

ORIGINAL ARTICLE

Caretakers of value: A theory of human personhood

Philip Woodward 

Department of Philosophy, Dunleavy Hall, Niagara University, Lewiston, New York, USA

Correspondence

Philip Woodward, Department of Philosophy, Dunleavy Hall, Niagara University, Lewiston, NY 14109, USA.
Email: pwoodward@niagara.edu

Funding information

John Templeton Foundation

Abstract

According to a traditional view, humans are superior to their non-human terrestrial companions because they alone are “rational animals.” Although the traditional view is presupposed by our social and legal institutions, it has been called into question by modern science: Darwin himself claimed that humans differ in degree rather than in kind from animals, and recent discoveries in comparative animal cognition have seemed to confirm Darwin’s assertion. Sustaining the traditional view in light of these discoveries calls out for a careful comparison of the human mind with its precursors, the vertebrate mind and the mammalian mind. Psychological capacities shared across the animal kingdom are repurposed in humans, rendering humans uniquely capable of rationality. To be rational is to respond appropriately to value. Humans as such are not rational animals, but have the potential to become rational animals. This potential marks a morally relevant difference in kind between humans and non-human animals.

1 | PERSONHOOD: THE TRADITIONAL VIEW VS. CONTEMPORARY BIOLOGY

Human beings are persons.¹ Other animals are not. This distinction marks a morally significant boundary among sentient beings. Or so it is widely believed. Is this belief still tenable in light of evolutionary and comparative psychology? That is the question I explore in what follows.

¹Mature and healthy adult humans, that is. What to say about non-mature or seriously disabled humans? I return to this question in the final section.

The concept of personhood in Western philosophy has a fascinating history.² When Boethius declared in the *Theological Tractates* that a person is “an individual substance of a rational nature,” he was drawing together threads from three ancient intellectual traditions: (a) from the Greeks (especially Aristotle), the doctrine that the rational soul is the form (*morphe*) of the human body, in contrast to the merely “sensate” souls of animals; (b) from the Romans (especially Cicero and Seneca), the use of the term “persona” to express the uniquely dignified role that human beings play on the world’s stage (“persona” originally denoted the thespian’s mask); and (c) from the Christian theologians, an emphasis on the underlying, unshareable, and enduring individuality of the person, an emphasis which had been developed in the context of theological reflection on Christ’s two natures and the shared nature of the three Trinitarian persons. (It was the Greek theologian’s term *hypostasis*—literally, *that which underlies*—that Boethius translated as “*persona*.”) Thus, antiquity bequeathed to subsequent Western philosophy a personhood-concept with at least three crucial components: rational nature (from the Greeks), dignity (from the Romans), and selfhood—that is, the enduring substratum of that dignified nature (from the Christians). Later contributors to the philosophy of personhood, such as Gracian, Aquinas, Locke, and Kant, can be read as attempting to get clear on the nature of these elements and how they are interrelated. Kant’s is the culminating discussion in this tradition. Kant expounds rationality in terms of *moral autonomy*—the ability to understand moral oughts and to regulate one’s actions accordingly. Persons are beings capable of moral autonomy. But to be capable of moral autonomy is also to be a bearer of dignity, understood as a type of worth different in kind from the worth of all other beings.

Throughout its history, the philosophy of personhood has focused on the shared rational nature and high moral status of persons, rather than on the *non*-rational natures and *lower* moral statuses of non-persons. But the contrast with non-human animals has always been implied and occasionally made explicit (Descartes is a famously extreme case). Western philosophy up until the 19th century affirms, nearly unanimously, that humans, because they are rational, are as different from all animals as all animals are from plants, and, in virtue of this difference in kind enjoy a difference in rank.³ I’ll call this “the traditional view.” The traditional view in both its metaphysical and moral dimensions endures as moral common sense in the West. It explains, for example, why slavery has been legally abolished but carnivorousness has not.⁴

But with the rise of modern biology in the 19th century, the traditional view began to be challenged. A watershed was Darwin’s famous assertion in *The Descent of Man* that humans differ from animals only in degree and not in kind. Darwin meant by this that every psychological capacity found in humans (a) has precursors in non-human animals and (b) came into being via gradual rather than abrupt evolutionary steps. Insofar as the traditional view requires something like the Aristotelian claim that humans have souls of a different genus from those of animals,

²There are a number of fine recent tellings of this history. I am especially indebted to Zagzebski (2001), Spaemann (2006), and the essays in LoLordo (2019).

³An important exception is Porphyry, whose treatise in defense of vegetarianism proves the rule, written as it is in an iconoclastic tone of voice.

⁴Which is not to say that the traditional view is sufficient to justify carnivorousness. To say that humans exhibit elevated moral status does not entail that animals are of no moral significance whatever. The traditional view might be necessary but is not sufficient to justify eating animals.

then Darwin's conception of the human/animal difference is inconsistent with the traditional view.⁵

Darwin, it is important to note, was not so much announcing an empirical discovery as launching a research program. Nevertheless, the subsequent fruits of the research program Darwin launched have seemed to confirm his suspicions about the traditional view. Take any psychological capacity that is reputed to be uniquely human, and it can be found, in an attenuated form at least, somewhere in the animal kingdom. Communicating vocally? So do vervet monkeys.⁶ Parsing syntax? So do dolphins.⁷ Making tools? So do crows.⁸ Teaching the young? So do meerkats.⁹ Forming abstract concepts (such as sameness and difference)? So do pigeons.¹⁰ Remembering particular past episodes? So do rats.¹¹ Discerning other's mental states? Coordinating activities in complex ways? Responding empathetically to suffering? So do chimpanzees.¹² Some of these findings are disputed, but given how many are well-established, the advocate of the traditional view shouldn't bet on their being overturned.

Not only has modern science shown that animals possess something like rudimentary rationality, it has also shown that rationality in human beings is far from ideal. Reasoning does not influence human decision-making very often. Rather, the reasons that come to our minds are most often those that justify our beliefs and actions after the fact, and we tend to overestimate the quality of these reasons.¹³ Furthermore, non-rational factors—such as whether one has just found a coin, or whether one is in a hurry—*do* influence human decision-making to a considerable degree.¹⁴ Thus, empirical science has apparently narrowed the gap between animals and humans by both “up-grading” animal powers and “down-grading” human powers.

My aim in what follows is to defend the metaphysical dimension of the traditional view, with a brief comment on the moral dimension in the concluding section. Contra Darwin, and consistent with the findings just described, being a human means being a different *kind* of thing from any animal, for it means having a capacity for personhood that no animal possesses. Findings from comparative animal psychology and evolutionary psychology do not undermine the traditional view; rather, they flesh it out. They reveal the particular manner in which human psychology realizes personhood, a manner that repurposes the psychological resources we inherited from our ancestors who were not persons.

