It's Only Human

It’s Only Human

The Evolution of Distinctively Human Cognition

ARMIN W. SCHULZ

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide. Oxford is a registered trade mark of Oxford University Press in the UK and in certain other countries.

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America.

CIP data is on file at the Library of Congress

ISBN 9780197800157

DOI: 10.1093/9780197800188.001.0001

Printed by Integrated Books International, United States of America

For Kelly, James, and Elizabeth:
distinctive humans of the best kind

Contents

Preface

This book pulls together many different strands of my recent research, and, as such, there are many different people who were integral in shaping it. Particularly notable among these are Colin Allen, Sarah Brosnan, Cecilia Heyes, Edouard Machery, Ron Mallon, and Bill Wimsatt. In fact, in line with the core argument laid out in what follows, any insights contained in this book should be seen to be built on the foundations laid by countless others ("it's turtles all the way down"). While these other scholars may not all be mentioned here, I am still very, very grateful to all of them. This also includes three anonymous referees for Oxford University Press who provided extremely helpful comments on a previous draft of this book and audiences at Northwestern University and the University of Pittsburgh who made many thought-provoking comments that greatly improved several key ideas expressed here. Many thanks to all of them.

Chapter 4 greatly expands and revises my "Enhancing Thoughts: Culture, Technology, and the Evolution of Human Cognitive Uniqueness" (*Mind & Language*, 2022, 37: 465–484, https://doi.org/10.1111/mila.12320). Chapter 5 is a revised version of my "The Origins of Distinctively Human Mindreading: A Bio-Social-Technological Coevolutionary Account" (*British Journal for the Philosophy of Science*, forthcoming). Chapter 7 is a revised version of my "Tools of the Trade: The Bio-Cultural Evolution of the Human Propensity to Trade" (first published in *Biology and Philosophy*, 2022, 37: 8, https://doi.org/10.1007/s10539-022-09837-2 by Springer Nature). Parts of Chapters 2 and 8 draw on my "Human Curiosity Then and Now: The Anthropology, Archaeology, and Psychology of Patent Protections" (first published in A. Killin & S. Allen-Hermanson (Eds.) *Explorations in Archaeology and Philosophy*, 2021, Dordrecht: Springer, pp. 67–84).

1
Introduction

Hainan gibbons—a member of the group *Hylobatidae nomascus*—and Billy Gibbons—a member of the group *Z Top*—have a lot in common. They are both primates, tool users, social organisms, and singers (Deng & Zhou, 2016; Guan et al., 2018). However, only Billy Gibbons uses a guitar to accompany his singing, only Billy Gibbons collects cars, and only Billy Gibbons sells a brand of hot sauce ("Billy Gibbons," February 11, 2023).[1] Indeed, Billy Gibbons, as a member of the human species, is a peculiar social animal for responding to the world in many cognitively unique ways, including by relying on representations of what products will sell well, the norms of musical traditions like blues rock, and complex technologies like seven-gauge guitar strings. In short: humans are not just any old thinking thing—they are a special kind of thinking thing.

Of course, saying that humans are cognitively unique is bordering on a triviality. Every type of organism is unique in its own ways: only gibbons use long arms to move by brachiation through tropical and subtropical forests in southeastern Asia while engaging in terminal branch feeding (Guan et al., 2018). What is far from a triviality is characterizing the nature of human cognitive uniqueness and explaining its evolution. Why do humans think in the ways that they do, and why do other organisms not do the

[1] It may be thought obvious that they also differ in the fact that Hainan gibbons are a group of organisms, whereas Billy Gibbons is an individual organism. However, since Hainan gibbons are one of the rarest mammals on earth—currently numbering around 30 individuals (Zou et al., 2022)—it is, unfortunately, not guaranteed that they will continue to differ in this regard.

It's Only Human. Armin W. Schulz, Oxford University Press. © Oxford University Press 2025.
DOI: 10.1093/9780197800188.003.0001

same? How did we end up being the kind of thinkers that we are—and exactly what kind of thinkers are we? While it is easy to come up with some candidate answers (social living, cultural learning, tool use), spelling out these answers and justifying them is not easy. Doing so is the aim of this book.

Given that we are humans, it is probably unsurprising that we would find this to be an inherently interesting project: we care about getting to know ourselves better. Knowing more about how we think and act tells us something important about who we are. It is thus to be expected that the nature and explanation of our distinctive ways of thinking and acting is a central topic in much of the history of philosophy and literature and is the core subject matter of anthropology.

However, more than this is true: knowing more about how we think and act is useful for helping us make better decisions. If I know that I will be unable to resist eating chocolate if any is around, I can take steps to avoid this—by not buying any. More generally, if we know that, qua being human, we are prone to feel X or think Y when faced with circumstances C, then we can use this knowledge to better handle circumstances C: we can try to counter-balance X and Y with other thoughts (if we want to avoid X and Y), or we can avoid being in circumstances C if we cannot counterbalance X and Y. In this way, the investigation of distinctively human cognition is also instrumentally useful for us: it can help us find ways to make us lead happier lives, with less strife and suffering.

Beyond these human self-centered reasons, the determination of the nature and explanation of distinctively human cognition is important due to the fact that humans have an outsized influence on the planet. Through our thoughts and actions, we shape the world around us. Our trading and industrializing activities affect the climate. Our engineering activities change the physical and biological structures on the planet. Our technological progress affects the lives of non-human animals. Understanding the nature of human cognition is therefore important also for helping us deal with these issues. Climate change, war, and disease prevention are not just issues that *happen to* humans. These are issues that *happen because of* human ways of thinking and acting. Knowing more about the latter can help us do a better job managing the former.

In short: making progress in understanding what distinctively human cognition consists in, and how it came to be what it is, is both intrinsically and instrumentally important—and that for both us humans and other living things on the planet. This makes an inquiry into distinctively human cognition timely and fascinating. However, from the get-go, it is important to be clear on the goals of any such project.

Detailing all of the distinctive ways in which humans think and act would go beyond the confines of one book—indeed, it would go beyond the confines of an entire library. Every human thinks and acts in slightly different ways, and, depending on the grain of analysis, there are indefinitely many aspects of these different ways of thinking and acting that could be noted. For example: we could focus on the fact that humans produce sounds for entertainment, or that they play musical instruments, or that they play string instruments, or that they play guitars, or that they play electric guitars, or that they bend the string just so when playing the electric guitar. In general, listing and explaining all the ways in which humans are cognitively unique would be an endless and fruitless task.[2]

These different ways of thinking and acting need to be organized and systematized: central human ways of thinking and acting need to be distinguished from peripheral ones, and the former need to be given an appropriate characterization and explanation. For this reason, the goal in what follows needs to be understood properly: it is not a *full* account of *all* aspects of the nature and evolution of

[2] Carruthers (2006) lists 22 aspects of distinctively human cognition, but Carruthers (2013, p. 326) notes explicitly that this list "should really be much longer." Brown (1991) gives a list of human universals including several hundred items, but, as will be noted momentarily, even this is a narrower list than what could be seen to be relevant here.

distinctively human cognition. This sort of account cannot be provided here or anywhere.

Rather, the goal in what follows is different. It is to sketch an explanatory schema of sorts: the goal is to make clear the *principles* to appeal to when making sense of the nature and evolution of distinctively human cognition. While the application of these principles to some key aspects of distinctively human cognition is and needs to be part of this elucidation, many other such applications will be left open. That is, the aim of this book is first and foremost to detail a *methodological framework* for approaching the question of the nature of human cognitive uniqueness. Of course, this does not mean that is not looking to provide key details of the nature of distinctively human cognition. The point is just that the present inquiry is not a bean-counting exercise ("this is uniquely human, and that is uniquely human. . . ."), but an attempt to provide a principled method for explaining the nature of distinctively human cognition.

In this context, it is also important to note that the present account is theoretical in nature. It is based on synthesizing work in many different fields—anthropology, psychology, economics, and philosophy, to name just a few—into a coherent overarching picture. In this sense, it is not based on a specific set of data (as is true of Tomasello, 2021) or a particular ethnographic or comparative cultural study (as is true of Henrich, 2015). Instead, it is driven by argumentative, theoretical considerations. That is to say, this is a work of philosophy: it is trying to develop a big picture overview of the kinds of thinking things that humans are.[3]

This, though, does not mean that what follows is pure speculation. In fact, the difference between pure speculation and argument-driven theoretical evidence is greater than that between argument-driven theoretical evidence and argument-driven empirical evidence. After all, every experiment needs be embedded

[3] For a set of anthropological perspectives that are broadly consistent with the theory laid out here, see, e.g., part 2 of Desmond and Ramsey (2023).

in auxiliary assumptions and theoretical background knowledge. There is no *principled* difference between empirically driven inquiries and theoretical ones; the question is really only whether the relevant arguments have sufficiently strong logical and empirical backing. While there is not always a hard and fast way to tell which arguments are compelling and which are not, the key here is just that it is not justifiable to brush aside the rest of the book as mere philosophical speculation.[4]

These points are important to stress, as projects like the present one are often dismissed as being overly speculative and scientifically unconvincing. From Kitcher's (1985) dismissal of "pop sociobiology" to Buller's (2005) takedown of "Evolutionary Psychology," many authors have been very critical of attempts to explain the evolution and nature of human cognition (see also Dupre, 2001). In particular, these authors are concerned that, if done carefully, theories of distinctively human cognition result in insights that are minor and narrow only. Since thoughts do not fossilize, since we do not know a whole lot about how ancestral hominins lived, and since the minds of other animals are very difficult to read, it is difficult to make claims in this context that are nontrivial *and* empirically convincing (or so these critics worry). Put the other way around: there is a concern here that conclusions that are exciting and groundbreaking are bound to be the result of wild and dangerous speculation only.[5]

Undoubtedly, these critics have raised some important issues: it is true that thoughts do not fossilize, that we do not know a whole lot about how ancestral hominins lived, and the minds of other

[4] Sterelny (2021) argues that theoretical accounts are not to be brushed aside because they are not so easy to provide: it is not trivial to get *all* the relevant details right. This may be the case, but something more is true as well: there are also constraints on what it takes to provide a compelling account *of a given set of data points*. Not all such accounts need to be created equally: for example, some fit better to *other* such accounts. See also Wylie (2002).

[5] For a related set of criticisms concerning evolutionary biological accounts of human nature more generally, see, e.g., Lewens (2015) and Downes (2018).

animals are very difficult to read. However, these facts should not be overemphasized either. All too often, works like those of Kitcher (1985) and Buller (2005) have left readers with the impression that exploring the evolution and nature of human cognition *must be* scientifically shallow and uninteresting. This is not the case, though. Rather, the facts that thoughts do not fossilize, that the living situations of ancestral hominins are unclear, and that non-human minds are difficult to scrutinize (among others) are *difficulties* that a theory of the characterization and explanation of distinctively human cognition has to grapple with. However, this does not mean that they are fatal to such a theory. This is parallel to the situation in other sciences: the fact that we cannot (yet) travel to other planets makes astronomy more difficult, but not impossible.

Indeed, by drawing on the rich toolkit of contemporary interdisciplinary science—from game theoretic economic models and comparative psychological experiments to gene–culture coevolutionary studies and archeological explorations—these problems can, while not being overcome, at least be lessened so that a compelling theory of distinctively human cognition can be developed. The rest of this book makes this clear in detail. For now, it is sufficient to note that the goal in what follows is to provide a rigorous, argument-driven, but also empirically well-grounded novel theoretical framework for making sense of distinctively human cognition—nothing more, but also nothing less.

At the heart of this framework is a complex feedback loop among innate, evolved mental structures, processes of cultural learning, and sociotechnological development. This perspective combines the most compelling elements of existing treatments of distinctively human cognition, including modular, nativist processes and states, cultural and individual learning, and 4E (embedded, extended, ecological, and enactivist) forms of cognition. The upshot is a remastered, twenty-first century nativism that is furthermore intricately interwoven with a twenty-first century techno-empiricism. As I try to show in the rest of this book, this is an exciting, fruitful, and empirically compelling theory of distinctively human cognition.

The book is structured as follows. Chapter 2 sets out an inventory of what makes human cognition special. It argues for a multidimensional, culturally variable account of human cognition that acknowledges that the differences between human and nonhuman cognition are matters of degree, not kind. However, the chapter also shows that this sort of view still allows for the notion of distinctively human cognition to be theoretically and empirically meaningful.

