

Agency and the First-Person Perspective

Lynne Rudder Baker

University of Massachusetts Amherst

It is no news that you and I are agents as well as persons. Agency and personhood are surely connected, but it is not obvious just *how* they are connected. I believe that being a person and being an agent are intimately linked by what I call a ‘first-person perspective’: All persons and all agents have first-person perspectives. Even so, the connection between personhood and agency is not altogether straightforward. There are different kinds of agents, and there are different kinds of first-person perspectives. On the one hand, all persons are agents, but not all agents are persons; on the other hand, all moral agents are persons, but not all persons are moral agents.

In this paper, I hope to sort out the relations between personhood and agency. In the first section, I’ll discuss general aspects of agency. Then, I’ll briefly summarize my view of persons and first-person perspectives. Next, I shall draw out the implications of putting together the discussions of agency and first-person perspectives. Finally, I’ll take a closer look at moral agency.

Aspects of Agency

There are four intertwined concepts to explore:

agency, action, intentional explanation, and practical reasoning. Let's begin with agency: What is an agent? An agent is an entity that is able to do things, where we may understand the relevant sense of 'doing things' as follows:

(DT) An entity x does something [in the sense relevant to agency] iff x brings about something that can be adequately explained only by reference to x's beliefs, desires or intentions.

Anything that has the ability to do things in the sense of (DT) is an agent. The things that agents do may be characterized correlatively as actions.¹ (Some of the things that we do—e.g., digesting food or growing older—are not explainable by beliefs, desires and intentions and hence are not actions at all; such things are not relevant to this discussion.) Actions are things that can be done in conformity with (DT), and agents are entities that can do things in conformity with (DT). If someone does something that conforms to (DT), then she does something that is explainable only in terms of her attitudes. I'll call explanations that explain actions in terms of attitudes 'intentional explanations'.

Actions that intentional explanations explain are manifestations of agency. Suppose that you are driving a

¹ Actions are constituted by events. Actions and events are both property-exemplifications.

car and see a child running into the road, and you swerve off the road. Your swerving off the road manifests your agency and has an intentional explanation (you wanted to avoid hitting the child). However, if instead you had had a heart attack that caused you to swerve off the road, your swerving would not have had an intentional explanation, and would have been no action at all. What makes an intentional explanation explanatory is this: If the agent had had different attitudes, she could have avoided doing what she did. (If you had wanted to hit the child, you would not have swerved.)

What connects the agent's attitudes to her actions in intentional explanations is that the agent's attitudes can be used in—perhaps primitive—practical (means-end) reasoning that concludes with the agent's acting. So, we have the following thesis:

(AE) If x brings about something that can be adequately explained only by reference to x's beliefs, desires or intentions, then x can engage in primitive practical (means/end) reasoning.

From (DT) and (AE), it follows that the ability to do things [in the sense relevant to agency] is conceptually tied to the ability to engage in primitive practical (means/end) reasoning. That is,

(PR) If x does something [in the sense relevant to agency], then x can engage in (at least) primitive practical (means/end) reasoning.

The explanatory power of intentional explanations derives from the fact that good intentional explanations capture the practical reasoning of the agent from her point of view: they give the reason that she did what she did.

The practical reasoning that intentional explanations capture may be simple, hardly worth spelling out: I want to get warmer and believe that the best way to get warmer is to move closer to the fire; so I move closer to the fire. Let me make four further points about practical reasoning:

First, practical reasoning is always first-personal: The agent reasons about what to do on the basis of her own first-person point of view. It is the agent's first-person point of view that connects her reasoning to what she actually does. Nevertheless, the agent need not have any first-person concept of herself. A dog, say, reasons about her environment from her own point of view.² She is at the origin of what she can reason about. She buries a bone at a

² Vonk and Povinelli argue that chimpanzees' minds are limited to their environment. J. Vonk and D. J. Povinelli, "Similarity and Difference in the Conceptual Systems of Primates: The Unobservability Hypothesis," in *Comparative Cognitive Experimental Explorations of Animal Intelligence*, E. Wasserman and T. Zentall, eds. (Oxford: Oxford University Press, 2006): 363-387.

certain location and later digs it up. Although we do not know exactly what it's like to be a dog, we can approximate the dog's practical reasoning from the dog's point of view: Want bone; bone is buried over there; so, go over there. The dog is automatically (so to speak) at the center of her world without needing any self-understanding.

