was told of his death." Three days later, in response to her caretaker's questions about the cat, Koko signed "Cry" and "Sleep cat." "When she saw a picture of a cat who looked very much like All Ball, Koko pointed to the picture and signed, 'CRY, SAD, FROWN.'" Three months later, when asked,

74. Jane Vessels, "Koko's Kitten," *National Geographic* 167 (1985): 110–13. There are many accounts of the remarkable abilities of the great apes: see, e.g., Frans de Waal, *Good Natured: The Origins of Right and Wrong in Humans and Other Animals* (Cambridge, MA: Harvard University Press, 1996). However, Koko's relationship with All Ball provides the clearest illustration I am familiar with of animal caring.

75. Francine Patterson and Wendy Gordon, "The Case for the Personhood of Gorillas," in *The Great Ape Project: Equality beyond Humanity*, ed. Paola Cavalieri and Peter Singer (New York: St. Martin's, 1994), 67–68.

“How did you feel when you lost Ball?” Koko signed “Want,” “Open trouble visit sorry,” and “Red red red bad sorry Koko-love good.” Even after several years she still signed “Frown” and “Sorry” upon seeing pictures of Ball.<sup>76</sup>

Koko’s gentle handling of Ball shows that she was able to understand his needs and interests—very different from her own—to a remarkable degree and to respond to them adequately. A full gamut of her emotions was focused on Ball, including joy in his company, fondness, excited anticipation of his visits, and, most remarkably, grief over his loss. Unlike many other animals who may also seem to “care”—about their young, about their owners, and so forth—this gorilla’s attitudes and reactions are not stereotypical. They go far beyond a rudimentary pattern of attachment found, for example, in dogs: protest at the perceived threat of separation from the object of attachment, sadness when separated, elation at reunion. Koko’s reactions track the distinctive needs and interests of the kitten, to which Koko could not respond so well merely through biologically prescribed instincts. Rather, Koko appears to be knowingly attuned to All Ball’s needs. Given Koko’s elementary mastery of language, it doesn’t seem far-fetched to presume that she has enough cognitive ability to form the concept of importance and that her nonstereotypical responses and actions toward All Ball are dictated by her implicit perception of All Ball, the focus of her emotions, as important. By the criteria advanced here, Koko is a carer and therefore a person worthy of full moral respect.

Accordingly, the recognition that FMS ultimately rests on the capacity to care requires revisions in both currently established approaches to interpersonal FMS. The preservationist approach holds that all humans have interpersonal FMS and nonhumans do not, but once we see that the basis of FMS is less demanding than traditionally thought, it turns out that some nonhuman animals have interpersonal FMS.<sup>77</sup> Sim-

76. A journal entry over five years after All Ball’s death: “Koko comes across a picture of herself and All Ball in a photo album. K: THAT BAD FROWN SORRY [emphatic] UNATTENTION [Koko’s negation of the sign for attention, covering her eyes with her hands—clarified in the internet version of the article]” (ibid., 68).

77. Since I labeled this view “preservationist,” it appears odd to suggest that the view can be revised and still remain preservationist. My idea is this: the goal of preservationists is to justify the commonsense approach to FMS, and especially the conviction that all humans have interpersonal FMS. In doing so, they typically adopt a certain theoretical strategy: they appeal to valuable capacities of ordinary adult humans and to the value of potential and/or species possession of such capacities. I have tried to show that, once we attend to intuitions about intra-agential FMS, this preservationist approach can be sustained only if we recognize the capacity to care as an ultimate ground of FMS. The resulting view preserves the key element of common sense: the idea that all humans have interpersonal FMS. But it forces some revisions of common sense; namely, animals capable of caring, such as great apes, now also turn out to have interpersonal FMS. (Note that by

ilarly, the established revisionists must now draw moral boundaries in a new place and accept that we have special duties toward all beings (human or not) that have the capacity to care and that we lack such duties to beings (human or not) incapable of caring. True, as I forewarned, my proposal does not settle the entrenched debate between the two camps. The revisionists will deny interpersonal FMS to human beings who lack the capacity to care, while the preservationists will continue to insist that such humans have interpersonal FMS owing either to their potential for caring or to the human species' capacity to care. Still, their disagreement has been bridged to some extent, since both camps now at least have to agree on one set of cases which has hitherto exacerbated their debate: nonhuman animals capable of caring, such as the great apes, must be treated with respect normally accorded to persons.

---

suggesting this revision I do not side with revisionists in their core debate with preservationists: I still remain neutral on whether potential and/or species capacities can ground interpersonal FMS.)