I begin by clarifying the notion of “rational nature” that will guide my discussion of personhood. Then, I describe minds of three very general types found in nature: the Vertebrate Mind, the Mammalian Mind, and the Human Mind. I show how the latter two types repurpose resources inherited from the type prior. I characterize the crucial differences between

⁵This is not to say, of course, that evolutionary theory in general is inconsistent with the traditional view. It is just to say that Darwin's peculiar anthropological views, which he supposed to follow from evolutionary theory, are inconsistent with the traditional view. (Thanks to an anonymous referee for inviting clarity about this point). This paper is one attempt to demonstrate the compatibility of evolutionary theory with the traditional view.

⁶Cheney and Seyfarth (1980).

⁷Herman et al. (1984).

⁸Bluff et al. (2010).

⁹Hoppitt et al. (2008).

¹⁰Wasserman and Young (2010).

¹¹Crystal (2010).

¹²Tomasello (2019), de Waal (2006).

¹³Mercier and Sperber (2017).

¹⁴See Doris et al. (2020).

the Human Mind and the Mammalian Mind in four stages: (1) the *accumulative* stage, (2) the *ampliative* stage (both of which involve only differences in degree), (3) the *additive* stage, and (4) the *transformative* stage (at which point we arrive at a difference in kind). Finally, I compare the resulting picture of the mature Human Mind with the normative ideal of rationality. I conclude that humans are indeed endowed with the unique capacity to be “caretakers of value”—to be persons.

2 | THE RATIONAL IDEAL

According to the traditional view, what differentiates human beings from other animals has something to do with rationality. But we can distinguish between many varieties of rationality. The first and oldest distinction is between theoretical and practical rationality. “Theoretical rationality” most generally construed is any cognitive activity aimed at (propositional) knowledge, especially where this involves the mastery of general categories (or “intelligible forms,” in an older vocabulary). More narrowly, it can mean engaging in processes of reasoning, that is, making inferences. Famously, Aristotle celebrated theoretical rationality as the best and highest expression of human nature (it is what is most “godlike” in us), but he did not mean by this that we should make as many inferences as possible. For Aristotle, the pinnacle of theoretical rationality is not inference but “contemplation” (*theoria*): apprehending the truth, being fascinated by it.

“Practical rationality” has to do not with knowledge but with action. But it can have to do with action in multiple ways. *Procedural* practical rationality is means-ends reasoning, or more simply, “strategizing.” *Substantive* practical rationality is the intellectual apprehension of appropriate ends of acting (though it is debatable whether such apprehension should best be understood as an achievement of theoretical or of practical reason). A further sense has to do with assessing potential actions according to a normative standard, which is Kant’s notion.

Which of these notions of “rationality” is definitive of personhood? Presumably, not all are equally so. For example, I have a hard time taking seriously the idea that persons as such are masters of deductive logic. The trouble is not only that most of us humans are very bad at logic, but also that deductive logic is nearly always used, like other forms of inference, in the service of some further intellectual aim.¹⁵

There are more plausible proposals. In fact, we have already encountered two: Aristotelian *theoria* and Kantian moral autonomy. As accounts of the “highest” expression of human rationality, both proposals have won plenty of adherents. But they are clearly not compatible accounts, since *theoria* is a form of theoretical rationality whereas moral autonomy is a form of practical rationality.

But perhaps we don’t have to choose a side. *Theoria* and moral autonomy have something in common, something that cuts across the theoretical/practical divide. That something is *responding appropriately to value*. For Kant, the apex of rationality is willing in accord with normative requirements; to be rational just is to respect value and to order one’s actions accordingly. But Aristotelian rationality is equally value-responsive. After all, it is hard to

¹⁵Compare Zagzebski (2001, p. 405): “One problem with using rationality as the defining property of personhood is that some of what is involved in being rational seems to be irrelevant to being a person, for example, the ability to perform mathematical calculations. We can easily imagine a race of intelligent beings who are resourceful and sensitive investigators of their environment, but who never develop mathematics.”

understand why *theoria* would have any appeal if the object of contemplation was not something of value. *Theoria* connotes wonder and admiration, rather than clinical assessment or critical interrogation. Thus, *theoria* as much as moral autonomy is the rational creature's proper response to something of value.

I propose, then, that the notion of rationality operative in discussions of personhood ought to be, and has often been implicitly, the notion of *responsiveness to value*, or more specifically, the notion of responding comprehendingly, autonomously, and appropriately to value. Persons apprehend the value of valuable things, and then they make it their business to give those things their due. Humans, qua persons, are *caretakers of value*.

This is, of course, a normative ideal. We human persons manifestly do not apprehend everything of value, nor do we always respond appropriately when we *do* so apprehend. Thus, the notion of a person is not a “threshold” concept, but a “ceiling” concept: It is defined in terms of its ideal instances.

En route to understanding how we rationally sub-optimal humans might fall into the class of persons so understood, I want to begin with an account of what the ideal (finite) person would be like. There are, I want to suggest, three families of capacities that must be actualized in such a being.

First, such a being will be endowed with certain cognitive capacities. For she will need to know (1) what is valuable, and (2) how to respond.

(1) Knowing what is valuable will involve knowledge of all the “final values,” the bearers of ultimate, non-instrumental value. Moreover, it will involve knowledge of the relative prioritization of the final values, insofar as they are commensurable with each other.

(2) Knowing how to respond will involve several components. Different values call for different kinds of response (awe, respect, protection, promotion, imitation, union, consumption, and so forth)—some involving a kind of distancing from the valued thing, some involving a kind of closeness to it, and some a blend. So, the ideal person will need to understand how different final values count as reasons for these different sorts of responses. Furthermore, where the proper response involves goal-directed activity—for example, the consumption of food, the promotion of others' interests—she will need to know two more things: first, the variety of actions at her disposal; second, which of these actions will have the desired result vis-à-vis the final value(s) at stake. Knowing both these things requires understanding the causal structure of her world—ideally, systematic and exhaustive understanding of that structure.

The ideal person will need other psychological endowments beyond the cognitive. She will also need certain *affective* capacities. For one thing, sometimes the proper response to a good includes an affective dimension: awe and delight are examples of such responses. Even where the proper response is not essentially affective, the ideal person will desire to respond appropriately, and will be pleased so to respond. Or, to put it another way: Unless the good *attracts* a person, she has not fully apprehended its goodness.

And finally: the ideal person is not responsive to value in a mechanical way. For imagine a creature that is disposed to behave in contextually appropriate ways to various goods, but does so automatically. The mere deployment of dispositions isn't sufficiently active to count as *care-taking*, which implies bearing responsibility. So, the ideally rational being will also need certain *volitional* capacities—the capacity to freely and consistently choose actions that are appropriate with respect to the final values that she knows to be such and that move her affectively.