Chapter 3 sketches the major existing treatments of the nature and evolution of distinctively human cognition: modular, nativist views, cultural learning–based views, and symbol processing–based views. It lays out the benefits of these different views but also shows why they, on their own, fall short of providing a compelling account of the nature and evolution of distinctively human cognition. The chapter ends by providing a list of open questions that we still need answers to.

Chapter 4 then argues that underlying the evolution of distinctively human cognition is a complex, looping dynamic. The chapter describes this dynamic and shows how it can answer the questions left open from the previous chapter. At the heart of this dynamic is a positive feedback loop: sophisticated cultural learning makes possible the creation of tools that increase the sophistication of representational decision-making, which in turn allows for yet further increases in the sophistication of cultural learning and tool manufacture. The next three chapters then illuminate and expand on key aspects of this dynamic.

Chapter 5 notes that humans stand out from other organism in their ability to mentalize: they attribute mental states with high degrees of abstractness and complexity, and they do so frequently in situations of major importance. However, their mindreading abilities are also less than fully reliable and culturally variable. This chapter proposes a novel account to explain these facts. Since human mindreading is cognitively very costly, further cognitive and social tools are needed to make it work efficiently (or at all), even given an evolved psychological machinery for mindreading. The upshot of this is a feedback loop involving mindreading, cultural learning, and sociotechnological development.

Chapter 6 begins by arguing that human cognition is unique in its dependence on often highly complex and abstract *moral* concepts. It is also widely known that there is massive variation in the ways in which humans think when it comes to morality—at least on a surface level. However, while the outlines of the ways in which humans rely on moral concepts are thus clear, it is not clear what *explains* this reliance on complex and abstract moral concepts in human thought. Taking steps toward answering this question is the goal of this chapter. To do this, it argues that we need to see these two questions—Why do humans rely on complex moral concepts when other species do not? Why do they do it so differently within their own species?—as deeply related. Understanding why humans evolved in such a way as to often think in terms of complex and abstract moral concepts provides us with the key elements for an understanding of what leads human moral concepts to differ across humans and vice-versa.

Chapter 7 starts by noting that humans are also standouts in their propensity to trade. More specially, the kind of trading found in humans—featuring the exchange of many different goods and services with many different others for the mutual benefit of all the involved parties—far exceeds anything that is found in any other creature. However, a number of important questions about this propensity remain open. First, it is not clear exactly what makes this propensity so different in the human case from that of other animals. Second, it is not clear why other animals did not acquire this propensity to the extent that humans did. Third, it is not clear what explains the fact that the extent to which humans engage in trade is culturally highly variable. The chapter argues that at the heart of the human–animal divergence in this propensity is the particular sociocultural environment in which humans evolved. This has led them to sometimes, but not always, acquire the cognitive technology (writing, algebra, tallying devices, money, etc.) to support a sophisticated disposition and capacity for reciprocal cooperation and deep and wide concepts of property and exchange value.

Chapter 8 of the book develops some applications and conclusions of the account sketched in the previous chapters. It begins by summarizing some key results from the discussions of Chapters 4–7 and returns to the questions left open in Chapter 2. It then considers three applications of this account to some issues of contemporary importance: the relationship between distinctively human cognition and increasingly sophisticated artificial intelligence (AI), the role of religion in human thought and action, and the relationship between extensive patent regimes and human innovation.

The final chapter of the book pulls together all the strings developed in the previous chapters and summarizes the overall picture that results. It also revisits some of the methodological themes laid out in the book.

All in all, this book is a deep dive into the evolution of human cognitive uniqueness, the result of which is the insight that underneath human cognitive uniqueness is not one key evolutionary innovation but a tangled web of processes and states dynamically responding to each other. I hope it is of value to researchers in many different areas—for, in some ways, the evolution of human cognitive uniqueness touches many different topics.

2

Human Cognitive Uniqueness

An Inventory

In a broad sense, setting out ways in which humans are cognitively unique is easy. Only humans play tennis, only humans write philosophy books, only humans act in romantic comedies. However, as also noted in the previous chapter, what is far from obvious is *characterizing* the nature of human cognitive uniqueness: What categories do the many specific things that only humans do fall under, and which uniquely human cognitive abilities underlie these categories? Answering these questions is important for two reasons.

On the one hand, this is a useful first step toward *explaining why* humans are unique in these ways. It is generally difficult to explain a phenomenon without getting clearer on what the phenomenon is that is the target of the explanation. Of course, it is at least sometimes possible to get clearer on exactly what a phenomenon is *by* providing an explanation of it (indeed, as will become clearer in the rest of this book, something like this is also true in the present context). However, it is typically a good idea to provide at least a rough outline of what it is that we seek to explain; otherwise, this is theorizing in the dark. Hence, in what follows, I shall provide exactly this kind of outline.

On the other hand, clarifying the nature of distinctively human cognition is inherently interesting. Human cognitive uniqueness is a premise in many inquiries, from anthropology and primatology to zoology and psychology; hence, providing a characterization of it is useful for providing foundations to these inquiries. This is

It's Only Human. Armin W. Schulz, Oxford University Press. © Oxford University Press 2025.
DOI: 10.1093/9780197800188.003.0002

especially so since it is controversial whether this premise can be spelled out in a compelling way at all—that is, whether there even is such a thing as a human nature, cognitive or otherwise (for a good overview of this debate, see, e.g., Downes, 2018; see also Heyes, 2012b; Machery, 2008; Penn et al., 2008).[1] As this chapter shows, though, it *is* possible to provide a compelling characterization of distinctively human cognition. This is thus an important and noteworthy result in and of itself (see also Carruthers, 2013; Tomasello, 2021). I return to this point in later.

For these reasons, providing an inventory of the cognitive features that are uniquely human turns out to be a highly worthwhile activity. This is the goal of this chapter. Section 2.1 presents some important background considerations about constructing an inventory of key distinctively human cognitive traits. Section 2.2 then sets out this inventory. Section 2.3 expands this inventory by sketching further distinctively human cognitive suites resulting from *combinations* of the other elements of this inventory. Section 2.4 takes a step back and brings out the relevance of providing such an inventory.

2.1 Some Background Considerations

Before beginning to construct an inventory of distinctively human cognition, it is important to lay out the major background considerations on which this construction rests. Doing so makes clearer what the principles are behind the inventory and thus helps to avoid misunderstandings.[2]

[1] Of course, the existence and characterization of human nature is a more general question than that of the existence and characterization of distinctively human cognition: for example, the former may need to appeal to bipedalism, opposable thumbs, and nakedness. Still, as made clearer in the last section of this chapter, this does not mean that an inquiry into distinctively human cognition is not interesting in its own right.

[2] In the famous quip, any classification system is satisfactory as long as it contains a category of "miscellaneous." However, a classification system would not be satisfactory if that were the only category it contained.

The first point to note here is that the focus in what follows is on *cognition*, not behavior. The goal is to explain how humans think, not how they act. This matters since, apart from their cognitions, there are many reasons for why people do what they do—including chance, physical constraints, and biological causes. Consider a drummer playing a seven-stroke roll on the snare drum. Why did they play the roll *just so*? Partly, this is due to the way the sticks bounced off the drumhead, which is influenced by the minor, random physical indentations of the head and sticks, as well as general (and partly chancy) physical principles. It is also influenced by the shape of the drummer's hand and body. On top of this, of course, it is influenced by what the drummer *wanted* to play: a quiet roll with an accent on the last hit, rather than a loud, even roll, say.

For present purposes, though, only this last set of causes of the behavior is relevant: the goal in what follows is the characterization and explanation of distinctively human ways of thinking. Of course, since these distinctively human ways of thinking are still (often) causes of human behavior—even if not the only causes—consultation of the latter is often important. For example, while explaining why humans engage in *particular hand movements*—a particular way of playing a seven-stroke drum roll, say—is not part of the project, it is part of the project to explain why people have the *motivation* to engage in these hand movements and why they have the *fine motor control* abilities to act on these motivations. In this sense, then, the investigation of human behavior will often be a key element in what follows. However, it is important to emphasize that, in the first instance, the focus is on cognition, not behavior: the target of the analysis is not the explanation of behavior but only of its psychological causes.[3] In this way, the present inquiry shares

[3] It is an insight at the heart of the (philosophy of) the social sciences that human cognition is deeply intertwined with human action (Rosenberg, 2012). Some of what people do is *defined* by people's thoughts: a wink is different from a blink solely by virtue of the agent's intentions. In this richly psychological sense, therefore, it often *is* the explanation of action that is the goal of the analysis here: at stake here is explaining why

focus with work in evolutionary psychology (Carruthers, 1998; Cosmides & Tooby, 1987) and differs from work in human behavioral genetics (which tends to focus on specific behavioral patterns; Plomin, 1990), say, or some work in anthropology and philosophy of anthropology (which tend to focus on explaining human institutions or ways of acting) (see, e.g., Sterelny, 2021).

The second major point to note before setting out an inventory of human cognitive uniqueness is that, since it is impossible to provide a complete description of all the unique ways in which humans act and think (for the reasons made clear in the previous chapter), the goal in the inventory set out below is just to state some of the major instances of human cognitive uniqueness. The focus will be first and foremost on those factors that are particularly useful to provide the foundation for the development of the principled account of the evolution of distinctively human cognition in the rest of the book. The goal is to make more concrete what we are trying to explain by providing some salient examples of the relevant phenomena. Unsurprisingly, those examples comprise the ones most typically cited in this context: mindreading, morality, reciprocal cooperation, culture, language, and tool use (to name a few). Still, it is worth emphasizing that no attempts at a systematic and complete literature review are made here; the goal is just a sketch of some of the—for present purposes particularly useful—elements of distinctively human cognition.

That said, given that the inventory set out in the rest of this chapter is not an unsorted laundry list but a structured, principled categorization, other elements of distinctively human cognition can easily be added to the account at a later date. (For example, this is true for religion and music, which will get a more detailed treatment in Chapter 8.) In this sense, this inventory is less a finished product and more of a model to be used continuously. Indeed, as

people wink, not why they blink. However, this does not alter the point in the text precisely because actions and behaviors, on this construal, are not the same.

also noted in the previous chapter, since the goal here is to provide a *principled* account of distinctively human cognition, the inventory can be seen as a list of samples to be expanded as needed.

The third point that it is important to stress here is that, in what follows, it is not presumed that everything that is uniquely human is also something that is shared by all humans. "Uniquely human" here is intended to refer to a significant feature of the thoughts and actions of some humans—and one that is not shared in quite this way by non-human animals. In other words, the inventory sketched in this chapter is not an inventory of human universals: it is broader than that and also includes features that differ across human groups. There are two reasons for this.

On the one hand, variation is a key part of human thought and action and needs to be placed at the starting point of any inquiry into human cognitive uniqueness. Even in classic and stereotypically human ways of thinking and acting, there is much intraspecific variation. For a particularly obvious example, consider the fact that only humans read and write—but many humans do neither. However, not seeing literacy as something that is distinctively human (due to its lack of universality) is not compelling. On the other hand, there are important bio-cultural-technological mechanisms underlying this variation in human thought and action. Indeed, bringing this out is a key aim of the present book. To show this, though, it is important to avoid restricting the present inquiry to human universals (if there are any).

The fourth point that is worth stressing from the beginning—though it will also be made clearer below—is that no claim is made here that everything that is uniquely human is different *in kind* from what is found among non-human animals. The differences in question can be merely ones in degree: humans may be doing certain things to a greater or lesser extent than non-human animals. For example, the fact that some non-human animals may show at least precursors of moral cognition is not taken to mean that moral cognition is not a key aspect of what makes humans cognitively unique. As will be made clearer in Chapter 6, we think morally more often, about more bioculturally important issues, in a more complex and abstract manner. The fact that this is only a difference in degree does not lessen its importance. (I return to this issue at the end of Chapter 4.)[4]

I return to the last two points in Section IV. For now, though, it is best to get started on using these background considerations to set out an inventory of distinctively human cognition that can provide the foundation for the discussion in the rest of the book.

2.2 An Inventory of Human Cognitive Uniqueness

To develop this inventory, it is best to begin by noting that human/non-human cognitive differences can be divided into two categories: low-level and high-level. Low-level cognitive traits, as they are understood here, refer to cognitive abilities that are nonrepresentational, generally subpersonal, and generally automatic (Fodor, 1983; Kahneman, 2003; Leitgeb, 2004). Classic examples of low-level cognitive traits are an organism's discounting rate (the degree to which future rewards are devalued compared to present rewards), its degree of impulse control (the extent to which it can resist smaller present rewards for bigger future rewards), and its extent of executive functioning (how much concentration and attention it has available for how long, and how much control it has over these abilities).