Second, as the dog example also illustrates, practical reasoning does not require that the agent have a natural language. Although, as I'll argue later, there is an important difference with respect to agency between entities that have a language and those that do not—theorists and ordinary people alike successfully explain behavior by attributing practical reasoning to creatures that have no natural language. Without assuming that such creatures engage in primitive practical reasoning, we are left with no explanation of their behavior at all.

Third, the practical reasoner—primitive or not—may be unaware that she is reasoning. Even beings like you and me, who do have language and can engage in sophisticated practical reasoning, are often unaware of our own reasoning. When I get to the Kardinal Köng Haus where the conference is held, I open the door. And I open it intentionally: I want to go inside and believe that the best way to get there is to open the door. Although I do not think about my attitudes or the door, my action is explained by

the intentional explanation and the simple practical reasoning that it captures.

Fourth, some actions are accidental or unintended, but insofar as they are actions at all, their explanations invoke attitudes used in practical reasoning. If, as I moved closer to the fire, I accidentally stepped on your foot, then my stepping on your foot was an action of mine—an unintended action. What makes my stepping on your foot an action at all is that it is explained in part by my attitudes: I wanted to move closer to the fire and I was being careless about where I stepped. That is, one of the things I did (e.g., moving closer to the fire) was intended, and another thing that I did (e.g., stepping on your foot) was not intended. But both my moving closer to the fire and my stepping on your foot were actions of mine.

Or consider a slightly more complicated example: suppose that a nurse who was supposed to bathe a toddler accidentally scalded the toddler in her care. Suppose that the nurse did not intend to scald the baby but mixed up the hot-water-tap and the cold-water-tap. Even though she did not do what she intended, her scalding the baby still has an intentional explanation: She wanted to give the baby a bath, and believed that turning the hot-water-handle 90 degrees and the cold-water-handle 180 degrees was the best way to give the baby a bath. So, she turned one handle

90 degrees and the other handle 180 degrees. However, since she mixed up the taps, she ended up scalding the baby. So, the scalding—although wholly accidental—needs an intentional explanation that invokes a mistaken belief about which tap was hot. One cannot do anything by accident unless one can also do things intentionally. In this way, we may also see mistakes as manifestations of agency.

In sum, we have several interrelated concepts: agency, action, practical reasoning and intentional explanations. These intertwined concepts may be characterized as follows:

An agent =_{df} an entity that is able to do things (as in (DT)) that are explained only in terms of her primitive practical reasoning (as in (PR))

An action =_{df} something that an agent brings about (as in (DT))

An intentional explanation =_{df} an explanation of an agent's action in terms of the agent's intentional attitudes used in practical (means-end) reasoning (as in (PR))

Minimal Agency

Agency comes in various degrees. The generic 'agent' as I have just characterized her might be thought of as a

minimal agent. So,

minimal agency =_{df} the ability to do things explainable only by attitudes used in practical (means-end) reasoning.

Rational Agency

Some agents are more than minimal agents. A rational agent (as I'll call her) is a minimal agent who has the second-order ability to evaluate her beliefs and desires. She not only has beliefs, desires and intentions, but also she knows that she has beliefs, desires and intentions. A rational agent not only reasons about what to do, but also can rank preferences and goals, evaluate her beliefs and try to resolve conflicts among them. She can deliberate, not only about what to do, but also about what beliefs, desires and intentions she ought to have. She can have second-order desires about her desires. She can decide what kind of person she wants to be, and make efforts to achieve the character that she wants to have. Not only can she identify her desires as her own, but also she can have (second-order) desires about the (first-order) desires that she wants to have.³ So,

rational agency =_{df} minimal agency + possession of appropriate second-order attitudes (e.g., desires

³ See Harry Frankfurt, "Freedom of the Will and the Concept of a Person," *Journal of Philosophy* 68 (1971): 5-20.

about one's own desires)

Moral Agency

To be a moral agent is to be accountable for what one does, to be subject to judgments of praise and blame. Moral agency requires not only that one have second-order attitudes generally, as a rational agent does, but also that one have a very specific second-order attitude: To be a moral agent, one must be able to appreciate the fact that she does things and has done things in the past (as in (DT)).

moral agency =_{df} rational agency + realization that
one does things and has done things in the past

How much of what she does is a moral agent morally responsible for? On my view, *a moral agent is defeasibly responsible for everything that she does* (in conformity with (DT)). By saying that a moral agent is defeasibly morally responsible for everything she does, I mean that ordinarily a moral agent is morally responsible for everything that she brings about that can be adequately explained only by reference to her beliefs, desires or intentions. In the absence of some special defeating condition, a moral agent is morally responsible for her actions. Being morally responsible for one's actions is the default state for agents who realize that they do things and have done things in the past.