This, then, is our ideal of a “rational nature”: the orientation of certain cognitive, affective, and volitional capacities toward the good. Let us now draw our attention to the least developed analogues of these capacities—that is, to the most rudimentary animal mind. I will then provide

a sketch of how nature might “build up” from such a mind to a *human* mind, a mind with psychological endowments in the ballpark of the rational ideal just described.

3 | THE VERTEBRATE MIND

I begin with what I'll call the “Vertebrate Mind.” (Let me be clear that this is a toy model of a rudimentary mind. Some but not all vertebrates will exemplify it). The Vertebrate Mind orients an animal in the world, helping the animal to find food, avoid threats, and so on. Significantly, the Vertebrate Mind exemplifies a tri-part structure analogous to the one exemplified by the ideally rational mind. First, it has rudimentary “cognitive” capacities, or rather, “C-capacities” (a category wide enough to span the rudimentary and ideal versions, and everything in between). In particular, the Vertebrate Mind includes *sensation*, the ability to sense relevant stimuli in the environment, and *associative learning*, the ability to adjust behavioral dispositions in response to the pairing of stimuli and reward (or harm).¹⁶

Second, it has rudimentary affections, or “A-capacities”: positively and negatively valenced feelings that attract or propel it with respect to certain stimuli, for example, hunger, thirst, pain, and fear.¹⁷

Third, it has rudimentary mechanisms for behavioral control, or “V-capacities”: It produces behaviors that are appropriate, given its perceptual stimuli and affective goals.

4 | THE MAMMALIAN MIND

The Mammalian mind comprises the capacities of the Vertebrate Mind but with important augmentations (Again, this is a toy model, and it errs on the side of generosity—so probably closer to the mind of a chimpanzee than of a vole. Note, too, that certain avian and cephalopod minds are just as sophisticated¹⁸). The most striking additions fall under the headings of “executive functions” and “pro-social” behavioral dispositions, if we use the prevailing nomenclature in empirical psychology.

Let's begin with additions to Mammalian C-capacities. In the Mammalian Mind, *sensation* is transformed into *perception*. Perceptual episodes are not just *caused* by the world but are *about* the world; they *represent* the world. One familiar type of representation is symbolic representation, and many researchers, both in philosophy and in empirical science, take mental representation to be a form of symbolic representation. But it is not. Symbols bear an extrinsic relationship to what they represent, so they must be interpreted to be understood. But perceptual episodes do not need to be interpreted to be understood. They “make present” their representational objects in a way that symbols do not.¹⁹ Another way to describe the sensation/perception distinction is to say that perception involves the application of concepts to what is sensed. The standard suite of basic perceptual concepts is known as “core cognition,” and it includes the concepts *object*,

¹⁶Shettleworth (2013, p. 17).

¹⁷See Asma and Gabriel (2019, p. 8).

¹⁸See e.g. Balakhonov and Rose (2017) and Godfrey-Smith (2016).

¹⁹The locus classicus for this point is Bonjour (1998), ch. 6.

quantity, space, animacy, and agency.²⁰ Although these concepts are initially employed in perceptual episodes, their cognitive role need not be restricted to perceptual categorization. They can also be deployed “offline,” in isolation from perceptual episodes. The Mammalian Mind thus includes the basic conceptual resources necessary for thinking—that is, cognitive processing not directly triggered by perceptual inputs.

The Mammalian Mind also includes the basic *functional structure* necessary for thinking. Here, two of the three canonical “executive functions” are relevant: *working memory* and *cognitive flexibility*. The former is the ability to hold information in mind and manipulate it. The second is the ability to consider alternative possibilities, both in the sense of (a) alternative underlying explanations of the same appearance in different contexts, and (b) rival plans of action that need to be weighed. Concept-possession, working memory, and cognitive flexibility together add up to a rudimentary form of *intelligence* in mammals. It is not very impressive by human standards: Disjunctive syllogism and hypothetical syllogism may constitute its upper bound.²¹ But such rudimentary intelligence nevertheless affords the animal an expanded capacity for flexible behavior, based not on associative learning but on learning by explicit trial-and-error. (Of course, mammals learn associatively, too).

Regarding A-capacities: The Mammalian Mind adds a secondary emotional layer, while retaining the primary layer. These secondary emotions include grief, play, and empathy (or “care”).²² Importantly, these second-tier emotions are largely social. An animal’s affective state is responsive to conspecifics: It is disposed to feel well because a fellow animal is present (attachment), to feel badly because a fellow animal is absent (grief, or the animal analogue), and to feel well or badly because a fellow animal is displaying signs of itself feeling well or badly (empathy proper). This all amounts to a major modification of the animal’s motivational structure: Other animals *matter* to it. Of course, they might not matter very much, when weighed against other affective motivators. Mammal sociality is rather callous by human standards.

The Mammalian Mind has updated V-capacities, too. In particular, it includes a rudimentary version of the third canonical executive function, “inhibitory control.” In humans, the term is used to cover a range of volitional activities. When applied to animals, it usually just means *delayed gratification*. Most animals are very bad at delaying gratification, but a few can do so for a little while if the payoff is big enough.²³ More generally, mammals need volitional control (not just “inhibition”) in order to bring their intelligence to bear on their behavior. If there is always a direct functional link between affect and behavior, the animal’s thinking won’t really matter at all; cognitive flexibility would not translate into behavioral flexibility. At the very least, volitional control interrupts this direct link, allowing a thought to dictate the animal’s behavior in place of an urge. Additionally, volitional control is involved in cognition itself, as an animal sustains attention on a particular task.

There are two ways to understand what is “executive” about executive functions. One way is to treat such functions as architectonic but nevertheless mechanical within a mechanical system.

²⁰The framework of ‘core cognition’ (Carey, 2009) or ‘core knowledge’ (Spelke, 2000) was first developed in the context of human developmental psychology, but was subsequently used to study other mammals. See Shettleworth (2013, p. 120ff). Different theorists provide slightly different lists.

²¹Laland and Seed (2021).

²²Asma and Gabriel (2019, p. 9). I prefer the term “empathy,” as “care” is apt to be understood anthropomorphically. “Empathy” can similarly mislead. The point is that we are still talking about a type of *feeling* that motivates.

²³“For a reward forty times larger than the immediate reward option, chimpanzees may wait up to eight minutes.” (Suddendorf, 2013, p. 109).

Another way is to treat them as spontaneous interventions *into* the mechanical systems. Volitional control, I maintain, must be understood in this latter way. Martine Nida-Rümelin makes the case in a particularly vivid and concise way:

If you observe a squirrel jumping from one branch of a tree to another, then the squirrel does not look to you like a mechanism that jumps as the result of some inner “mechanical” process. It looks to you as though the squirrel itself, the subject of experience, does the jumping. ... We see the movements of biological organisms that we implicitly accept to be conscious as being done by the conscious individual itself. A related claim is true for the way we experience our own doings. We experience our doings as brought about by ourselves. To assume that some inner processes cause our doings is incompatible with the content of the phenomenology of our experience. If these experiences of ourselves when we are active and our perceptions of others as being active are not illusory, then conscious individuals are active in their doings.²⁴

In short, if ever a minded creature is *active*—if it does more than ride the causal waves of the universe—then its behaviors are sometimes its very own, irreducible, spontaneous *doings*.