These abilities are marked by the fact that we tend not to be aware or in control of their operation—we just know that we prefer $100 today to $110 a year from now and may not be able to change that preference easily. Most importantly, they also operate without

[4] Indeed, arguably, mere differences in degree—and not kind—are at the heart of the biological world in general: Godfrey-Smith (2009). See also Chapter 4 for more on this.

relying directly on *mental representations* (i.e., "concepts"; on a common construal of this notion see, e.g., Margolis & Laurence, 2015; but see also Machery, 2009): we do not represent our discount rate and use that to infer that we prefer $100 today to $110. Rather, it is the opposite way around: we have to generate a representation of our discount rate from a consideration of our preferences.

It is worthwhile to note immediately, though, that it is not presupposed here that there is a strong and clean difference between these kinds of cognitive abilities and the more high-level representational ones that will become important again below. For starters, there is still much debate about how best to understand the nature of mental representation (Dretske, 1988; Fodor, 1990; Millikan, 1984, 2002; Papineau, 1987; Prinz, 2002), and it might turn out that some or all low-level traits end up being representational *in some sense* as well.[5] Similarly, people can and do alter many of their preferences and valuations and may sometimes be quite strongly aware of them. In general, the difference between what is automatic, subpersonal, and nonrepresentational and what is not is one of degree only. However, the fact that the low-level/high-level distinction is one of degree does not mean it is not a useful distinction to make. Rather, it just needs to be conceived properly: low-level and high-level cognitive traits are parts of a spectrum, and the classification of a trait as low-level (or high-level) should not be seen to mark it as *essentially* different from other traits.[6]

On top of these points, it is important to note that some cognitive differences between human and non-human animals crosscut the low- and high-level distinction. This is due to the fact that low-level and high-level traits can combine to make up a complex

[5] For example, it is possible to appeal to purely perceptual concepts (Barsalou, 1999; Millikan, 1984, 2002; Prinz, 2002). See also below.

[6] Note also that this distinction does not concern whether a form of cognition is conscious or not (whatever, exactly, that means): many representational forms of cognition are still subconscious. In this sense, the setup here is different from that in Donald (2001).

cognitive suite (see also Carruthers, 2013; Heyes, 2018, 2012b). For example, a difference in the extent to which an organism is able to be a "tool user" might depend on *all* of the following: differences in the attention span the organism has (to concentrate on using a tool), its willingness to forgo present rewards to engage in complex learning tasks with uncertain future outcomes (to learn how to use the took), and fine motor control (to actually manipulate the tool), as well as its ability to represent the causal processes underlying the tool (to use the tool in the right circumstances and for the right purposes). The former two are low-level traits, but the latter is a high-level trait. These points will become important again below, but, for now, it just important to note that the same low-level differences can matter for several high-level differences, further supporting the usefulness of the low-level/high-level distinction.

With this in mind, consider some of the major low-level cognitive differences between human and non-human animals, first, and high-level differences, second. After that, some key cognitive suites are laid out.

2.2.1 Low-Level Differences

Humans can be characterized by the existence of a number of low-level cognitive differences compared to non-human animals. These differences are not always seen as central to what makes humans distinctive thinkers, but they do deserve some discussion and consideration nonetheless—especially due to the fact that they play a role in many distinctively human cognitive suites.

2.2.1.1 Attention and Concentration

Compared to non-human animals, humans have a distinctive attention and concentration span (Desrochers et al., 2016; Guo, 2016; Robertson et al., 1996). Many animals can focus on certain aspects of their environment intensely and for long periods: for example, cats and many other ambush predators can focus their attention on their target for long periods and "drown out" environmental distractions. Humans are thus not special in their ability to focus on parts of their environment and remain thus focused for a significant period of time. Rather, they are special in a different set of ways.

First, they are unique in the *range of aspects of their environment* that they can focus on. Humans can focus pretty much on anything—video games, musical instruments, or distant planets. Their attentional and concentrational processes are unconstrained in a way that is not true for other organisms whose attentional and concentrational processes are focused on prey, predators, mates, or food (Heyes, 2018).

Second, humans can *direct* their attention and concentration more freely. While many animals can learn to concentrate on something if doing so is rewarded appropriately, humans are special in being able to concentrate on something even in the absence of a reward (such as Galileo attending to the movement of the planets even though this was not being rewarded—on the contrary). Put differently, it is not just that our attention and concentration can be directed *at* a wider variety of objects—we also have relatively more *control* over what they are directed at, and we are intrinsically motivated to focus on a wide variety of objects (Kidd & Hayden, 2015).[7]

2.2.1.2 Patience

Many animals are exceedingly patient while engaged in various tasks. However, humans are unique when it comes to the *scale* of

[7] Note that there may also be some reason to think that humans are unique in being more easily able to *jointly* attend to something (Tomasello, 1999, 2021, 2022). However, as noted below, this is better seen as a higher-level trait, and there is also some debate exactly how pronounced this difference is really (Boesch, 2005). See below for more on this.

their patience. For example, they can wait *much of their lives* for certain rewards (going on a trip around the world after they retire from work). It is not just a matter of postponing a reward to the foreseeable future: humans can postpone a reward to the unforeseeable future. This affects their learning rates: for humans, rewards for tasks do not need to be temporally close to the task and thus they are able to learn things that other animals cannot (Allen, 2018; Heyes, 2018).

2.2.1.3 Social Toleration

Even highly social animals tend to be very picky about the presence of others in many circumstances. In contrast, humans tolerate and in fact invite—at least sometimes—the presence of others into very intimate spaces. For example, human experts allow apprentices access to their work spaces where they can acquire much useful knowledge from the "master" (Sterelny, 2012a). This is not something that is found on this scale in non-human animals: the latter tolerate the presence of others in more limited circumstances (Sterelny, 2021), whereas humans tolerate the presence of more others in more circumstances.

2.2.1.4 Fine Motor Control

Relative to their size, humans are able to manipulate quite small objects extremely carefully (Carruthers, 2013). Other animals are more restricted in the extent to which they are able to do so. By contrast, humans can control their hands to make baskets, blankets, clothes, and other tools. This is, of course, partly due to physiological differences: since, famously, humans have opposable thumbs, we can hold and manipulate physical objects in ways that other animals cannot. However, it is also partly a cognitive difference: we can use our hands and fingers in very specific ways—we can learn to play guitar, say, or learn how to manipulate a pen in highly specific ways to write Mandarin characters.

2.2.2 High-Level Differences

It is also plausible to see humans as characterized by a set of distinctive, high-level cognitive traits: psychological abilities that tend to be representational, personal, and voluntary. These distinctive high-level traits tend to feature centrally in debates concerning uniquely human cognition, and while—as just noted and as will also be made clearer momentarily—they do not exhaust the sphere of distinctively human cognition, they are indeed very important aspects of the latter and do need to be carefully considered here.

Exactly what mental representations are is still debated.[8] However, for present purposes, any generally satisfactory account will do. What is important here is just that mental representations can involve concepts of different degrees of complexity. On one extreme, they can involve purely perceptual concepts: the organism makes its behavior dependent on how it perceives the world (Millikan, 2002; Schulz, 2018b). On the other extreme, they can involve highly abstract and complex concepts—that is, concepts that only have a highly tenuous connection to perceptual states and/or which are constructed out of other (perhaps themselves complex) concepts (Fodor, 1990; Prinz, 2002). Key among these latter concepts are indexical ones, self-referential concepts, aesthetic and moral concepts, scientific and metaphysical concepts, epistemic and mental concepts, and mathematical concepts (Carey & Spelke, 1996; de Hevia et al., 2014; Gilbert, 2018; Millikan, 1989; Papineau, 2003). Two further points about representational cognition are crucial to note here.

[8] Similarly, it is also still debated whether human (or non-human) thought should be seen to be representational at all (Brooks, 1991; Chemero, 2009; Ramsey, 2007; Van Gelder, 1995). However, the consensus is strongly in favor of some kind of representationalism—and that for good reason (Allen, 1992; Schulz, 2018b; Stich, 1990). However, it is important to stress that the present acceptance of representationalism is consistent with many different versions of this sort of view and does not need to commit to narrow forms of a language of thought hypothesis, for example—a point that will be made clearer below and in Chapters 3 and 4.

First, the distinctiveness of human cognition is not well seen to consist in its being conceptual—or even conceptual to a high degree of abstraction—*tout court*. Many non-human animals also rely on mental representations (Carruthers, 2006; Millikan, 2002; Schulz, 2018b; Sterelny, 2003).[9] Some non-human animals might even sometimes rely on relatively complex and abstract mental representations (Allen, 1999, 2014; Allen & Bekoff, 1997; Andrews, 2015, 2012): for example, they may represent a conspecific as *dead* (Allen & Hauser, 1991), or they may meta-represent conspecifics as being *intentional agents* (Andrews, 2018; Tomasello, 2022).

Rather, what is uniquely human is the reliance on particularly abstract and complex concepts in particularly many different decision situations (Millikan, 1984, 2002; Papineau, 2003; Schulz, 2018b).[10] A large number of the decisions humans make feature a large number of particularly complex and abstract concepts (e.g., whether a [[RANDOM] [SAMPLE]] is [REPRESENTATIVE] of its [POPULATION], or whether an [INHERITED] [[PROPERTY] [ARRANGEMENT]] is [JUST]). In other words, the distinction between human and non-human cognition here is one of degree, not of kind. It is not that humans can do something—such as have abstract thoughts—that non-human animals cannot. Rather, the

[9] The evolution of representational cognition, in general, is plausibly driven by cognitive efficiency (Lieder & Griffiths, 2020; Schulz, 2018b). Relying on mental representations allows an organism to streamline its cognitive and neural system in specific ways: instead of relying on a large battery of individual perception–action connections that need to be maintained and updated individually, it "chunks" some of these connections and/or *infers* the appropriate behavioral response to the situation it is in. This streamlining allows an organism to save costly neural and physical resources as well as to save time adjusting to changes in its environment (Lennie, 2003; Levy & Baxter, 1996; Lieder & Griffiths, 2020; Niven & Laughlin, 2008; Schulz, 2018b; Wang et al., 2016). Importantly, though, these benefits of the reliance on mental representations are offset with costs: reductions in decision-making speed and increases in the reliance on cognitive resources like concentration and attention (Coolidge & Wynn, 2009; Epstein, 1994; Greene, 2008; Lieberman, 2003; Ramsey, 2014; Schulz, 2018b; Wynn & Coolidge, 2011). It is the balance of these costs and benefits that, plausibly, is a key driver of the evolution of representational decision-making (Lieder & Griffiths, 2020; Schulz, 2018b). This applies to human as well as to non-human animals.

[10] This is not to say that all decisions that humans make must feature many highly abstract and complex concepts—just that many (including many everyday ones) do.

point is that humans can think more thoughts that feature more concepts, many of which are more abstract or complex than what is true of non-human animals. The issue that needs to be explained, therefore, is why humans *frequently* think with *particularly abstract and complex concepts*, as well as *which concepts*, in particular, they tend to so employ.

Second and relatedly, it is reasonable to see human representational cognition as *systematic*. It is not just the case that humans can build concepts out of other concepts, some of which are highly abstract. It is also plausible that different conceptual mental representations are related to each other in a structured manner. That is, humans are thus that, if they are able to form representations of the form "Dogs like cats," they can also form representations of the form "Cats like dogs" (though whether they will in fact form either of these representations is an open question) (Calvo & Symons, 2014; Fodor, 1975; Fodor & Pylyshyn, 1988; Penn et al., 2008).

However, it needs to be noted that it is controversial to what extent humans are unique in this systematicity of thought (as has been suggested by Penn et al., 2008; Read, 2008; see also Carruthers, 1998).[11] On the one hand, there seem to be limits to the kind of conceptual combinations that are entertained by humans (as well as by non-human animals). On the other hand, the exact extent of systematicity and compositionality in other species is controversial, as is its importance (Beck, 2018; Calvo & Symons, 2014). Still, independently of how this debate turns out, it is reasonable to see humans as *often* generating novel thoughts and concepts by recursively combining old thoughts and concepts. Whether or not there is a discontinuity here (as argued for, e.g., by Penn et al., 2008), it is at least true that there is a large difference here—even if that difference is only continuous. This issue will

[11] In the background here is the large debate surrounding the reasonableness of something like a "generality constraint" on concept possession: see, e.g., Camp (2004).

become important again in the next chapter, but, for now, this is all that is needed.

With this in mind, consider some of the major high-level, distinctively human cognitive traits commonly discussed in the literature (see, e.g., Carruthers, 2013; Tomasello, 2021). While some of these traits will be discussed in more detail later on in the book, it is useful to start with a brief overview of them here.