Traditionally, the conditions that defeat a charge of moral responsibility are ignorance and interference—an agent’s ignorance of relevant nonmoral circumstances of the action or interference by some outside force or agent with what the agent does. There are standard examples of each kind of defeating condition. An example of ignorance of a relevant nonmoral feature is ignorance that a bridge in front of the school bus you are driving has been swept away. Suppose that you are an attentive school bus driver, careful of your young charges. On one unhappy day, you plunge the school bus into the raging river. Your driving the bus into the river was an action of yours: it is intentionally explained by your attitudes: you wanted to get the children home safely, and you thought that driving as you did was the best way to get the children home safely (as in the past). Nevertheless, you are not morally responsible for the injury to the children. There was no way for you to have known that the bridge had just been struck by lightning and given way. Your ignorance of the circumstances defeats the charge of moral responsibility.

Another kind of defeating condition is interference with what the agent does. A well-worn example of interference as a defeater is the farfetched case of the mad neuroscientist who manipulates the agent’s brain in such a way that the agent does something that she would not otherwise have done. Coercion and compulsion also count

as defeating conditions that interfere with the agent's action. Unfortunately, but inevitably, we cannot enumerate all the things that defeat charges of moral responsibility. This is a broadly Aristotelian view of moral responsibility, and I believe that it is correct.

I have discussed three varieties of agency: minimal agency, rational agency, and moral agency. Let's now turn to persons and first-person perspectives.

Persons and First-Person Perspectives

What, then, is a person? A person, I believe, is a being who has a first-person perspective essentially. You and I have robust first-person perspectives; we have a unique conceptual ability to think of ourselves as ourselves—not just to discriminate between ourselves and others, but to conceptualize the distinction between ourselves and others. We have the conceptual ability to think of ourselves without the use of any name, description or other third-person referring device. I can distinguish between the thoughts 'I am glad that I'm happy' and 'I am glad that LB is happy'. I could still be glad that I was happy even if I had amnesia and did not know that LB was happy. To distinguish thoughts about myself as myself from thoughts about myself as LB or as the tallest woman in the room or as the person in the mirror, I must have a robust first-person perspective.⁴

⁴ We can signal attribution of a robust first-person perspective

We are not only subjects of experience, but we also know that we are subjects of experience. Not only do we have beliefs and other attitudes; we also know that we have them. We not only interact with things in our environment, but we also know that we do. A robust first-person perspective enables us to know that we (we, ourselves) interact with things in our environment, and to know that we (we, ourselves) are subjects of experience. A robust first-person perspective is a property that requires sophisticated conceptual ability.

I anticipate a strong objection: A human infant does not have the conceptual resources to think of herself as herself. She is a subject of experience and she interacts with things in her environment, but she does not have a first-person concept of herself as herself. Does this imply that human infants are not persons, that a person does not come into existence until there is a conceptual ability to think of herself as herself?

No. Although we language-users have robust first-person perspectives, human persons begin existence with rudimentary first-person perspectives.⁵ A being has a

either to oneself or to someone else by the device of an asterisk '*'. To attribute a robust first-person perspective to myself as in 'I believe that I am in Vienna,' I'll say 'I believe that I* am in Vienna; to attribute to you a robust first-person perspective as in 'You believe that you are in Vienna,' I'll say 'You believe that you* are in Vienna.'

⁵ There is no precise moment when a person comes into existence. Everything in the natural world comes into existence

rudimentary first-person perspective if and only if (i) it is a sentient being, (ii) it has a capacity to imitate, and (iii) it behaves in ways adequately explainable only by attribution of beliefs, desires and intentions. These are the kinds of properties—person-making properties, unlike, say, the property of having a heart—that we specifically associate with being a person. Note that I do not define rudimentary first-person perspectives in terms of potentiality. The property of being a sentient, intentional entity with a capacity to imitate is not to have some kind of potential. The term ‘Capacity’ should be understood as an in-hand capacity, not just a capacity to develop a capacity.⁶

There is a good deal of evidence from developmental psychologists that human newborns meet these conditions for having rudimentary first-person perspectives. (i) Human infants are obviously sentient. (ii) They have been shown to naturally imitate tongue protrusions and mouth openings as young as 42 minutes old.⁷ (iii) And human infants display obviously defensive behavior in response to an approaching gradually—persons, organisms, artifacts, artworks.