5 | THE HUMAN MIND

I will describe the transition from the Mammalian to the Human Mind in four “stages”:

1. *The accumulative stage*, in which Mammalian capacities are quantitatively increased.
2. *The ampliative stage*, in which these accumulated capacities interact with each other in complex ways.
3. *The additive stage*, in which a qualitatively new capacity is superadded.
4. *The transformative stage*, in which this new capacity transforms the nature of the capacities shared with the Mammalian heritage.

I invoke these “stages” heuristically, but I expect that they have corollaries in both phylogenetic and ontogenetic development. The first two stages are consistent with the Mammalian-Human difference being only one of degree. Whatever is controversial here will be a matter of detail, I suspect. But the third and fourth stages take us into highly contested territory.

5.1 | The accumulative stage

The uniqueness of the Human Mind is built on the foundation of increases of C-capacities on two fronts: executive functioning and social cognition.²⁵

Human executive function is more robust than mammalian executive function in all of its dimensions: Humans can think longer and harder and about more things than can mammals. One particularly important advance is the ability to engage in what cognitive scientists call “mental time travel”: recalling past episodes (“episodic memory,” which is rare in the animal kingdom,

²⁴Nida-Rümelin (2007, p. 208).

²⁵See Suddendorf (2013) and Tomasello (2019).

maybe entirely absent²⁶) and imagining future episodes. The human imagination, in other words, is leaps and bounds better than the mammalian imagination, and this development—along with other quantitative improvements in executive function—opens up new possibilities for procedural rationality. Whereas mammals can solve the occasional problem, humans deliberate and create.

The basic social-cognitive skill that humans have but mammals lack is, according to Michael Tomasello (2019), joint attention. A prototypical example of joint attention occurs in the context of a child playing with an adult. Each participant is aware of the object of their individual attention—some toy, say—but also of the other's awareness of that object. Joint attention is crucial because it forms the basis of human communication. Every successful linguistic exchange is an invitation extended and received to attend together to some information or other. What makes joint attention possible is not a settled matter, but at least two ingredients seem essential. One is the mundane fact that humans have whites around their irises which enable gaze-following. But another (which exploits the first cognitive advancement) is a much-improved capacity for “mindreading” (or “theory of mind”). While some mammals can track what another creature intends to do or knows, humans track enormous amounts of such information, about many specific individuals, keeping it updated in real time.²⁷

Plausibly, there is an additional accumulation of A-capacities that is relevant here, viz., a greater motivational weighting of the social emotions. For example, while chimpanzees extend help to others, human children are much more eager to help, from an earlier age.²⁸ Alternatively, it is possible that accumulations of C-capacities are sufficient, without alternations to A-capacities, to translate existing social emotions into such helpfulness.

The basic Mammalian-Human differences with respect to executive function and social cognition are all quantitative, and moreover, almost if not entirely restricted to C-capacities. But special things start to happen when these, at this advanced human quantitative threshold, begin to interact with each other and with the rest of the Mammalian inheritance.

5.2 | The ampliative stage

Language, learned by exploiting enriched C-capacities and in the context of joint attention, is the catalyst for a plethora of mutually enhancing feedback loops among the various capacities in the Human Mind.²⁹

Two features of language make its possession into something of a cognitive super-charger: It is *symbolic*, and it is *communal*. Recall that core cognition—the conceptual repository available to Mammals—is not a matter of symbols in animals' heads. The concepts that make up core cognition are “grasped” rather than symbolized. But what can be symbolized does not have to be grasped. A symbol, once given a conventional meaning by a linguistic community, can be used to

²⁶Laland and Seed (2021).

²⁷Dunbar (2021).

²⁸Tomasello (2014).

²⁹There are essentially two camps regarding the prerequisites for language learning. One camp, associated with Noam Chomsky, maintains that there is a uniquely human language-learning module. Another holds that executive function and social cognition are together sufficient. In other words, language learning is a form of social pattern-detection. See Tomasello (2019, p. 127ff). I will assume the latter, though of course it will do no violence to my aims if the former turns out to be correct.

express that meaning by any member of the community—whether or not that language-user is even capable of grasping of the meaning.

Why is this important? Language is the tool whereby the Mammalian inheritance of core cognition (e.g., the concepts *object*, *quantity*, *space*, *animacy*, and *agency*) is leveraged into the full suite of concepts that humans use to make sense of the world. Susan Carey (2009) gives the canonical example of this process. Core cognition includes a system for tracking only very small quantities (up to 3 or 4 items), and yet humans learn to count. How? First, it turns out, by rote recitation of the number series, then by mapping magnitudes from core cognition onto the number series, and then—*eventually*—grasping the successor rule (successive numbers are the same as the previous number + 1). More generally, words serve as placeholders for concepts that are acquired via such processes of “structure mapping.”³⁰ And this leveraging of language does not just occasion the acquisition of discrete concepts. Syntactic structures serve as placeholders for various semantic relationships that contents can bear to each other. Language as a system of syntactic rules serves as an entree into a system of semantic relationships: the space of reasons.³¹

In sum: language “outsources” meaning; it puts at one’s disposal all of the representational resources of one’s community, prior to, or even in the total absence of, grasping meanings for oneself. And the result is that a human can think about not just those categories that are built into the architecture of her mind but about anything at all.

A human’s greatly expanded C-capacities have a recursive, ampliative effect on A-capacities. In particular, humans develop a tertiary tier of emotions (without losing the primary and secondary tiers) that develop in response to a world richly conceptualized: feeling anxiety about an upcoming exam, feeling hopeful about reconciling with a friend, etc.³² Importantly, not only do these emotions have cognitive *sources*, they have cognitive *contents* as well. As Patricia Greenspan puts it: “Emotions that represent their objects in some positive or negative light (as most do) may be said to have a content expressible by an evaluative proposition.”³³ Not only do such emotions *feel* good or bad, but such valences represent something else *as* good or bad. In other words, in the Human Mind, emotions begin to represent *value*.

How do they come to do so? Answering that question turns on the question of how humans acquire their normative concepts, and this is a disputed matter, but there is widespread agreement that social interaction plays an important role. Here is one account, from Carpendale and Lewis (2020):

[Moral obligation] is already implicit in the human developmental system as a result of the nature of early relationships. Infants are treated as persons, as participants in interaction. It is the product of treating others as persons and responding to them in everyday activity... Caring and mutual affection are embedded in the structure of the human developmental system. These strong emotional bonds are the seed for mutual respect, which is already there in communication, and develops increasingly into moral obligation.³⁴

³⁰See Gentner (2010).