2.2.2.1 Personal Memory

Humans are not distinctive in the *amount* of memory they have (how many things they can remember).[12] Many animals can store the locations of many food sources, hiding places, and other features of their environment. What makes humans unique is the types of memory they have. In particular, humans are unique in being able to remember their conspecifics particularly well: humans remember who did what when, and they do so for long periods of time. Animals can associate others with certain valences (dangerous, friendly, etc.), and they can remember who shared food in the past (Carter & Wilkinson, 2013). However, humans are unique in being able to remember a large variety of facts about their conspecifics: what they wore years ago, what they said, who they were talking to. That is, human personal and social memory is uniquely expansive (Chater, 2012; Hammerstein, 2003).[13]

2.2.2.2 Causal Cognition

It cannot be ruled out that non-human animals can represent some events A and B as being causally related. However, there is no question that humans are especially good at appreciating and

[12] Read (2008) argues that humans are unique in the amount of working memory they have; this may be true, but, as noted earlier, it is not clear that this truly marks a discontinuity between human and non-human animals here (see also Wynn & Coolidge, 2011).

[13] Humans may also be unique in the kind of episodic memory they have—but this is far from well-established, and the exact nature of episodic memory is still controversial as well. See also Schulz and Robins (2022) and Gershman and Daw (2017).

learning about the difference between two events being merely correlated and them being causally connected (Gopnik et al., 2004; Papineau, 2003).[14] Indeed, the representation of causal relations is found in many cultures (Abarbanell & Hauser, 2010), underlies the human ability for thinking about the physical environment (Carey & Spelke, 1996; Carruthers, 2006, 2013), and is a key element in human tool use and manufacture (a point to which I return below).

2.2.2.3 Social Cognition

It is not that humans are the only organisms that think about what others are thinking (Andrews, 2012; Parker et al., 1994; Suddendorf & Butler, 2013; Toda & Platt, 2015; Tomasello et al., 2005; Tomasello & Herrmann, 2010; though see also MacLean, 2016; Penn et al., 2008; Povinelli, 2003). However, even granting that humans are not unique in *being* mindreaders, they are unique in being mindreaders *of a specific sort*. In particular, humans can and do attribute mental states to others in ways that are cognitively very sophisticated: they attribute mental states at the third or even fourth level of intentionality, and they attribute complex mental states that are composed out of other, not necessarily atomic, mental states ("he is just hangry;" "that's just wishful thinking on her part") and which are highly abstract (without straightforward links to behavior, such as *nostalgia*). Moreover, humans use mental state attributions very often in their decision-making, and these attributions are often crucial factors determining the outcomes of these decisions.

2.2.2.4 Moral Cognition

Non-human animals at least arguably have some proto-moral concepts (de Waal, 1996; Flack & de Waal, 2000; Kitcher, 1998;

[14] Of course, this is consistent with the metaphysical claim that causation just is correlation of a specific kind. Here, this would merely mean that humans can distinguish different kinds of correlations from each other.

Monsó & Andrews, 2022; Prétôt & Brosnan, 2015; but see also Machery & Mallon, 2010). However even if the existence of proto-concepts in non-human animals is granted, it is uncontroversially true that humans rely *more frequently* on *more* and *more abstract and complex* moral concepts. Much of human thought and action plays out in a nexus of moral concepts—what we ought to do or be is a constant element of our practical deliberations. Further, human concepts like [FAIRNESS], [DUTY], or [VIRTUE] can attach to a wide variety of things, many of which are not straightforwardly observable. Human moral concepts can also involve a larger number of other concepts (such as [[HISTORICAL] [EFFORT]], [CHARACTER TRAIT], [[BEHAVIORAL] [EXPECTATION]]). Finally, among humans, *more* moral concepts are used—beyond [FAIRNESS] and [WELL-BEING], humans rely on concepts such as [RIGHT], [DUTY], [LIBERTY], [VIRTUE], and [MORAL RESPONSIBILITY].

2.2.2.5 Property Cognition

Many animals recognize property, at least in a rudimentary manner (Eswaran & Neary, 2014; Gintis, 2007; Kummer, 1991; Tibble & Carvalho, 2018). While the details of how to characterize property concepts—in general and in humans specifically—will be made clearer in Chapter 7, it is here sufficient to note that the human concept of [PROPERTY] is deeper and wider than that of non-human animals (Kanngiesser et al., 2020). Humans do not just consider who can prevent others from accessing which resources. Instead, they distinguish theft from gifting or borrowing on the grounds of whether the *claims* to the relevant good—whatever it may be—have changed (Brosnan, 2011; Gintis, 2007; Kummer, 1991; Scorolli et al., 2018; Tse, 2008, p. 287; Wilson, 2020). Very similar remarks hold for the concept of [EXCHANGE VALUE]: humans can conceive of virtually anything as at least potentially tradeable (artistic experiences, air rights, futures) and not just certain things (tokens, say) for certain other things (foodstuffs, say) (Brosnan et al., 2008; Oberliessen & Kalenscher, 2019; Parrish et al., 2013).

Overall, then, it is possible to group distinctively human cognitive traits into those are relatively more subpersonal, automatic, and nonrepresentational—such as attention and concentration, patience, social toleration, and fine motor control—and those that are relatively more personal, controlled, and representational—such as those of personal memory, causal cognition, social cognition, moral cognition, and property cognition. However, this is not where an inventory of human cognitive uniqueness should stop for, as noted earlier, these low-level and high-level cognitive traits combine to produce *suites* of distinctively human cognitive traits.

2.3 Uniquely Human Psychological Suites

A number of the most commonly noted features of distinctively human cognition are in fact collections of cognitive traits (Carruthers, 2013; Heyes, 2012b). These cognitive suites are the result of the presence of some low-level and/or some high-level cognitive traits that combine to form a novel, collective trait. Before considering some of these psychological suites in more detail, it is useful to note three points.

First, for present purposes, the exact metaphysical status of these suites is not greatly important. So, it can be left open here whether these suites are genuine (emergentist) cognitive traits in their own right or merely tightly integrated collectives of other cognitive traits. (Indeed, it is even possible to see some of the above higher-level cognitive traits as suites in their own right.) Determining this is made complex by the fact that the individuation of cognitive traits—like that of biological traits more generally—is far from straightforward (Baum, 2013; Nichols & Stich, 2003). Fortunately, such an individuation is not needed here. What matters here is just that some of the above low- and high-level cognitive traits are now recognized to interact in specific, integrated ways to produce a distinctive, novel set of cognitive and behavioral phenomena.

Second, many of these suites are interrelated with each other. In fact, this will be one of the major themes in the rest of this book. For now, though, these interrelations will not be central; what matters here is just providing brief characterizations of these suites *with which* their interrelations can then be further elucidated later on.

Third, there are two further distinctively human traits that it is worthwhile mentioning here, even though they will not get an explicit discussion in this chapter: religion and music.[15] These traits combine aspects of moral cognition, social cognition, cooperation, cultural learning, attentional and motor processes, and social structures of different sorts. The particular features of these cognitive suites come out best once the new account of the evolution of distinctively human cognition at the heart of this book has been laid out. Discussion of these traits is therefore postponed to Chapter 8 (though they also get some explicit consideration in Chapters 4 and 5). At any rate, it is worth emphasizing again that no pretense is made here that a "complete" list of distinctively human cognitive suites is given (were that even possible): the goal is just a model inventory that can be expanded as needed.

2.3.1 Language

The human ability to speak a language is a classic example of a human cognitive uniqueness—no other animal has it (though they may have some related abilities) (Friederici, 2017; Pinker & Bloom, 1990; Russon & Andrews, 2010; Sterelny, 2012c, 2017). It is now widely—though, as noted momentarily, not exclusively (see, e.g., Piantadosi, forthcoming)—seen to be the product of a

[15] Another one is humor, but the latter is not as yet very well understood. See also Olin (2016).

specialized set of mechanisms for learning words and grammatical structures partly based on innate representations (Chomsky, 1968, 1988; Fodor, 1983). However, it is also now recognized to be embedded in the uniquely human motivations to cooperate with and tolerate the presence of others. Speaking a language is inherently cooperative (Bloom, 2002; Grice, 1989; Skyrms, 2010; Sterelny, 2012c; Tomasello, 1999, 2021), and it needs to be understood as such: it is premised on the motivation to exchange information with others. In turn, this requires us to be sufficiently comfortable in the presence of others to make such an information exchange possible. Furthermore, speaking a language may be inherently connected to mindreading and certainly requires significant abilities of attentional control and the presence of a dedicated set of mental representations (see, e.g., Carruthers, 2006; Fodor, 1983; Pinker & Bloom, 1990). Finally, it dovetails with the human ability for cultural learning because it is premised on information exchange.

For what follows below, it is useful to note that while this book follows mainstream opinion in affirming the existence of specific, innate, representational presuppositions of language, this is not in fact an essential part of the argument here. In particular, the human ability to communicate in a natural language could be easily subsumed under the interactionist model laid out in the rest of the book even without these innate presuppositions: language would then be like other forms of learned abilities in humans, such as video-gaming or skydiving.[16] For this reason, the present acceptance of the standard, Chomskian nativist view of language should not be seen to unduly bias the argument here—though, as noted in Chapters 3 and 4, it does have some important sources of support.

[16] This would then bring the present view of language closer to that of Everett (2017)—though the details would still be spelled out very differently.

2.3.2 Cultural Learning

Humans learn many different things from others, from particular behaviors to high-level psychological traits (Boyd & Richerson, 2005; Henrich, 2015; Heyes, 2018; Nisbett et al., 2001; Piccinini & Schulz, 2019). A number of different models for how this cultural learning can proceed are defended in the literature. Some have argued that it is largely based on a specific underlying psychological trait, such as the sharing of joint attention (Tomasello, 1999) or an innate sense of pedagogy (Csibra & Gergely, 2011). Others have questioned this and argued that the ability for human cultural learning is just based on the confluence of a number of other specifically human cognitive traits, such as high social toleration and cooperation to allow for apprenticeship of sorts (Sterelny, 2012a), strong general learning abilities (Heyes, 2012b, 2016, 2018), or something else (Carruthers, 2006, 2013).[17]

However, for present purposes, it is not necessary to settle these questions. Independently of exactly what the nature of human cultural learning is—whether it is underwritten by a specific cognitive trait or the result of other psychological mechanisms—it is here sufficient to note two key points about the human ability and disposition to learn from others.

First, while some other animals (such as crows, rats, and chimpanzees) also seem to engage in various kinds of cultural learning (Brown, 2018), humans learn more and more complex facts from more and more different models (Boyd et al., 2011; Creanza et al., 2012; El Mouden et al., 2014; Henrich, 2015; Henrich & McElreath, 2011; Heyes & Galef, 1996; Laland & Janik, 2006; Sterelny, 2012a; Tennie et al., 2009; Tennie & Over, 2012;

[17] Similarly, individuals can learn from their parental generation, their own generation, or both, and they could learn from individual models (their best friend) or sets of people (what their friends do on average), and they can be subject to various biases in their learning (Boyd & Richerson, 2005; Godfrey-Smith, 2009, chap. 8; Henrich & McElreath, 2007; Heyes, 2018).

Tomasello, 1999; Tomasello & Herrmann, 2010; van Schaik & Pradhan, 2003). This is one of the main reasons that the appeal to cultural learning is widely thought to be a key element in the explanation of human cognitive uniqueness.

Second, human cultural learning can be cumulative. A behavioral or psychological variant can be improved successively by letting improvements achieved up to the present become the basis for future improvements (Tennie et al., 2009; Tomasello, 1999).[18] In turn, this brings the acquisition of behavioral or cognitive variants with many different and complexly interacting parts within the scope of human cultural learners—though they would be out of reach of individual learners (Boyd & Richerson, 2005; Henrich, 2015; Heyes, 2018; Legare, 2017, 2019; Tennie et al., 2009; Tomasello, 1999; Tomasello et al., 2005).

2.3.3 Reciprocal Altruism

Human cooperation is unique for being extensively reciprocal and extending beyond kin (Avsar, 2020, p. 16; Cosmides & Tooby, 1992; Griffin et al., 2004; Hammerstein, 2003; Hammerstein & Noe, 2016; Schulz, 2018a; Sober & Wilson, 1998; Sterelny et al., 2013; West et al., 2011). Indeed, reciprocal interactions provide the foundation of much of human social behavior (Boyd, 2018; Kaplan et al., 2012): through trade, humans have built social structures on a scale that is unheard of in the rest of the biological world (Avsar, 2020, pp. 52, 60–63, 90–95; Nowell, 2010; Ofek, 2001; Sterelny, 2012b, 2021). The propensity to trade, as a specific form of the capacity for reciprocal altruism, will be explored in more detail in Chapter 7; for now, it is sufficient to focus on the more general cognitive suite of reciprocal altruism.