⁶ The distinction between an in-hand capacity and a capacity to have a capacity is made vivid by Robert Pasnau in *Thomas Aquinas On Human Nature* (Cambridge: Cambridge university Press, 2002): 115

⁷ “These data directly demonstrate that a primitive capacity to imitate is part of the normal child’s biological endowment,” say Alison Gopnik and Andrew N. Meltzoff in “Minds, Bodies, and Persons: Young Children’s Understanding of the Self and Others as Reflected in Imitation and Theory-of-Mind Research,” in Parker, Mitchell and Boccia, eds., *How Babies Think: The Science of Childhood* (London: Weidenfeld and Nicholson, 1999): 30.

object.⁸ Since they are sentient, intentional beings with an ability to imitate, human infants have rudimentary first-person perspectives and are persons.

This conclusion invites another objection: Many nonhuman mammals—not only nonhuman primates, but also dogs and cats and other animals—also meet the conditions for having rudimentary first-person perspectives. If the onset of a rudimentary first-person perspective in a human organism marks the beginning of a person, why does the onset of a rudimentary first-person perspective in a chimpanzee, say, *not* mark the beginning of a person? The answer is that chimpanzees are not of kinds that support robust first-person perspectives.⁹ What distinguishes a human infant from a chimpanzee is that the human infant is *of a kind* that typically supports a robust first-person perspective. The human infant is a person constituted by a human organism, as I'll explain. The chimpanzee is simply an organism.

The relation between the human organism and the human person (or human being) is what I call 'constitution':

⁸ The defensive behavior by 10-day-old infants had three components: eye widening, head retraction, and interposing of hands between face and object. *Development in Infancy*, T.G.R. Bower, (San Francisco: W.H Freeman and Co., 1974): 84.

⁹ Daniel J. Povinelli and Christopher G. Prince report that "there is little evidence that chimpanzees understand anything at all about mental states." "When Self Met Other," in *Self-Awareness*, Ferrari and Sternberg, eds.

Constitution is a perfectly general relation of unity-without-identity between things of two basically different kinds: A piece of bronze and a statue; a piece of plastic and a driver's license; a human organism and a human person. An organism is essentially biological, but not essentially first-personal; a person is essentially first-personal, but not essentially biological. (A post-mortem person could have a nonbiological resurrection body.)¹⁰ So, a person is not identical to the organism that constitutes her.

The difference between a human infant and a chimpanzee may be made clearer by a distinction between having a property nonderivatively and having a property derivatively.¹¹ An entity *x* has a property *F* derivatively if and only if *x* has *F* in virtue of its constitution relations to something that has *F* nonderivatively. When a human organism develops a rudimentary first-person perspective, a new being—a person—comes into existence. The human organism then constitutes a person, who has the rudimentary first-person essentially. The organism then has the rudimentary first-person perspective derivatively—in virtue of constituting something (a person) that has it nonderivatively. When a chimpanzee develops a first-person perspective, no new being comes into existence. So,

¹⁰ See my "Persons and the Metaphysics of Resurrection," *Religious Studies* 43 (2007): 333-348.

¹¹ Not all properties may be had derivatively. See *the Metaphysics of Everyday Life*.

the chimpanzee's rudimentary first-person perspective is not derivative—that is, it does not depend on any constitution relations. Fundamentally, a chimpanzee is an animal—a being with a first-person perspective contingently; fundamentally, a human infant is a person—a being with a first-person perspective essentially.

To put it metaphorically, when a human organism develops a rudimentary first-person perspective, she passes it off to the person that the organism comes to constitute; and the organism has the rudimentary first-person perspective derivatively (in virtue of constituting a person) and contingently (in virtue of existing before being able to support a first-person perspective). In contrast, the person has the rudimentary first-person perspective nonderivatively and essentially. When the chimpanzee develops a rudimentary first-person perspective, the chimpanzee has nothing to pass it off to; nothing further is constituted.

A rudimentary first-person perspective is a functional property that has different roles in human and nonhuman organisms. In nonhuman organisms, a rudimentary first-person perspective has a purely biological role in survival and reproduction. In human organisms, a rudimentary first-person perspective has an additional role: to bring into existence a person that the organism then constitutes. The

newly-existing person has a rudimentary first-person perspective nonderivatively, and the human organism has it derivatively. Typically, the newly-existing person goes on to develop a robust first-person perspective.