³¹See Tomasello (2019, pp. 120–121), Asoulin (2019).

³²Asma and Gabriel (2019, p. 9).

³³Greenspan (2004, p. 204).

³⁴Carpendale and Lewis (2020, p. 43).

The suggestion here is that interpersonal interaction transforms a certain kind of *feeling*—delighting in the presence of another—into a certain kind of *knowing*—that the other is valuable. The very feeling itself becomes a means of apprehending normative reality. In support of this idea, Dahl and Killen (2018) report that the helping behaviors of toddlers have much less to do with what others need and more to do with what will engage others interactively, but that this pattern is replaced by more need-centered helping in the next few years of development. That is, the development of human moral psychology starts with taking delight in interacting with others, proceeds to an awareness of others' concerns, and culminates in the grasping of norms of rights and fairness—from attachment to partiality to impartiality.³⁵

Now, emotions are not the only way that humans represent value. We have coined the word “value,” for example. And this means that our capacity to represent value can come untethered from particular emotional experiences and be appropriated into one's cognitive economy. One can form beliefs about what matters, in other words. One can, furthermore, form such beliefs on the basis of explicit moral instruction. Subsequently, tertiary emotions can be generated that share representational contents with these moral beliefs. Thus, human emotions are to the world of value what sensory qualities are to the world of perceptibles. We know there is a world out there even when we are not sensing it, but the world is made manifest to us only when our consciousness is saturated with visual hues and sonic pitches and the like. Likewise, we know that we inhabit a world of goods and bads and rights and wrongs, and we reason about these matters, but their reality is made manifest to us only when our consciousness is saturated with longings and revulsions and enchantings and so on.

The effects of the foregoing on V-capacities are dramatic. Recall that inputs to mammalian V-capacities are (1) affective goading (what the animal wants or likes, including the company of other animals) and (2) procedural planning, of a rudimentary sort. With the explosion of language-induced cognition, the sheer quantity of such inputs, in both categories, is massively expanded. But there will also be a new *type* of input: awareness of the good, as such. In addition, an awareness of *time* will greatly expand the time-scale relevant to the exercise of control. Mammals navigate situations; humans pursue projects. Thus, the outputs of human V-capacities are not *behaviors*, but rather *intentions*, including temporally extended varieties such as “standing” intentions or “distal” intentions.³⁶

This is an awful lot for a Human Mind to manage. Perhaps it could be managed with an exponentially larger working memory capacity, but our mammalian heritage has not bequeathed such resources to us. Rather, we manage our volitional life by pushing as many decision-making processes as possible out of consciousness entirely, automating them. One name for this is “secondary modularization”—the mind's construction of automatic procedures dedicated to various cognitive tasks, analogous to the “primary modularization” that underwrites core cognition and other mammalian cognitive tasks.³⁷ Other names are “habit” and “expertise.”

A final source of dramatic ampliative change comes in the form of social interaction. We have already mentioned two ways that social interaction shapes the Human Mind: It provides one

³⁵See Schaubroeck (2019) on the way that partial love might afford awareness of the grounds of moral duties, or perhaps even constitute those grounds. Schaubroeck bucks the dominant trend in moral philosophy of pitting partial love against impartial morality—an opposition that is implausible anyway, in light of the evidence from moral developmental psychology.

³⁶For a discussion of these categories, see Mele and William (2009).

³⁷On “secondary modularization,” see Burkart et al. (2017).

with language, a cognitive super-charger; and it supplies the context in which *value* is first apprehended and conceptualized. Of course, the formative influence of one's social world really kicks in once these initial contributions have already been made. Humans *teach* one another, broadly, extensively, explicitly, and universally, whereas mammalian teaching is extremely rare.³⁸ One of the things that humans teach one another is *how to reason*. Left to themselves, individual humans are not great at thinking things through. While we are capable of grasping semantic connections among concepts (as are some mammals), we do not exploit those connections in deliberation particularly well. We make lots of mistakes, we fail to consider alternatives, and so forth. It is when we start giving reasons *to each other* that we get a better sense for what is relevant and what is not.³⁹

5.3 | Additive stage

The Human Mind, as we have depicted it so far, is already very different from the Mammalian Mind. It thinks differently: about anything at all; it feels differently: evaluatively, in ways that are shaped by its beliefs; and it wills differently: it concerns itself with the good and with the future. It is debatable, I think, whether the introduction of normative elements into A-capacities and V-capacities amounts to a *kind* difference rather than a mere *degree* difference; probably sorting this out would require a more precise account of the transition than I have given. But it doesn't matter, because what I have described so far is not yet the mind of a person. It would be something like the mind of a cognitively sophisticated moral robot. For a person is more than an instance of a nature: A person is a *self*. And selfhood requires the addition of a crucial element absent from our discussion so far: *self-consciousness*.

To motivate this claim, I begin with two evocative statements in support of a Mammalian-Human kind difference. First, Robert Spaemann:

Human beings...exist by distinguishing their being from their specific way of being, their specific "nature." Their nature is not what they *are*, pure and simple; their nature is something that they *have*. And this "having" is their being.⁴⁰

The second is from Christine Korsgaard:

[We] exert a deeper level of control over own [sic] movements when we choose our ends as well as the means to them than that exhibited by an animal that pursues ends that are given to her by her affective states, even if she pursues them consciously and intelligently. Another way to put the point is to say that we do not merely *have* intentions, good or bad. We assess and adopt them. We have the capacity for

³⁸Laland and Seed (2021). One of the rare cases of "teaching" identified in the animal kingdom (as mentioned in the introductory section above) occurs among meerkats. It amounts to mature animals' providing the young with opportunities to develop skills in hunting—for example, letting a pup complete the killing of a scorpion whose stinger has been removed by an adult. "It is often assumed," write the investigators, that "teaching requires awareness of the ignorance of pupils and a deliberate attempt to correct that ignorance, but viewed from a functional perspective, teaching can be based on simple mechanisms without the need for intentionality" (Thornton & McAuliffe, 2006, p. 229). Clearly, this is not an instance of teaching in the full human sense.

³⁹Mercier and Sperber (2017).

⁴⁰Spaemann (2006, p. 31).

normative self-government, or as Kant called it, “autonomy.” ... The distinctive character of human action gives us a whole different way of being in the world.⁴¹

The person, the rational human individual as such, has a “whole different way of being in the world.” What is this different way of being? Spaemann characterizes it as a kind of distance between the human and her nature. The crucial difference between a non-rational and a rational one is not so much the capacities that characterize each nature but the relationship that the rational animal has to those capacities—the relationship of “having,” rather than of *being*, of definition.