[18] This is sometimes called "Tomasello's ratchet" (Tennie et al., 2009; Tomasello, 1999).

The human ability for reciprocal altruism is underwritten by a suite of cognitive traits.[19] Somewhat controversially, Cosmides and Tooby (1992) argue that it depends on a dedicated psychological machinery for detecting cheaters in situations of social exchanges. While there has been much debate concerning this idea (Barrett, 2015; Buller, 2005; Machery & Barrett, 2006), this debate also does not need to be settled here. What matters here is just that, in order to be capable and disposed to cooperate reciprocally, a set of high-level cognitive traits is needed—including mindreading abilities, personal memory, and *perhaps* some distinctive further representational expectations (such as sensitivity to violations of exchange conditions). Exactly what these high-level traits are can be left open here. On top of this, a number of lower-level cognitive traits need to be in place as well, including impulse control, social tolerance, and patience (Avsar, 2020, p. 72; Brosnan, 2011; Hammerstein & Noe, 2016; Kabadayi & Osvath, 2017; Rosati et al., 2006; Mulcahy & Call, 2006; Ofek, 2001, p. 142; 2013; Stevens & Stephens, 2008). Without the latter, humans could not act on motivations for reciprocal cooperation even if they were able to recognize or have them.

2.3.4 Tool Use

The manufacture and use of tools—that is, objects designed to fulfill a given function—is found in several different organisms (see, e.g., Cheke et al., 2011; Fragaszy et al., 2013; Hansell & Ruxton, 2008; Haslam, 2013; Mann & Patterson, 2013; McGrew, 2013; Sanz & Morgan, 2013 Shumaker et al., 2011). However, humans are able to build and use tools of particularly high degrees of complexity (Muthukrishna & Henrich, 2016; Režek et al., 2018; Shea,

[19] This goes some ways toward explaining why, in many non-human cases, the conditions making reciprocation mutually profitable seem satisfied yet there is no reciprocal cooperation (Boyd, 2018): these animals lack the needed cognitive traits.

2017; Vaesen, 2012; van Schaik & Pradhan, 2003; Wadley et al., 2009; Wimsatt, 2007). So, humans are able to build and use tools whose function it is to build *other* tools—they build robots that build computers that navigate airplanes. More abstractly, humans have also developed a device for storing and transmitting precise information about (nearly) anything: written language (Gibson & Ingold, 1993; Mullins et al., 2013b).[20]

As with the other cognitive suites here mentioned, there is debate about exactly what underlies this human ability for tool use and manufacture. It is widely agreed that it depends on cultural learning, causal cognition, fine motor control, and the ability to focus attention on freely chosen objects. Potentially, though more controversially, it also depends on a specific set of mental representations concerning the nature of technology (Gopnik et al., 2004; Martinez, 2013; Sterelny, 2012a; Tennie & Over, 2012). However, this last point can again be left open here: the key for present purposes is just that there is a set of specifically human cognitive traits—whatever, exactly, this set is of comprised of—that allows humans to build technology of high degrees of sophistication.

2.3.5 Creativity and Play

The final cognitive suite worth mentioning here concerns the distinctively human ability for creativity and play. Curiosity is a combination of the motivation to engage in exploratory behavior for its own sake and the ability to act on that motivation; for complex actions, this requires memory, patience, causal knowledge, and fine motor control, among other things (Bar et al., 2023; Kidd & Hayden, 2015; Loewenstein, 1994; Oudeyer & Kaplan, 2007). Not all organisms are equally curious (Glickman & Sroges, 1966; Kidd & Hayden, 2015;

[20] Even more abstractly (perhaps), humans also developed social institutions that can function as cognitive technology. This point will be made clearer in Chapter 4 (and the rest of the book).

Reader & Laland, 2003): indeed, humans are thought to be standouts even among mammals and even among primates (Glickman & Sroges, 1966; Kidd & Hayden, 2015; Reader & Laland, 2003).

In the first place, recent work in developmental psychology has shown that human infants are strongly and frequently driven by curiosity. From an early age onward, humans are happy to spend significant periods of time searching for insights about the world and how to manipulate it, all without getting much in the way of reward for it (Cook et al., 2011; Gopnik & Schulz, 2004; Kidd et al., 2012; Kidd & Hayden, 2015; Oudeyer & Smith, 2016). Recent work in neuroscience further confirms the importance of curiosity to human thought and action (Bar et al., 2023).

Importantly, moreover, work in archaeology suggests that curiosity being a strong and frequent motivator is not something that is restricted to young kids, but seems to characterize human living more generally.[21] So, for example, on a conservative estimate, intricately designed and decorated musical instruments appear in the archaeological record from about 40,0000 years ago (Adler, 2009;Conard et al., 2004, 2009; Cross, 2012; Killin, 2018; Lawson & d'Errico, 2002; Morley, 2013). Now, it is possible that some of these were built for signaling purposes: by building complex instruments, an instrument-maker may have displayed her motoric and cognitive skills and thus have hoped to obtain external rewards in the form of pay or social capital (Hayden, 1998; Kohn & Mithen, 1999; Premo & Kuhn, 2010; Renfrew & Scarre, 1998; Skyrms, 2010; Sterelny, 2012a). However, it is far from clear that this was always—or indeed ever—the case: there is little (if any) further evidence that musical instruments were indeed taken as marks of cognitive or motoric skill. It just as plausible—if not more so—that these instruments are simply the result of curious

[21] There is, of course, also reason to think that early humans collected information to be better able to obtain external rewards at a later date (see, e.g., Bergemann & Välimäki, 2008). However, since this does not concern *curiosity*—i.e., largely internally rewarded information-seeking behavior—this is not so relevant here.

experimentation: early humans may just have been interested to see if they could make a bone (say) make certain sounds (Adler, 2009; Cross, 2012; Killin, 2018; Morley, 2013).

Indeed, this idea can be extended even further back in time. Functionally useless, decorative handaxes appear in the archaeological record from about 500,0000 years ago (Mithen, 2005). Again, it is *possible* that these were designed to signal group membership, motoric and cognitive skills, or social status (Kohn & Mithen, 1999). However, again, there is little to support to this external reward–based view further. It is just as plausible—if not more so—that these artifacts were also built simply because their makers were interested in seeing if they could shape a rock just so, or how a rock shaped just so would feel if held or thrown (Currie, 2011; Wynn, 1993). In short, the act of making decorative handaxes—another very laborious exercise—plausibly was its own reward.

The idea that humans are inherently strongly curious is further supported by the fact that contemporary adult humans innovate even when by themselves and when doing so is not strongly socially rewarded (Kidd & Hayden, 2015). For example, humans appear to feel pleasure—an internal reward—when engaging in unrewarded information-seeking tasks (Kang et al., 2009; Kidd & Hayden, 2015; Perlovsky et al., 2010). It is plausible that this has been the case since very early times in hominin evolution. In short, it is plausible that, for a variety of reasons, humans are exceptionally curious and that this ability is constituted by motivations for intrinsic exploration, fine motor control, patience, and the ability to focus attention and concentration freely.

2.4 The Existence of Distinctively Human Cognition

All in all, therefore, it is plausible that human cognitive uniqueness can be characterized by the fact that it draws on a set of distinctive low-level cognitive traits—including attention and concentration, patience, social toleration, and fine motor control—and high-level traits—including personal memory, causal cognition, social cognition, moral cognition, and property and exchange cognition—that furthermore interact to build up distinctively human cognitive suites—including language, cultural learning, reciprocal altruism, tool use, and creativity.[22] As noted earlier, no claim here is made that this is an exhaustive list of distinctively human cognitive traits, nor is it claimed that this list is distinct in kind from what is found among non-human animals (although see also the discussion in Chapter 4). However—and this is the key point to note here—the list still shows that there *is* such a thing as distinctively human cognition. The latter is a phenomenon that deserves to be theorized about in its own right.

This is not a trivial point; indeed several scholars have explicitly denied it (see, e.g., Downes, 2018; Dupre, 2001; Kitcher, 1985). There are many aspects to this controversy: for example, some have questioned whether appeals to a human nature hide unjustifiable essentialist presuppositions (Kitcher, 1985; Sober, 1980). Others have debated whether requiring human nature to be universal is necessary to avoid making the approach theoretically uninteresting (see, e.g., Downes, 2018; Machery, 2008). In particular, it may be thought that allowing for intra-specific variation in human nature—which is needed to avoid problematic forms of essentialism—yields what is merely a trivial laundry-list of features that some humans display some of the time. This, the argument concludes, is not a theory; it is just a list of the phenomena to be explained.

In this context, what the above points make clear is that a variation-focused and gradual approach is not only reasonable, but also interesting and far from trivial. As the above inventory (and indeed the rest of this book) shows, it is in fact the case that variability

[22] For a related view, see also Laland and Seed (2021).

and continuity in human nature can be theoretically handled and illuminated in an interesting way.

In particular, the above set of cognitive traits is more than a laundry list of human–non-human animal differences: they are at the center of human living. To make sense of why humans do the things they do, we must determine how we think. Human actions are driven by human minds, so an understanding of the former depends on an understanding of the latter. This makes them different from some other human traits—the shape of our noses, say. The shape of our nose is not at the center of human living: explaining why and how people interact with their environment can be done without appealing to the details of this shape. In contrast, the consideration of distinctively human cognitive abilities like mindreading, morality, and trading *is* central to explaining why humans do what they do.

Similarly, while it is of course true that humans are biological organisms just like any other, and, while the ways in which humans think and act are not supernatural and have homologs in other organisms, this does not mean that the above traits are not interesting: they are still at the heart of explanations of how and why humans live the way they do. Indeed, showing this is one of the aims of the rest of this book.

Finally, it is of course also true that human nature comprises many other traits as well, beyond the cognitive ones at center stage here: for example, bipedal walking, functional nakedness, and omnivory. However, it is clear that language, reciprocal cooperation, mindreading, morality, tool use, and rich cultural learning are central aspects of human nature, too. This makes their investigation theoretically well-grounded and worthwhile, even if some other traits deserve closer investigation, too.

It is furthermore important to note that the present characterization of distinctively human cognition is not built on evolutionary presuppositions (as is true of some accounts of human nature: Machery, 2008). That is, it is not here assumed, from the start, that distinctively human cognition has a particular evolutionary history—one based on natural selection, say. This has the benefit of leaving the details of this history open: it can involve genetic evolution, but also cultural evolution (as well as drift and other nonselective factors). Indeed, showing that this evolution involves a complex interplay of different factors—including genetic, cultural, and technological ones—is the aim of the rest of this book. In this way, the present characterization of distinctively human cognition has the advantage of being theoretically relatively modest.

What is needed next is therefore exactly this: providing an account of how humans are standouts in all of the cognitive capacities mentioned here. The next chapter begins the investigation of what such an account should look like.

3

Explaining Distinctively Human Cognition

The Lay of the Land

Several major theories have been developed to explain the evolution of distinctively human ways of thinking. These theories generally fall on a spectrum from highly modular, nativist accounts to accounts heavily focused on cultural learning—although, as will be made clearer below, there are also some outliers that are not well classified as either nativist or learning-based. However, a closer look at these accounts reveals that, while they all contribute some important insights to the discussion, they are unable on their own to provide a fully compelling explanation of distinctively human ways of thinking. Showing this is the aim of this chapter.

In doing this, though, it is important to note that the goal in what follow is not a detailed review of every attempt to explain distinctively human cognition; rather, the focus will be on three major types of accounts, which together present a reasonably clear outline of the lay of the land here. Some further, more specific accounts (e.g., ones targeted at mindreading specifically, such as Zawidzki, 2013) will be discussed in later chapters. Further, the goal in what follows is not a detailed dissection of every aspect of the accounts that are being considered; such a detailed treatment is more easily given in the context of the discussion of specific cognitive traits such as mindreading or morality. Accordingly, this is a task for the later chapters in this book. Rather, the goal here is to *motivate* why a new treatment of the evolution of distinctively human cognition is

It's Only Human. Armin W. Schulz, Oxford University Press. © Oxford University Press 2025.
DOI: 10.1093/9780197800188.003.0003

needed in the first place. For this, a briefer, more general discussion is better suited.