So, although a mature person has a robust first-person perspective, a person comes into existence when a human organism develops a rudimentary first-person perspective. In short,

(HP) x constitutes a human person at t if and only if x is a human organism (nonderivatively) and x has a rudimentary or robust first-person perspective at t.

(HP) concerns only human persons, who are necessarily embodied (though they do not necessarily have the bodies that they in fact have). Perhaps there are nonhuman persons—silicon persons (constituted by fusions of silicon items) or God (constituted by nothing at all). (HP) is silent about nonhuman persons.

I suspect that the development of a robust first-person perspective goes hand-in-hand with learning a natural language. Certainly, the evidence we have of robust first-person perspectives is linguistic: The being who asserts, “I wish that I were a movie star,” not only refers to herself (by means of the first occurrence of ‘I’), but also attributes to

herself a first-person reference (by means of the second occurrence of 'I'). This kind of assertion manifests a robust first-person perspective.

As a toddler learns more about her physical and social environment, and as she learns to talk, her conceptual abilities explode. She can entertain many more and different kinds of thoughts. When she discovers that she is a subject of experience among other such subjects, she acquires a robust first-person perspective. But the person she is has existed since the organism that constitutes her developed to the point of supporting a rudimentary first-person perspective.

In sum, persons have first-person perspectives essentially. Persons are constituted by organisms that have first-person perspectives nonessentially.

Persons and Agents

The big dividing line among agents is between those who have robust first-person perspectives and those who have only rudimentary first-person perspectives. I shall show that all agents have rudimentary first-person perspectives and that all beings with rudimentary first-person perspectives—human or not—are minimal agents; but only beings with robust first-person perspectives are rational and moral agents.

Let me produce some evidence for the conclusion that all persons and some nonpersons are minimal agents.. First, consider some evidence that human infants (persons with rudimentary first-person perspectives) are minimal agents. Infants as young as two months old engage in problem-solving behavior. When an experimenter sets up a light so that an infant can turn it on by moving her head to the left, most infants learn quickly that moving their heads to the left turns on the light. They turn on the light over and over and then the rate of left-head-turning drops dramatically. Then the experimenter changes the contingency to right-head-movement. When the infant turns her head left again, the light fails to come on. Then the infant makes a rapid succession of left-head-turnings. Sooner or later the infant turns her head to the right and the light comes on. This is followed by a high rate of right-head-turnings, which then subside. With changes in the contingencies, an infant can master complex series of movements, such as right-right-left-left.

The infant seems to have little interest in the light itself; she merely glances to see whether it is on, without paying further attention. As one experimenter remarked, "The infant seems to be testing hypotheses and trying out sequences of movements in order to discover which one operates at the moment. When the correct sequence is discovered, it is tested a few times and then dropped....It is

quite obvious from the behavior of the infants that the light source is not the motivating factor...[It] seems that the pleasures of problem solving are sufficient to motivate behavioral and mental activity in young infants.”¹² As problem-solvers, human infants are minimal agents.

Now consider some scientific evidence that some nonhuman animals are minimal agents. The journal *Science* reported on work that showed that bonobos and orangutans not only can use tools to get a fruit treat from a mechanical apparatus, but also they can plan ahead.¹³ They were first trained to use a tool to get a fruit treat from a mechanical apparatus. Then, the apes were given tools, some suitable and some unsuitable for the task of getting the fruit; next, they were taken out of the test room into a waiting room and brought back to the test room after an hour. Significantly more often than predicted by chance, the apes took with them a suitable tool for getting the treat and brought it back with them after the waiting period.

So, we have strong evidence that both human infants (persons) and nonhuman higher animals (nonpersons) are

¹² T.G.R. Bower, *Development in Infancy* (San Francisco, W.H. Freeman and Co., 1974): 8-9. Thanks to Gareth B. Matthews for showing me this source.

¹³ “Apes Able to Think Ahead,” *Scientific American.com*, May 19, 2006. http://www.sciam.com/print_version.cfm?articleID=000C4698-F1DE-146C-B1DE83414B... Accessed 5/23/06. The apes “selected, transported, and saved a suitable tool not because they currently needed it, but because they would need it in the future.”

agents. This evidence gives empirical backing to my definitions of ‘rudimentary first-person perspective’ and ‘minimal agent’, together with the thesis (PR).¹⁴ From these definitions it follows that any entity that has a rudimentary first-person perspective—whether human or not—is a minimal agent. The evidence suggests that the definitions are in fact satisfied by actual entities.