Likewise, Korsgaard claims that the human ability for “normative self-government” is explained, not by a unique type of motivation, but by a unique relationship that she has to her motivations. They are not “hers” automatically; she *adopts* them, or not.⁴² And she can do this because of the unique way she is aware of them.

Korsgaard calls that unique relationship “self-consciousness.” It is now standard to appeal to self-consciousness as a distinguishing feature of personhood, an emphasis that goes back to John Locke.⁴³ But what can get lost in discussions of this distinguishing feature—which Korsgaard preserves—is the dramatic novelty of this psychological capacity. What Korsgaard has in mind is apt to be conflated with two lesser psychological phenomena: (1) meta-cognition, in the sense of having mental states that are about other mental states; and (2) what Elisabeth Schechter (2018) calls “implicit self-awareness,” or the ability to have self-concerning attitudes. (The famous “mirror-test” of self-consciousness is really a test of implicit self-awareness, argues Schechter). Nor, I think, should self-consciousness be identified, as it sometimes is, with (3) the possession of an I-concept.⁴⁴ Presumably, no one can deploy, or perhaps even possess, an I-concept unless one is self-conscious, but self-consciousness is the more fundamental psychological phenomenon.

What, then, *is* self-consciousness? Scientists of consciousness study not only the contents of consciousness but the “level” of consciousness. The paradigm here is the distinction between dreaming and wakefulness. I want to suggest that self-consciousness is as different from merely being conscious (even in sophisticated ways) as wakefulness is from dreaming. It is a new *level* of consciousness. (And it is one we drop in and out of, just as we drop in and out of being awake). It is the dawning of an inner light, a “coming home” to oneself as a self. Unless it is present, the subject is not self-governed, but rather nature-governed—even as she executes the deliverances of her capacities in spontaneous action. And if it is present, “it creates a new kind of psychological being,” as Schechter puts it.⁴⁵

⁴¹Korsgaard (2006, pp. 112–113, 117).

⁴²In places, Korsgaard does suggest that what marks the difference is a unique type of motivation, viz., a moral one, for example: “A form of life governed by principles and values is a very different thing from a form of life governed by instinct, desire, and emotion” (117). No doubt this is right, but it cannot mark the non-rational/rational distinction, since the *self* needs to be doing the governing, not the “principles and values.” There are, in other words, two ways a being could fail to exhibit “normative self-government”: if its motivations fail to be normative, and if the self fails to be what governs.

⁴³A person, according to Locke, is “a thinking intelligent being that has reason and reflection, and can consider itself as itself, the same thinking thing, in different times and places; which it does only by that consciousness which is inseparable from thinking, and as it seems to me, essential to it: it being impossible for any one to perceive without *perceiving* that he does perceive.”

⁴⁴Such is Lynne Rudder Baker’s (2013) account of self-consciousness, or what she calls “the robust first-person perspective,” which she claims is unique to persons.

⁴⁵Schechter (2018, p. 192).

From a neuroscientific perspective, levels of consciousness are puzzling. The character of one's phenomenal experience differs dramatically depending on one's level of consciousness, yet underlying brain function is comparatively stable across changes to these levels. REM sleep, for example, is sometimes known as “paradoxical sleep,” because the brain is as active during REM sleep as during wakefulness.⁴⁶ Evidently, smallish brain changes can result in dramatic changes to the quality of consciousness. Thus, I propose that at some point in human evolution, a small genetic change resulted in an abrupt expansion of the possibilities of human consciousness—and self-conscious creatures came into existence for the first time.

5.4 | Transformative stage

In “Additive Theories of Rationality: A Critique,” Matthew Boyle argues that the kind difference between rational and non-rational animals cannot be accounted for simply in terms of the addition of a new capacity, because the rational part of the resulting mind would be at odds with the non-rational part. “What entitles us to hold that this reflects a fracture within a single subjective standpoint,” he asks, “rather than a struggle between two essentially distinct standpoints for control of a certain body?”⁴⁷ Boyle mentions Korsgaard as someone whose theory of the rational/non-rational difference could lead to this problem.

I am not treating self-consciousness as a separate psychological *faculty* added on to the rest, so my proposal is already in a Boylean spirit. Moreover, while I do not think that self-consciousness is merely a result of ampliative interactions among human faculties, it is consistent with my picture that self-consciousness was already required as a catalyst of some of the ampliative processes I have described. For example, it is plausible that human social cognition requires a kind of self-other differentiation that is only possible for a self-conscious being.⁴⁸ Thus, any capacities built on the foundation of social cognition—including the transition from affectively liking someone to valuing her/him—require self-consciousness.⁴⁹

Nevertheless, we can imagine that one who lacked self-consciousness and suddenly had “the lights turn on,” so to speak, could very well feel alienated from her own cognitive, affective, and volitional states. (This is the familiar reaction typified by characters in science fiction who are freed from various kinds of mind-control). This type of alienation is impossible for a being who lacks self-consciousness. There is, I suggest, a diachronic process of unification that occurs in a self-conscious agent. To use Korsgaard's language, the self-aware agent *assesses* and *adopts* certain of her inclinations, rejecting others. She builds habits. She makes commitments. She constructs a practical identity. Or tries to; the process can fail. If it succeeds, the psychological unity that results is, at least in part, of her own making.⁵⁰

⁴⁶See McNamara (2019).

⁴⁷Boyle (2016, p. 549).

⁴⁸As Schechter (2018) argues convincingly.

⁴⁹Boyle himself implies that the acquisition of value concepts is only possible in an already-rational mind: “While it may be correct to say that a non-rational animal's desires present their objects as attractive (for instance, as promising pleasure or promising to relieve some distress), it cannot be correct to say more specifically that they present their objects as desirable (i.e., as meriting desire): this way for something to be attractive lies beyond the scope of a nonrational mind.” Ibid., p. 539.

⁵⁰Compare Schechter (ibid., p. 186): “The executive aspect of self-consciousness provides the basis for the self-constructing agent, who is more autonomous than non-self-conscious agents in being able intentionally to guide and direct her own behavior by conforming it only to principles of action that she herself accepts.”

6 | CONCLUSION: BECOMING CARETAKERS OF VALUE

According to the traditional view, we humans are *persons*: rational selves, and thus bearers of dignity. Non-human animals are not persons; they lack rationality proper, as well as that self-consciousness which enables self-hood. In our endeavor to make sense of this traditional view, we started from the notion of “rational nature” understood as an ideal: a *caretaker of value*, one whose cognitive, affective, and volitional capacities are oriented toward the good. We then described rudimentary Vertebrate Minds, more sophisticated Mammalian Minds, and finally Human Minds, which accumulate, amplify, add to, and finally transform the Mammalian psychological inheritance—until a different kind of mind emerges.