The structure of the chapter is as follows. Section 3.1 critically discusses massively modular, nativist accounts of distinctively human cognition. Section 3.2 does the same for cultural learning–based ones. Section 3.3 considers an alternative account based on relational symbol processing. Section 3.4 brings the discussion together and summarizes what questions remain unanswered by the existing accounts.

3.1 Massively Modular Explanations of Distinctively Human Cognition

The first account of distinctively human cognition is nativist and massively modular in structure.[1] According to this account, the major aspects of distinctively human higher-level cognition are the result of dedicated neurocognitive mechanisms that operate largely in isolation from each other. In turn, these mechanisms have been selected for in hominin evolutionary history as a result of the socioecological environments humans lived in. (For some classic versions of such an account, see, e.g., Barrett, 2015; Carruthers, 1998, 2006, 2013; Spelke, 2022; Tooby & Cosmides, 1992.) A detailed version of such an account focused specifically on mindreading will be laid out in Chapter 5; for now, though, it is sufficient to note some general principles.

The main reason why evolutionary pressures are thought important to the evolution of such modular minds is a combination of natural selection and the limitations of domain-general learning mechanisms. Throughout their evolutionary history, humans

[1] This account was a much discussed theory in the 1990s and the early 2000s, but has since then seen less development. However, several of its key elements are still much discussed (see, e.g., Carruthers, 2013; Spelke, 2022), and it certainly still deserves to be taken seriously.

needed to respond to specific environmental contingencies and to do so from an early age onward. This includes physical aspects of the environment—whether two material objects can occupy the same space at the same time, say (Carey, 2011; Carey & Spelke, 1996)—and biological aspects—whether an albino tiger is still a tiger (Gelman, 2009). However, the key environmental aspects humans needed to respond to were social: as humans, we need to deal, first and foremost, with other humans, and we need to do this by seeing them as psychological, social beings (Tomasello, 2022). Since we lack major physical adaptations to our environment, we are dependent on others to make our way through the world. Doing this successfully requires determining what others know, what they want, and what their social roles and expectations are (Sterelny, 2003; Tomasello, 1999, 2021, 2022; Whiten & Byrne, 1997).

Making these kinds of determinations, though, is far from trivial: similar behaviors can result from many different mental states and social roles. If person A acts aggressively toward me, is the best response an attempt at reconciliation, a refusal to engage with them, or counter-aggression? The answer may be different for different forms of aggression, even by the same conspecific: it will depend on the motivations and beliefs of the conspecifics, our respective social situations, etc.

This makes it difficult to rely on domain-general learning mechanisms to acquire the right inferential strategies. In the worst case, this kind of learning may be impossible: in the famous phrase, the environmental stimulus may be too impoverished to make this possible (Carey & Spelke, 1996; Carruthers, 1998, 2006, 2013; Cosmides & Tooby, 1992). If we needed to be taught all the circumstances in which it is reasonable to attribute *jealousy* or *wishful thinking* to someone else (for example), we would need a massive set of learning data and a lot of time—none of which we have. More weakly, though, even if domain-general individual or social learning were in principle possible here, this kind of learning would seem to be inferior to the existence of nativist expectations (this is a version of the Baldwin effect: Weber & Depew, 2003). Learning takes time and may lead to costly errors—and this is especially so for domain-general learning. Hence, organisms with the right innate psychological expectations and processes have an adaptive advantage here because they can avoid paying these temporary and error-related costs (Carruthers, 2006).

However, this does not mean that modular accounts do not have any room for learning. In particular, the modular cognitive mechanisms they posit to have evolved need not be static, in the sense that the representations and processes they are based on may still be the result of much learning. The point is just that this learning is scaffolded in such a way that a particular set of learned representations or processes is *quickly and reliably* obtained (Carruthers, 2006; Spelke, 2022; Tooby & Cosmides, 1992). Putting this the other way around, human evolution may have led to the development of innate, specialized learning mechanisms that efficiently and easily lead us to acquire the appropriate mental states and processes, rather than to the development of these states and processes themselves.[2] Either way, though, the upshot is a reliably developing set of cognitive traits dedicated to specific domains and leading to the forms of distinctively human thought inventoried in the previous chapter.

It is furthermore important that this account conceives of these reliably developing, domain-specific sets of cognitive traits as "modules." The exact definition of a module differs across different versions of the account (for good overviews, see, e.g., Callebaut & Rasskin-Gutman, 2005; Carruthers, 2006; Nichols & Stich, 2003)—a point to which I return in the next chapter. However, for present purposes, these details are not greatly important. What matters here is just that, on any version of the modular, nativist

[2] It could also be that these learning mechanisms are domain-general—a form of Bayesian learning, say—but that they are set up (e.g., with an appropriate distribution of priors) to make the learning sufficiently efficient and reliable.

account of distinctively human cognition, the nature of specifically human thought lies in the fact that evolutionary pressures led to the existence of relatively independent cognitive traits that enabled humans to respond in biologically appropriate ways to specific environmental contingencies.

This account gets at several important points about distinctively human cognition and its evolution. In the first place and most importantly, the account notes correctly that it is plausible to see humans as having evolved a dedicated set of psychological expectations and processes. For the reasons just noted, it is just not plausible to see all forms of distinctively human cognition as resulting from individual or social domain-general learning. While it may not be entirely clear yet exactly which aspects of human cognition are innate in a relevant sense, it is plausible that a considerable number of them do have such an innate basis.

Second (and not unrelatedly), the account does well in handling the fact that many forms of distinctively human cognition can be found, at least in some degree, in most human cultures. In virtually all cultures, people mindread each other (instead of treating others as inanimate objects or animals), in virtually all cultures people engage in some kind of moral-normative thought, in virtually all cultures people build and use tools in some form (Brown, 1991; Carruthers, 1998, 2006, 2013; Tomasello, 2021). These forms of thought develop reliably and cross-culturally. This does suggest a canalized, evolutionary basis for these capacities: in fact, the modular hypothesis *predicts* the existence of this kind of cultural variation.

However, the account also faces several challenges. Most obviously, while the account is good at recognizing the existence of reliably and cross-culturally developing human cognitive traits, it is not as clear that it is good at handling cross-cultural *differences*. This matters because while virtually all cultures contain mindreading, moral-normative thought, and tool use, the particular versions of these can still differ drastically across cultures (Downes &

Machery, 2013; Henrich, 2020; Machery, 2017; Machery & Stich, forthcoming; Schulz, 2021).

To account for this, the modular theory has developed the concept of an "evoked" culture: a cultural difference that stems from a common psychology being embedded in different environments, leading to different outcomes (Carruthers, 1998, 2013; Tooby & Cosmides, 1992). The trouble with this is that it is not always clear that this response does justice to the situation: the amount of variation observed is often better accounted for on the assumption that it is fundamental rather than merely the result of a universal psychological mechanism that is merely employed in different environments (Schulz, 2021; Ward, 2022). Making this clearer is easier done in the context of the discussion of specific cognitive abilities, and I will therefore return to this point in later chapters. For now, though, it can be noted that the modular account needs to be *expanded* to explain human cognitive diversity. On its own, it is not yet clear that it is able to do so in an appropriate manner.[3]

Importantly, even if this problem could be solved, another problem remains. The modular account leaves open how we can *use* any of the innate expectations and processes we have stored about a particular topic. In general, this cannot be taken for granted as it can be quite difficult to achieve. For example, applying a specific set of psychological representations to the world requires time, concentration and attention (Coolidge & Wynn, 2009; Epstein, 1994; Greene, 2008; Lieberman, 2003; Lieder & Griffiths, 2020; Ramsey, 2014; Schulz, 2018b; Wynn & Coolidge, 2011). Precisely because

[3] There is, of course, also considerable *individual* cognitive variation (i.e., variation that cross-cuts cultural differences). However, this is less of a concern for the modular account because the latter does not need to (implausibly) assume or predict that development is 100% uniform. Just as with all other biological traits, small variations in the timing of developmental events, minor genetic differences, and different environmental circumstances can lead to slight differences in traits: sunflowers generally have yellow petals of about 67 millimeters length, but there can be much individual variation concerning this (Mirzabe et al., 2018). In the present context, though, the worry is that much of cognitive variation in humans correlates with cultural differences. It is this that is not well predicted by the modular, nativist account.

the reliance on mental representation in decision-making generally calls for the further processing of perceptual states as well as the process of *inferring* what to do based on these representations, it tends to require significant amounts of cognitive resources. It is far from clear that we always have sufficient such resources available to us.

Again, the details of this point are best made clear in the context of the discussion of specific cognitive traits and will also become important in the next chapter. However, what matters here is just noting that the modular account is again somewhat underspecified: it presumes that, by positing the existence of innate modules, we have explained all we need to explain about distinctively human cognition. This, though, is not the case: being innately *prepared* to face the environment is one thing; actually be able to *use* these preparations is another.

Making all of these points more precise and applying them to specific cases is one of the goals in the following chapters. For now, though, it is sufficient to note the following points. The modular account is good in noting the importance of innate psychological preparations for dealing with our environment. It is also good at predicting the existence of many cross-cultural human cognitive communalities. However, the account leaves some questions open as well. In particular, without further expansions and clarifications, it is not yet clear (a) how to explain human cognitive *diversity* and (b) how to explain how we manage to actually *use* our innate psychological endowment (see also Schulz, forthcoming).

3.2 Cultural Learning–Based Accounts

On the other end of the spectrum are accounts emphasizing the fact that humans are exceptional cultural learners; these accounts take this fact to be the source of our distinctive cognitive capacities. As noted in the previous chapter, there are several different models of cultural learning leading to slightly different cultural-learning based accounts of distinctively human cognition.

At its most extreme, these accounts see the innate, evolved cognitive capacities of humans as nearly all of the lower-level, domain-general variety (Brown, 2018; Heyes, 2012a, 2013, 2016, 2018, 2012b; Sterelny, 2012a, 2021). Our dependence on others has made us exceptionally tolerant of their presence in our vicinity, we are exceptionally good at paying close attention to what others are doing, we have a strong personal memory, and we are quite good at delaying rewards. In turn, this has led to us becoming very good cultural learners: from tolerating apprentices nearby (say) we can shift to some explicit teaching until we build up institutions like schools that are explicitly dedicated to cultural learning.

Taking a step closer toward the modular theories of the previous section, other versions of this kind of cultural learning–based account agree that the socioecological pressures hominins have faced led to the evolution of some distinctive human cognitive capacities; they just think that the main or only such capacity (apart, perhaps, from language, which is typically seen as closer to a sensory modality—Fodor, 1983) is one for "natural pedagogy" (see, e.g., Csibra & Gergely, 2011) or "shared intention" (Tomasello, 1999, 2021, 2022; Tomasello et al., 2005; Tomasello & Herrmann, 2010; see also Legare, 2019; Tennie & Over, 2012; Tennie et al., 2009; van Schaik & Burkart, 2011).[4] Importantly, though, these accounts see the major upshot of these evolved abilities to be the fact that humans are exceptionally good cultural learners.

[4] Tomasello—especially in some later works (see, e.g., Tomasello, 2021, 2022)—also suggests that at the heart of uniquely human cognition is a distinctive psychological architecture that includes a normative module that is able to alter how we respond to our environment: we can change our wants and thoughts in light of feedback from our social group. However, this emphasis on issues of cognitive architecture is less central here: on the one hand, there are concerns about the details of the proposal (Moll et al., 2021), and, on the other hand and most importantly, it is really the social nature of humans that is at the heart of Tomasello's proposal. What makes human cognition unique, according to him, is the fact that we respond to cultural information in a way that other organisms do not. As noted in the text, exactly why that is can be left open here.

In the present context, there are two key aspects of this exceptional capacity for cultural learning in humans—however, exactly, it is realized. First, this capacity is important as it means that humans do not have to rediscover the best ways of dealing with their environment from scratch every generation. Instead, they can pick up on what previous generations found out. In this way, humans can fine-tune their behavior beyond what there may be genetic selection for: the locally adaptive ways of acting may be too frequently changing to making genetic adaptation possible, but cultural learning can still make these ways of acting available (Boyd & Richerson, 2005; Henrich & McElreath, 2011).