From the point of view of my theory of persons, there is an enormous difference between minimal agents on the one hand, and rational agents and moral agents on the other hand. Rational and moral agency require second-order attitudes; anyone who has a second-order attitude has a robust first-person perspective. It follows then that all rational and moral agents have robust first-person perspectives. Moreover, the converse is true for rational agents: all beings with robust first-person perspectives are rational agents. But I’m not sure about the converse for

¹⁴ Here is the argument:

1. If x has a rudimentary first-person perspective, then x is able to engage in behavior explainable only by attribution of beliefs, desires and intentions. (by definition of ‘rudimentary first-person perspective’)

2. If x is able to engage in behavior explainable only by attribution of beliefs, desires and intentions, then x is able to engage in practical (means-end) reasoning. (by PR)

3. If x is able to engage in behavior explainable only by attribution of beliefs, desires and intentions used in practical (means-end) reasoning, then x is a minimal agent. (by 1,2 and definition of ‘minimal agent’)

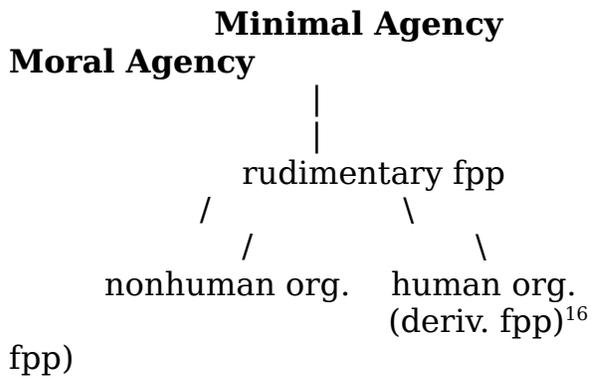
∴ 4. If x has a rudimentary first-person perspective, then x is a minimal agent.

moral agents. I don't know whether all beings with robust first-person perspectives are moral agents or not. If it is possible to have a kind of amnesia in which one does not know that she has done things in the past, although she still has other kinds of second-order thoughts, then a person so afflicted would not be a moral agent.¹⁵

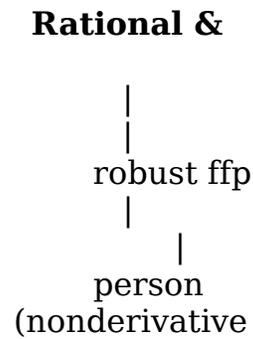
In sum, entity x is a minimal agent if and only if x has a rudimentary first-person perspective; x is a rational agent if and only if x has a robust first-person perspective, and x is a moral agent only if x has a robust first-person perspective. Since all rational and moral agents have robust first-person perspectives, all rational and moral agents are persons. I'll try to sum up these relations in three figures:

Fig. 1 illustrates that all and only beings with rudimentary first-person perspectives (persons or not) are agents; Fig. 2 illustrates that all rational and moral agents have robust first-person perspectives. Fig. 3 illustrates that all persons (beings with first-person perspectives essentially) are minimal, rational and/or moral agents.

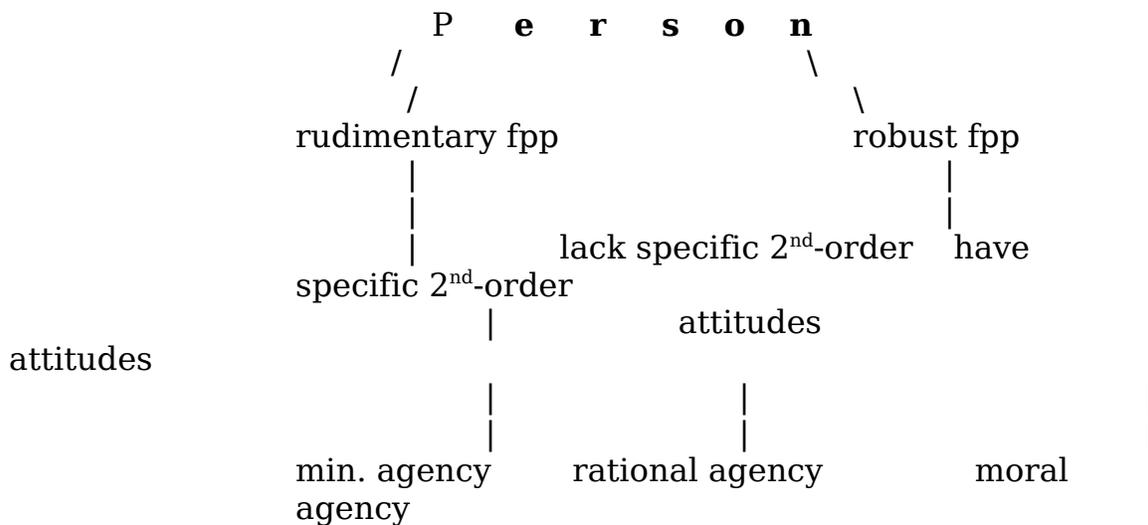
¹⁵ However, a schizophrenic with a robust first-person perspective and the knowledge that she had done things in the past would still be a moral agent; but if for everything she did, there was a morally defeating condition, there may be nothing for which she is morally responsible.