Have we shown how human beings could be persons, that is, realizers of the rational ideal? In one sense, no. Because the notion of a person—as a caretaker of value—is an ideal, having a Human Mind does not guarantee that one attains that ideal. One is not born, but rather becomes, a person—and not simply by standard maturational stages but by free acts of normative self-governance.

But in another sense, yes: The Human Mind is rational in the sense of being *capable* of such normative self-governance. The ideally rational mind, recall, will (1) know what is valuable and according to what priority ranking; (2) know how to respond appropriately, and the means necessary to do so effectively; (3) desire to so respond and delight in doing so; (4) freely and consistent choose to so respond. The Human Mind really is capable of all of this. But (we might say) *just barely*—in peculiar and limited ways. Or, from another angle, in stunningly economical ways—repurposing every available resource from its Mammalian heritage.

The Human Mind realizes (1) by repurposing Mammalian emotions (especially social ones) as representations of value.

It realizes (2) by repurposing Mammalian intelligence—in particular, (a) repurposing Mammalian problem-solving capabilities as general-purpose reasoning abilities, (b) expanding and calibrating these abilities by training automated cognitive modules (mental “habits”) while exchanging reasons socially; and (c) repurposing Mammalian core cognition in tandem with outsourced linguistic meanings, thereby expanding its representational repertoire indefinitely.

It realizes (3) by joining its affective system with the jerry-rigged cognitive system just described, creating a new “tertiary” tier of emotional responses.

It realizes (4) by repurposing Mammalian spontaneous volitional control as bona fide free agency—empowered by self-consciousness, and expressed in the training of automated behavioral modules (that is, habits).

In short, though we humans are not, *eo ipso*, persons, we have the potential to become persons: we have the requisite psychological endowments to make the journey. Our capacity to make that journey means that we are different kinds of things from our non-human terrestrial companions, despite our sharing so many of their psychological endowments.

I have been defending the metaphysical dimension of the traditional view—viz., the claim that humans, qua persons, are creatures of a different sort from all non-human animals; we stand apart. It is worth turning now to the moral dimension of the traditional view. The two dimensions are clearly related: It would be odd if human beings’ peculiar moral status had nothing to do with their peculiar nature. Nevertheless, there is more than one way to connect the dots between our nature and our moral status. Readers who have been convinced by what I have said so far are free to reject what I proceed to say in the remainder of the paper.

Humans' distinguishing feature, I have claimed, is their unique capacity to steward the good. What is so great about stewardship of the good? What makes persons, so understood, more valuable than non-persons?

Before I answer that question, I need to address a standard objection against any account that grounds human dignity in psychological capacities, as mine does. The objection comes in the form of a dilemma: Either human dignity is grounded in *sophisticated* psychological capacities, in which case a wide range of human beings are excluded, for example, infants, the cognitively impaired, and victims of dementia; or else it is grounded in *primitive* psychological capacities (sentience, say), in which case a wide range of non-humans are included. Thus, no appeal to psychological capacities can make good on the traditional elevation of all humans over all animals.⁵¹

A standard reply, in defense of appealing to sophisticated capacities, is to say that the ground of dignity is not the *possession* of the relevant psychological capacities but the *potential* to possess them, a potential which all humans have. But then the objector will observe that potentialities are all too easy to come by. If a severely cognitively impaired human has the potential to be as cognitively sophisticated as a typical human adult, then so does a chimpanzee—with sufficient technological fiddling.

I respond by rejecting both horns of the dilemma. It is neither possession nor potential that matters, but *propriety*. It is proper to a human, for example, to be able to make moral decisions, but it is not proper for a chimpanzee to do so. It is no sign of ill health or immaturity if a chimpanzee lacks that ability, whereas it is a sign of ill health or immaturity if a human lacks it—even if there is some sense in which a chimpanzee has the potential for that ability (after undergoing radical enhancement of some sort). There is thus a conceptual connection between the kind of thing a creature is and the mode of flourishing that is proper to it, whereas no such conceptual connection exists between a creature's kind and its potentials.

But what explains this difference in proper flourishing between humans and non-human animals? Why must any human become a person in order to flourish? There is no space for a full answer, but I can at least sketch the shape that such an answer could take. Human psychological development “comes to rest” at a different place than it does for a chimpanzee. A mature human's psychology is organically unified and harmoniously expressed only when, as I have proposed, her faculties, under her own self-conscious direction, can register and respond appropriately to value. Anything short of this will amount to either (a) an incomplete or frustrated process of psychological development, or (b) inter-psyche conflict of some sort—a sense, however vaguely felt, that one is at odds with oneself or is selling oneself short. (And this is true despite how difficult and rare it is for anyone to arrive at that place where the human psyche comes to rest. Hence the great moral restlessness of human beings). But not so for a non-human animal, for whom stable psychological unity and harmony do not require self-conscious stewardship of the good.

Returning now to the main question: what is it about persons, then, that elevates their moral status? Why must we respect persons in ways and to degrees that are not applicable to non-persons?⁵²

Part of the answer can be read directly off of our account of personhood. What is “dignified” about human beings is precisely their capacity to steward the good. Sarah Buss puts it this way:

⁵¹See Jaworska and Tannenbaum (2019) for one version of this familiar objection.

⁵²What follows summarizes a more extended discussion in my (under review).

Human dignity is grounded, at least in part, in the human capacity to transcend the concerns, needs, and demands of the self in paying disinterested tribute to what warrants this response. ... Precisely because human beings are capable of performing the sublime service of giving things of value their due, we have reason to acknowledge *their* value by treating them as constraints (though not necessarily *absolute* constraints) on what we can justifiably do.⁵³

What is of value deserves to be stewarded. It is thus a very good thing that reality includes such stewards. That which can pay proper respect to the good is, for that reason, something to which we ought to pay respect.

This can't be the whole story, however. To see why, consider the phenomenology of moral concern. When I am moved to extend help to a destitute person, what moves me? Not, I submit, the fact that she is able to apprehend and respond appropriately to the good. I might regret that her death would remove a caretaker of value from the world. But my concern is for *her*, not just for the role she plays, however noble of a role it may be.

Another problem is that the Buss-style story about dignity leaves out the *moral non-fungibility* of persons. It is not merely to *humanity* that I owe respect. Rather, I owe respect to each unique human person as such. But as far as the value of stewarding the good is considered, any old persons will do; unique selves as such do not matter. Thus, as Linda Zagzebski (2001) has argued persuasively, the rational *nature* of human persons (which I have identified with the capacity to steward the good) cannot be the full story about their dignity. A person is worth more than what her *nature* is worth.

The full story about human dignity includes the point that stewarding the good is dignifying, but it includes more besides. Here's the rest of the story (or the bulk of it, anyway).