Second and relatedly, this capacity for exceptional cultural learning is important as it enables humans to *build on* and *improve* knowledge acquired in prior generations. Humans are not forced to replay the mistakes of their ancestors: they can do better (Boyd et al., 2011; Chudek et al., 2013; Henrich, 2015; Henrich & McElreath, 2007). For this reason, the cultural learning–based accounts of distinctively human cognition place cultural learning at the source of specifically human thought: all the other aspects of distinctively human cognition are said to follow from the latter. Put differently, what the different cultural learning–based accounts have in common is the idea that there are only a small number of biologically derived cognitive differences between human and non-human animals (though they differ over exactly what these differences are). It is just that these accounts see these differences as allowing for cumulative cultural learning—which, in turn, makes the acquisition of extremely complex further traits possible that would be out of reach of individual learners.[5] Among these traits

[5] Sterelny (2021) suggests that ecologically scaffolded individual learning can provide many of the same benefits as cumulative cultural learning. These differences do not matter here, though—as noted in the previous chapter, reliably ecologically scaffolded individual learning is here seen as a form of cultural learning. For more on this niche constructionist form of cultural learning; see also Fogarty & Creanza (2017); Jablonka & Avital (2010); Odling-Smee et al. (2003); and Sterelny (2003, 2012a, 2018).

are the remaining ones inventoried in the previous chapter, including capacities for mindreading, moral cognition, reciprocal cooperation, and tool use (language being perhaps an exception—though see also Sterelny, 2012c).[6]

In short, at the heart of cultural learning–based accounts of human cognitive uniqueness—however, exactly, they are constituted—is the idea that the fulcrum on which the existence of distinctively human cognition rests is domain-general cultural learning: the latter is the source of all the other aspects of human cognitive uniqueness. This kind of account has several positive qualities.

First and most obviously, the account does justice to the extensive use of cultural learning in humans. As noted in the previous chapter, humans are extensive cultural learners, and this type of account correctly notes this. It does not need to be expanded to account for this fact; this fact is already at the center of the account.

Second, this type of account is also good in making sense of the amount of cultural variation there is in human cognition. As noted in the previous section, there is no question that there is a large amount of such variation and that making sense of the latter is a crucial part of any compelling account of distinctively human cognition. The cultural learning–based account can take the existence of this kind of variation in its stride: indeed, the account *predicts* this—a major strike in its favor (Hitchcock & Sober, 2004; Schulz, 2021).

However, there are also some negative aspects to the cultural learning–based accounts of distinctively human cognition. These fall into three categories.

[6] "Culture" has a wide extension here. While it can refer to large groupings of humans, such as "Westerners" or the Hadza (Henrich, 2020), it can also refer to smaller groups and subcultures, like the hippie community in 1970s Haight-Ashbury or contemporary *kawaii* adherents in Japan. The point is that the cultural learning-based accounts see distinctively human cognition as stemming from the fact that humans learn from others in their community; exactly what that community is can be left open.

First, there are the flipsides of the positive aspects just mentioned: while cultural learning–based accounts of distinctively human cognition do better than the modular accounts in terms of explaining cultural diversity in human cognition, they are less good at making sense of cultural universality in human cognition. On the face of it, if all distinctively human forms of cognition—including, perhaps, even cultural learning itself (Heyes, 2012a)—are culturally learned, then we would expect there to be massive cultural variation in all forms of human cognition. However, while we do find much such variation, we also find many cultural communalities: all human cultures seem to mindread, rely on some moral concepts, etc. (Brown, 1991; Carruthers, 2013; Tomasello, 2021). More generally, there is considerable developmental psychological evidence supporting the view that humans come equipped with a set of "core knowledge systems" (Spelke, 2022). While it is not impossible for cultural learning–based accounts to make sense of these facts (e.g., it may be that all human cultures face similar problems in a well-defined solution-space and thus converge on the same kinds of solutions), it is at least the case that this is not obviously expected on this account (Northcott & Piccinini, 2018). Much the same holds for the disadvantages associated with learning. In particular, as noted in the previous section, it is just not clear that (culturally) learning all distinctively human cognitive traits is either possible or sufficiently efficient to be adaptively favored over the evolution of innate cognitive modules (Carruthers, 2006).

A second set of worries for the cultural learning–based accounts centers on specific aspects of the latter accounts. In particular, it is not clear that these accounts can always truly *explain how* humans are able to be such strong cultural learners. So, as noted in the previous chapter, Tomasello's proposal that this rests on the fact that humans have a unique, innate ability for joint attention has been criticized as lacking a sufficiently strong empirical basis (Boesch, 2005) and for failing to be theoretically fully compelling (Carruthers, 2013; Papadopoulos, 2023). The proposal of the evolution of an innate capacity for natural pedagogy (Csibra & Gergely, 2011) has been criticized on similar grounds (Heyes, 2013, 2014, 2016). The ability of niche constructionist approaches to provide an account of truly cumulative cultural learning is also controversial (Brown, 2018), and the same goes for purely domain-general, low-level accounts like that of Heyes (2018), which have been questioned for being able to capture all aspects of distinctively human cognition (Jacob & Scott-Phillips, 2021). Many of these criticisms will be assessed in more detail in later chapters. What matters here is just that it needs to be noted that the extent to which the cultural learning–based accounts are successful in explaining human cognition—including, in particular, the very cognitive suite they are based on—is not obvious. It is just not obvious that the capacity for extensive cultural learning in humans can be taken for granted: it, too, needs an explanation. Because of this, the explanatory foundations of the account seem somewhat shaky.

Finally, the cultural learning–based accounts also struggle with the same point that the massively modular accounts struggled with. Independently of how we acquire the relevant psychological capacities, it is still not clear how we manage to *use* them. Whether culturally learned, learned with an innately scaffolded mechanism, or innately developing, the use of mental representation in action is cognitively costly. These costs cannot be taken for granted, and it needs to be explained how we can overcome them.

In this way, we arrive at the following overall assessment of the cultural learning–based accounts. These accounts are good in emphasizing the importance of cultural learning to the development of distinctively human cognition, and they are well positioned to make sense of the amount of cultural variation there is in human cognition. However, these accounts also still leave several questions open: (a) Why do all humans rely on some very similar mental processes? (b) How can we explain *all* aspects of distinctively human cognition—including cultural learning itself?

and (c) How do we manage to actually employ the cognitive traits we acquire through cultural learning?

3.3 Symbol Processing–Based Accounts

A third account of the evolution and nature of distinctively human cognition sits between the modular and the cultural learning–based accounts: it accepts the importance of learning, but also that of innate mental representational abilities for distinctively human cognition. Specifically, the account takes as its central element the particularly powerful systematic representational abilities of humans and derives the other aspects of distinctively human cognition from this.

As noted in the previous chapter, humans are distinctive not for having mental representations that guide their behavior, but for instead making their behavior dependent on particular abstract and complex such representations (for particularly well worked-out versions of such an account, see Penn et al., 2008; see also Clatterbuck, 2018; Deacon, 1997; Fodor, 1975, 1983; Tomasello, 2022; Whiten, 1995). As also noted in the previous chapter, not only can humans react to how they perceive the world or even to categorizations of their perceptions, but they also can relate different such categorizations to each other, meta-categorize them, and change them. In short, human representational thought is systematic.

The symbol processing account of distinctively human cognition sees this systematicity of human representational thought to be especially deep and wide: human mental representations can be related to each other in a systematic manner on a scale that is not true for non-human animals (Deacon, 1997; Penn et al., 2008). This matters—so the account goes on—as it allows humans to be more flexible agents who can react to a finer set of circumstances in a more efficient manner (Deacon, 1997; Penn et al., 2008; Schulz, 2018b; Tomasello, 2022). For example, humans can react not only to conspecifics looking at various spots in the environment. Instead, humans can react to whether "Alex likes Tyler and Tyler likes Alex," as opposed to "Alex likes Tyler and Tyler does not like Alex." In particular, they might use the latter representation, in combination with the representation that only people who mutually like each other should be invited to a meal, to infer that we should not invite both Alex and Tyler.

While this is interesting in and of itself, the symbol cognition–based account of distinctively human cognition goes further and uses these facts to ground other aspects of distinctively human cognition as well. So, according to the symbol cognition–based account, the distinctively human ability to communicate in a natural language is made possible by the fact that humans can recursively relate representations to each other—the hallmark of a language (Deacon, 1997; Fodor, 1975, 1990; Penn et al., 2008). The same goes for the distinctively human abilities for causal cognition and tool use: these are enabled by humans being able to group different events as "cause" and "effect," and then relate these event types to each other. Similar remarks go for property cognition, social cognition, and moral cognition (Penn et al., 2008; Wilson, 2020). Finally, the recursive nature of relational symbol systems makes it possible to take existing elements and combine them in new ways, thus opening the doors to much innovation and creativity.

As was true for the other accounts, the symbol processing account, too, undoubtedly has several positive aspects. In the first instance, the account is good as far as representational cognition is concerned. As noted in the previous chapter, it is plausible that the particularly abstract and complex representational abilities of humans are distinctive, and the symbol cognition–based account correctly emphasizes this. Second, the account has some theoretical plausibility, in that, if not through some form of symbol processing, it is not clear how else to understand many aspects of human cognition (Deacon, 1997; Penn et al., 2008). The systematicity and productivity of thought, while heavily debated, do call for some form of symbol-computational thought (Calvo & Symons, 2014). Especially if symbol processing is understood broadly to include some non-classical computation (Piccinini, 2010)—which is possible (Penn et al., 2008; Schulz, 2018b; but see also Barrett, 2008)—the symbol cognition–based account can accommodate this point well.

However, there are also several negative aspects of the account. First, as noted in the previous chapter (and also by many discussants of Penn et al., 2008), it is just not obvious exactly how uniquely extensive the systematicity of human thought is compared to non-human animals. Human thought is not fully systematic—we do not consider all the logical permutations of all the thoughts we have, for example—and some non-human animals also seem to be able to think in quite systematic ways (Barrett, 2008; Beck, 2018; Pepperberg, 2008; Tetzlaff & Carruthers, 2008). This makes the foundations of the symbol account less than fully solid; at least, more work is needed to fully shore it up here.

Second, the account does not make clear how and why humans are able to have such powerful symbol cognition abilities. As noted in the previous two sections, the use of mental representations to guide behavior is not costless. It is not clear how humans are able to pay these costs whereas other animals cannot. This is made even more puzzling due to the fact that both types of organisms can use *some* kinds of mental representations. What accounts for the enhanced representational abilities in humans? As matters stand, this question is not answered by the symbol processing account (a point admitted to by Penn et al., 2008; see also Barrett, 2008).

Third, it is not clear how the account can explain the low-level features inventoried in the previous chapter. As made clear there, just because these differences between human and non-human animals are merely ones of degree does not mean that they are not significant. However, the symbol processing account is very focused on higher-level cognition and does not speak to lower-level aspects of thought at all. This, though, is problematic: the fact that, structurally, the symbol processing account is forced to treat the explanation of distinctively human capacities such as fine motor control, patience, and freely directed attention as completely independent from the explanation of language, tool use, and reciprocal cooperation is not greatly compelling.

Finally, the symbol processing account leaves many details open even of the issues it discusses. So, it may be true that humans have extensive relational symbolic cognitive abilities but this does not explain the specific representational contents that humans rely on (see also Tetzlaff & Carruthers, 2008). That is, it is one thing to note that humans are good are systematic representational thought; it is another thing to explain the specific, representational cognitive modules that seem to have evolved in human history. Similarly, the symbol processing account, by itself, has nothing to say about the kinds of cognitive cultural differences there are in representational cognition. The account does note that cultural learning is important, but it does not make clear exactly what impact the latter has on distinctively human cognition.

All in all, therefore, while making a positive contribution to the discussion here, the symbol cognition–based account of distinctively human cognition also leaves several important questions open. Key among these questions are (a) why humans are able to rely on highly abstract and complex, structured relational symbol systems when other animals are not; (b) how distinctively human lower-level cognitive capacities can be incorporated into the explanation of distinctively human cognition; and (c) what explains which distinctive representations humans are born with and which they learn from their culture.

3.4 Upshot

Putting all of this together, this implies that none of the modular/nativist, cultural learning–based, or relational symbol-processing accounts by themselves can do justice to all the issues surrounding distinctively human cognition (Schulz, forthcoming). In particular, together, all of these accounts still leave open the following questions:

1. How can we explain both the cross-cultural similarities *and* differences in distinctively human cognition?
2. How can we explain that humans are able to frequently use the complex and abstract representations they rely on to interact with the world, given the costs of using these representations?
3. How can we explain *all* aspects of distinctively human cognition—including lower-level ones?