(fig.1)



(fig. 2)



(fig. 3)

Moral Agency

I want to conclude by taking a closer look at moral agency. I believe that moral agency is closely connected to having a robust first-person perspective, that beings with robust first-person perspectives are morally responsible for

¹⁶ That is, the human organism has a (rudimentary) first-person perspective derivatively in virtue of constituting something (a person) that has it nonderivatively.

what they do unless some defeating condition is present. Other philosophers believe that moral responsibility requires much more. They think that in order to be morally responsible, one must have ultimate control over one's actions. This opposing view, which I'll call 'libertarian free will',¹⁷ requires that we have *originative* power over the choices and actions that we are responsible for.¹⁸ This originative power requires that our choices and actions (at least those for which we are morally responsible) not have their origin in anything beyond the agent's control. As Roderick Chisholm, a prominent libertarian put it, "If we are responsible, and if what I have been trying to say is true, then we have a prerogative which some would attribute only to God: each of us, when we act is *a prime mover unmoved*. In doing what we do, we cause certain events to happen, and nothing—or no one—causes us to cause those events to happen."¹⁹

¹⁷ 'Libertarian free will' is shorthand for 'a libertarian conception of free will', just as 'Newtonian simultaneity' is shorthand for 'a Newtonian conception of simultaneity'. Peter van Inwagen has complained vehemently about my use of a term like 'libertarian free will'; so, I am stipulating what 'libertarian free will' is to denote. Since 'free will' is a term of philosophical art, it does not (*pace* van Inwagen) have an unambiguous pre-theoretical meaning. 'Libertarian free will' and 'compatibilist free will' are as innocent as 'Newtonian simultaneity' and 'Einsteinian simultaneity.' All these terms are clear and unambiguous.

¹⁸ See Robert Kane, *The Significance of Free Will* (New York: Oxford University Press, 1998): 4.

¹⁹ Roderick Chisholm, "Freedom and Action," *Freedom and Determinism*, Keith Lehrer, ed. (New York: Random House, 1966): 11-44.

Elsewhere, I have argued at length that no human being has this kind of libertarian power.²⁰ I do not believe that moral responsibility requires any uncaused events, or an agent's ultimate control over the actions and decisions for which she is morally responsible. It is a fantasy to think that one is independent of her upbringing, character, genes, and environment. We are the origin of the actions for which we are morally responsible only in the following limited way: we beings with robust first-person perspectives can deliberate and decide what to do; but we are not the origin of ourselves or of our characters or of the attitudes that make up our deliberations. Our first-person perspectives afford us the limited control over our characters and attitudes that we have; but, *pace* libertarianism, we are not the ultimate sources of what we do.²¹

Here let me just give a couple of examples of actions, which seem to me to be paradigmatic actions whose agents are morally responsible. Yet—if the libertarian view is correct—the agents in the examples are not morally responsible for their actions.

²⁰ See my "Moral Responsibility Without Libertarianism," *Noûs* 40:2 (2006): 307-330, and "Why Christians Should Not Be Libertarians: An Augustinian Challenge," in *Faith and Philosophy* 20 (2003): 460-478.

²¹ Chisholm, of course, is not alone as a libertarian. Other prominent libertarians include Robert Kane, Timothy O'Connor, Alvin Plantinga, Eleonore Stump, and Peter van Inwagen.

First, the leader of a certain tribe—call it the Xs—wants to avenge ancient wrongs inflicted on the ancestors of the Xs by the ancestors of the Ys, and he orders the Xs to massacre the Ys. The source of his order to massacre the Ys lies in the ancient wrongs inflicted on his ancestors. If his beliefs, desires, upbringing, character, understanding of history, and so on are sufficient for his giving the order, then he does not have the ultimate control that libertarians require for moral responsibility; and if libertarianism is true, then we should not hold him morally responsible for the massacre that he ordered.