Being a caretaker of value confers upon one a dreadful vulnerability. All creatures with interests presumably are owed some modicum of regard. But because our interests include apprehending value, and because we are self-conscious and thus reflective about those interests, we are capable of a different kind of flourishing (when we apprehend the good) and a different kind of suffering (when we are deprived of the good) than other creatures. It is thus not a person's *rational nature* that is the locus of moral concern but the person as rational *self*—as the vulnerable, self-conscious subject at the center of the storm, so to speak—who looks out at the world knowingly, lovingly, hopefully, and fearfully.

ACKNOWLEDGEMENTS

For helpful comments on earlier versions of this paper, I would like to thank Taylor Cyr, Hao Hong, Tim O'Connor, Tim Perrine, the audience at the 'Theological Anthropology in a Scientific Age' conference at Beeson Divinity School in January 2022 (and especially Josh Reeves for his invitation to keynote at that conference), and two anonymous referees for this journal. This project was made possible through the support of the Henry Center for Theological Understanding, funded by a grant from the John Templeton Foundation. The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the John Templeton Foundation.

CONFLICT OF INTEREST STATEMENT

The author has no conflicts of interest to report.

⁵³Buss (2012, p. 353 and 357).

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

ORCID

Philip Woodward  <https://orcid.org/0009-0003-0356-2789>

REFERENCES

- Asma, S. T., & Gabriel, R. (2019). *The emotional mind: Affective roots of culture and cognition*. Harvard University Press.
- Asoulin, E. (2019). Phrase structure grammars as indicative of uniquely human thoughts. *Language Sciences*, *74*, 98–109.
- Balakhonov, D., & Rose, J. (2017). Crows rival monkeys in cognitive capacity. *Scientific Reports*, *7*(1), 8809.
- Bluff, L. A., Troscianko, J., Weir, A. A., Kacelnik, A., & Rutz, C. (2010). Tool use by wild new Caledonian crows *Corvus moneduloides* at natural foraging sites. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1686), 1377–1385.
- BonJour, L. (1998). *In defense of pure reason: A rationalist account of a priori justification*. Cambridge University Press.
- Boyle, M. (2016). Additive theories of rationality: A critique. *European Journal of Philosophy*, *24*(3), 527–555.
- Burkart, J. M., Schubiger, M. N., & van Schaik, C. P. (2017). The evolution of general intelligence. *Behavioral and Brain Sciences*, *195*, 1–24.
- Buss, S. (2012). The value of humanity. *The Journal of Philosophy*, *109*(5/6), 341–377.
- Carey, S. (2009). *The origin of concepts*. Oxford University Press.
- Carpendale, J. I., & Lewis, C. (2020). Tomasello's tin man of moral obligation needs a heart. *Behavioral and Brain Sciences*, *43*, 19–20.
- Cheney, D. L., & Seyfarth, R. M. (1980). Vocal recognition in free-ranging vervet monkeys. *Animal Behaviour*, *28*(2), 362–367.
- Crystal, J. D. (2010). Episodic-like memory in animals. *Behavioural Brain Research*, *215*(2), 235–243.
- Dahl, A., & Killen, M. (2018). A developmental perspective on the origins of morality in infancy and early childhood. *Frontiers in Psychology*, *9*, 1736.
- De Waal, F. (2006). *Primates and philosophers: How morality evolved*. Princeton University Press.
- Doris, J., Stich, S., & Phillips, J. (2020). Moral psychology: Empirical approaches. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2020 Edition). <https://plato.stanford.edu/archives/spr2020/entries/moral-psych-emp/>
- Dunbar, R. (2021). *Friends: Understanding the power of our most important relationships*. Little Brown.
- Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science*, *34*(5), 752–775.
- Godfrey-Smith, P. (2016). *Other minds: The octopus, the sea, and the deep origins of consciousness*. Farrar, Straus and Giroux.
- Greenspan, P. (2004). Practical reasoning and emotion. In A. R. Mele & P. Rawling (Eds.), *The Oxford handbook of rationality* (pp. 206–221). Oxford University Press.
- Herman, L. M., Richards, D. G., & Wolz, J. P. (1984). Comprehension of sentences by bottlenosed dolphins. *Cognition*, *16*(2), 129–219.
- Hoppitt, W. J., Brown, G. R., Kendal, R., Rendell, L., Thornton, A., Webster, M. M., & Laland, K. N. (2008). Lessons from animal teaching. *Trends in Ecology & Evolution*, *23*(9), 486–493.
- Jaworska, A., & Tannenbaum, J. (2019). Personhood and moral status. In A. LoLordo (Ed.), *Persons: A history* (pp. 334–362). Oxford University Press.
- Korsgaard, C. M. (2006). Morality and the distinctiveness of human action. In S. Macedo & J. Ober (Eds.), *Primates and philosophers* (pp. 98–119). Princeton University Press.
- Laland, K., & Seed, A. (2021). Understanding human cognitive uniqueness. *Annual Review of Psychology*, *72*, 689–716.
- LoLordo, A. (Ed.). (2019). *Persons: A history*. Oxford University Press.
- McNamara, P. (2019). *The neuroscience of sleep and dreams*. Cambridge University Press.

- Mele, A. R., & William, H. (2009). *Effective intentions: The power of conscious will*. Oxford University Press.
- Mercier, H., & Sperber, D. (2017). *The enigma of reason*. Harvard University Press.
- Nida-Rümelin, N. (2007). Dualist Emergentism. In B. P. McLaughlin & J. D. Cohen (Eds.), *Contemporary debates in philosophy of mind*. Blackwell.
- Schaubroeck, K. (2019). Reasons of love. In A. Martin (Ed.), *The Routledge handbook of love in philosophy* (pp. 288–299). Routledge.
- Schechter, E. (2018). *Self-consciousness and 'split' brains: The minds' I*. Oxford University Press.
- Shettleworth, S. J. (2013). *Fundamentals of comparative cognition*. Oxford University Press.
- Spaemann, R. (2006). *Persons: The difference between 'someone' and 'something'*. Oxford University Press.
- Spelke, E. S. (2000). Core knowledge. *American Psychologist*, 55(11), 1233–1243.
- Suddendorf, T. (2013). *The gap: The science of what separates us from other animals*. Constellation.
- Thornton, A., & McAuliffe, K. (2006). Teaching in wild meerkats. *Science*, 313(5784), 227–229.
- Tomasello, M. (2014). The ultra-social animal. *European Journal of Social Psychology*, 44(3), 187–194.
- Tomasello, M. (2019). *Becoming human*. Harvard University Press.
- Wasserman, E. A., & Young, M. E. (2010). Same–different discrimination: The keel and backbone of thought and reasoning. *Journal of Experimental Psychology: Animal Behavior Processes*, 36(1), 3–22.
- Woodward, P. (Under review). *Homo adorans and the grounds of human dignity*.
- Zagzebski, L. (2001). The uniqueness of persons. *Journal of Religious Ethics*, 29(3), 401–423.