Second example, a philanthropist believes that if she gives millions of dollars to a certain village to purify its water supply, she will help many people. She wants to help people, and since she isn't interested in getting credit, she gives the money anonymously. In the causal history of her act of charity is her belief that that village had impure drinking water. That belief derived from what she read, and what she read was caused in part by the actual horrible conditions in the village—over which she had no control. So, if her beliefs and so on are sufficient for her charity, the philanthropist is not the ultimate originator. She has no ultimate control since the sufficient conditions for her charity—her beliefs, desires, character, wealth, and so on—had their sources outside her. Although it is natural to take the philanthropist to be praiseworthy, if libertarianism is

true, then she is not praiseworthy. Her philanthropy had sources beyond her control.

So, I do not think that appeal to libertarian free will is a viable way to understand moral responsibility. Beings with robust first-person perspectives who act knowingly and without interference are usually morally responsible for what they do. A robust first-person perspective gives us limited control over our desires. Unlike other animals, we can discover what our goals are; we can evaluate and try to change them. (Agents with only rudimentary first-person perspectives can act on their desires, but they cannot try to change their desires.) The partial control over our desires that is afforded by our robust first-person perspectives has the consequence that even if determinism is true, we are not just puppets; we have a hand in what causes our actions. We can modify our first-order desires that produce our intentions. We can try to become the persons we want to be, even though our wanting to be a certain kind of person, no doubt has causes beyond our control.

The heart of libertarianism is that an action for which one is morally responsible has no sufficient condition beyond the control of the agent. Without trying to analyze the concept of moral responsibility, let me just suggest a sufficient condition for an agent to be morally responsible for a choice or action X. This sufficient condition is a

modification of Harry Frankfurt's condition.²² I call it the 'Reflective-Endorsement' view:

(RE) An agent S is morally responsible for a choice or action X if X occurs and:

- (i) S wills X,
- (ii) S wants to will X,
- (iii) S wills X because she wants to will X, and
- (iv) S would still will X even if she knew the provenance of her wanting to will X.²³

That is, an agent is morally responsible for doing something even if her reasons for doing it were traceable back to factors beyond her control—provided that she would still have acted in the way that she did if she had known where her reasons had come from. The whole view of (RE) is studded with presuppositions of a robust first-person perspective. Conditions (ii), (iii), (iv) explicitly require that S have a robust first-person perspective.

(RE) shows that a sufficient (but not necessary) condition for an agent to be morally responsible for an

²² Harry Frankfurt, "Freedom of the Will and the Concept of a Person," *Journal of Philosophy* 68 (1971): 5-20.

²³ To say that S would still will X even if she knew the provenance of her wanting to will X is to say that if S had known that her wanting to will X had causal antecedents that traced back to factors beyond her control, and if she had known of those causal antecedents that they were in the causal history of her wanting to will X and that they were beyond her control, then: she still would have willed X, and wanted to will X, and willed X because she wanted to will X.

action is that the agent endorse the beliefs and desires on which he acts; he affirms them as his own. When he acts on such desires, he—the agent himself, not just some events occurring in him—is involved in what he does.²⁴ If I can say, “These desires reflect who I am, and this is the kind of person I want to be,” then surely I am morally responsible for acting on those desires—whether the sources of the desires were beyond my control or not. This makes determinism irrelevant to moral responsibility.

Conclusion

So, personhood and agency are related intimately, but not straightforwardly. We have seen that all persons are agents, but not all agents (e.g., chimpanzees, dogs) are persons; and we have seen how all rational and moral agents are persons, but not all persons (e.g., human infants) are rational and moral agents.

What justifies these complications is the fact that the Constitution View simultaneously captures the seamlessness of the animal kingdom and the ontological uniqueness of persons. Rival views either affirm the seamlessness of the animal kingdom and leave out the ontological uniqueness of persons (as naturalism typically does) or affirm the ontological uniqueness of persons and

²⁴ On the Constitution View, the agent has irreducible person-level states constituted (perhaps) by brain states with which the person-level states are not identical.

leave out the seamlessness of the animal kingdom (as dualism typically does). I know of no view besides the Constitution View that affirms both the unity of the animal kingdom and the ontological uniqueness of persons.²⁵

²⁵ I am indebted to Gareth B. Matthews, with whom I taught a graduate seminar on the first-person perspective, and from whom I have learned so much.