To make progress in answering these questions, the rest of the book shifts perspective. Instead of taking one particular cognitive ability—such as language, cultural learning, or relational symbol cognition—to be central, *and* instead of just positing the existence of highly specific modular cognitive structures, the account defended in the rest of this book emphasizes the *interaction* among all of these elements to generate distinctively human cognition. The book shows that the upshot of this interaction is a feedback effect that takes a number of small differences in degree between human and non-human animals and amplifies these into much larger (and partially even—as in the case of language, for example—in-kind) differences. Outlining this account is the goal of the next four chapters.[7]

[7] Other authors—including some of those mentioned in this chapter (see, e.g.,Barrett, 2008; Heyes, 2012b; Tomasello, 2021)—have hinted at such a picture, but none has spelled it out in the way done here.

4

Thinking Tools

In this chapter, I sketch the outlines of the interactionist processes underlying the evolution and nature of distinctively human cognition that form the backbone of the theory developed in this book. At the core of this interactionist theory is the idea that underlying human cognitive uniqueness is a complex, interlocking interplay between three different elements: (1) a set of evolved representational expectations, (2) a pronounced disposition for cultural learning, and (3) complex and variegated technology. All three are further underwritten by the unique set of lower-level cognitive abilities sketched in Chapter 2, which are in turn influenced by the feedback loop created by elements (1) through (3).

In this way, the present account makes clear how human cognition can end up differing quite markedly from non-human cognition even though its individual elements might initially and intrinsically only differ in relatively small degrees (see also Laland & Seed, 2021). Indeed, as I make clearer below, this account allows us to answer the questions left open by the previous accounts: by combining and expanding on the best of the previous perspectives, the present account can do what these perspectives cannot do on their own. In particular, the account shows that human cognition is characterized by a lot of cultural learning, but also by innate, evolved psychological modules and expectations; both of these are further externally enhanced and underwritten by technology.

It's Only Human. Armin W. Schulz, Oxford University Press. © Oxford University Press 2025.
DOI: 10.1093/9780197800188.003.0004

The goal of this chapter is to present the core workings of this account. To really see how it is able to make sense of distinctively human cognition—and how it differs from the existing treatments on this topic—it is best to apply the account to some of the key items of distinctively human cognition inventoried in Chapter 2. This will be the goal of the following three chapters. Still, it is useful to start with a general overview of the account; this is what this chapter provides. Doing so makes clearer how the different applications detailed in the next few chapters hang together and provides an explicit contrast to the accounts sketched in the previous chapter, to bring out the novelty here more explicitly.

The chapter is structured as follows. Section 4.2 sketches the building blocks of the account, drawing on the work done in Chapters 2 and 3. Section 4.3 is the heart of the chapter: it presents the interplay between these building blocks. Section 4.4 shows how this account can answer the questions from the previous chapter (and do so better than the alternative accounts). Section 4.5 summarizes the discussion.

4.1 The Building Blocks

The present account starts off where the modular, nativist account does, too. Hominins evolved in a unique set of environments that rewarded them to be psychologically well equipped to handle a number of unique challenges. These challenges center around the fact that humans are obligate social organisms without special biophysical adaptations to their environments and whose degrees of relatedness and convergence of biological interest were less than unity. In this sort of environment, hominins profited from tracking important social relationships, remembering who did what to whom, and being able to delay gratification, often for significant periods of time. More specifically, this sort of environment led to several sets of important psychological adaptations.

4.1.1 Evolved Representational Expectations

The present account further follows the arguments from the modular, nativist views that, in the kinds of environments that hominins evolved in, it was adaptive to be prepared with sets of representations—or at least ways of acquiring specific such sets of representations—that make interaction with these environments possible and/or efficient. Since these arguments were sketched in the previous chapter, there is no need to rehearse them here in detail. What matters is just that, theoretically, the evolution of innate representations has a lot to recommend it: it streamlines or even makes possible the acquisition of important cognitive inferential abilities. It is also empirically plausible: as noted in the previous chapter, there is significant data supporting this view (Schulz, forthcoming; Spelke, 2022). Because of this, evolved representational expectations should be seen as an important part of our cognitive building blocks.

However, it is now important to clarify the notion of modularity—or, rather, lack thereof—at stake here. The present account differs from the modular views of the mind sketched in the previous chapter by not focusing on *modularity*, but on the evolutionary process itself. This is important to note because it means that the present account has some unique features that enable it to avoid some of the problems affecting the traditional modular views of the mind.

As is widely noted, on the most stringent view of cognitive modularity—that of Fodor (1983)—there are only very few cognitive modules: sensory modalities, language, and perhaps a handful of motor modules. On this account of modularity, modules cannot access representations stored in other parts of the cognitive system (they are informationally encapsulated). While this view has several benefits—it is relatively precise and is able to account for some important features of human cognition (e.g., the Stroop effect)—it cannot be seen as a fully general account of the structure of the human mind. In particular, it is not able to account for the evolution of most of the domain-specific behavioral expectations or learning devices noted in Chapter 2. After all, it is just not clear how any domain-specific set of representations could affect decision-making without relying on other such representations. For example, it is not clear how a mindreading module could be informationally encapsulated: after all, another person could be thinking about all sorts of things.

For this reason, the modular accounts sketched in the previous chapter tend to rely on a weaker form of modularity, according to which modular cognitive processes are (a) typically domain specific and (b) not *usually* reliant on representations in other parts of the cognitive system, or not to a significant extent (Carey, 2011; Carey & Spelke, 1996; Carruthers, 2006). That is, the representations in question typically are just seen to be *the leading drivers* in getting us engaged in appropriate forms of behavior in certain specific situations. While they may require some further added representations from other parts of the cognitive system to be fully functional, they are still the most significant elements of the relevant causal processes.

Again, while this weaker account of "massive" modularity has some benefits—for example, it can be spelled out in a very general way that accommodates most of the relevant cognitive phenomena (see, e.g., Carruthers, 2006)—it also faces some challenges. The main one among these is that the resultant notion of "modules" can appear to be so permissive as to be virtually content-less. If "modules" need not be *fully* domain- or task-specific and may depend on representations in other parts of cognitive systems, it is not obvious to what extent seeing them as substantive, separate modules in the first place is compelling (Dupre, 2001; Samuels, 2006; Sterelny, 2003). Put differently, the weaker version of modularity is in danger of collapsing into a purely semantic point about how to label different cognitive traits or abilities. More generally, accounts of cognitive modularity face the challenge of having to navigate between the Scylla of being too stringent to account for the complexity of human cognition and the Charybdis of being too liberal to be theoretically and empirically meaningful. To what extent they are able to avoid these twin dangers is, at the very least, still an open question.

The present account of distinctively human cognition, though, sidesteps this entire issue. In particular, the account to be defended in this book focuses just on the fact that it is plausible that there are different sets of mental representations in human minds that (a) have an independent evolutionary history and (b) help us deal with specific aspects of our environment. So, as far as point (a) is concerned, the account here notes that the evolutionary pressures on the representations underlying distinctively human moral cognition, for example, are different (though related) from those underlying distinctively human property, which are different from those underlying mindreading cognition, and so on. This, though, does not mean that making decisions in these different domains may not recruit the very same representations or that they are informationally encapsulated. Rather, it is just that distinct selection pressures led to the accumulation of the total number of representations needed to provide a sufficient foundation for the kind of decision tasks humans faced. Put differently (and as noted in point (b) above), these different sets of representations *combine* to allow us to make decisions across the situations we face. Figure 4.1 makes this clearer.

In this way, there is no need to commit to these sets of representations being "modules" in a metaphysically deep way. At most, these sets of representations could be seen to be "phylogenetic modules" in the sense of having independent evolutionary histories (Felsenstein, 1985, 2004; Schulz, 2013). This, though, does not entail that they are cognitively separable traits in an important sense.[1] This relates to the point noted in Chapter 2 that

[1] This also makes them different from what some have called "Darwinian modules" (Machery, 2007).

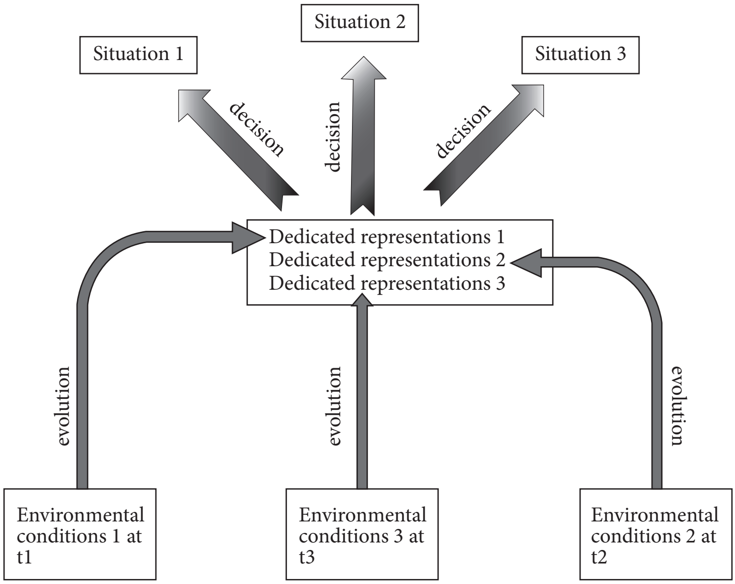

Figure 4.1 The separate evolution of different sets of representations generates a representational pool with which to guide inferences across a variety of decision situations.

determining what makes for a trait in biology, more generally, is not straightforward and that the present account stays neutral on this. The avoidance of a commitment to modules is an extension of this neutrality.

Furthermore, it is important to note that these independently evolved sets of representations are just that: sets of mental representations. They are not inference machines that somehow directly yield decisions by themselves. It is true that representational, high-level decision-making is, broadly speaking, *inferential* in nature: instead of simply mapping particular perceptual states to particular behavioral outcomes, organisms *infer* what to do, either by making an inference from their perceptual states to the state of the world, or by inferring what they ought to, given how

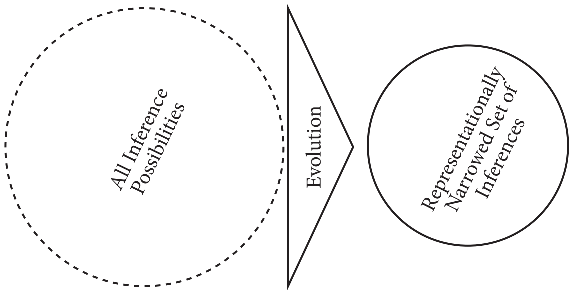

Figure 4.2 Open-ended set to smaller finite set.

they perceive the world, or both (Carruthers, 2006; Millikan, 2002; Schulz, 2018b; Sterelny, 2003).[2] However, this does not mean that the inference can be *automated*. What inference to draw from a relevant set of representations is highly situationally dependent. Indeed, this is precisely why it is not plausible to see these modules as being Fodorian in nature.

Instead, the proposal here is just that these sets of representations evolved to *guide* the relevant inferences by providing us with a set of psychological expectations about the world. Instead of having to consider an open-ended set of possible conclusions about the state of the world and our best reaction to it, we can shortcut this very large set considerably and reduce it to a more manageable set. Given the complexity of the situation, though, it is implausible to see this set as generally being reducible to one or a few options only—many options still remain on the table. It is just that this option set is now smaller. From an effectively infinite set, it has been reduced to a finite one. Figure 4.2 illustrates this.

[2] Note that it is not entirely clear to what extent this inference is *computational*—and what the relevant sense of "computation" here is—rather than *inferential* in a noncomputational sense (Piccinini & Scarantino, 2010). However, since settling this is not important for present purposes, I shall sidestep further discussion of this issue and just refer to "representational inferences" (or its cognates) in what follows.

Finally, it is important to note that the present account agrees with the symbol processing account that it is plausible that humans have evolved particularly extensive abilities for embedding their representations in a systematic framework. However, it differs from the symbol processing account in denying that, at least initially, this was a discrete, major difference. As noted earlier, the extent to which non-human animals are able to rely on structured representations is not yet clear (Beck, 2018; Pepperberg, 2008; Schulz, 2018b; Wasserman, 2008). For this reason, the present account does not place the ability for extensive, systematic symbolic cognition at its center.[3]

4.1.2 Foundations for the Evolution of Cultural Learning, Technology, and Language

The next key element of the present account is shared with accounts emphasizing cultural learning. Humans needed to adjust their behavior to the locally prevailing conditions, and thus, cultural learning became highly adaptive for them (Boyd & Richerson, 2005; Fogarty & Creanza, 2017; Henrich, 2015; Henrich & McElreath, 2007, 2011). In particular (as noted in the previous chapter), the fact that cultural learning is cumulative, but yet not dependent on the origination and then spread of suitable genetic variants, makes it adaptive for enabling organisms to determine biologically advantageous ways to think and act in environments with intermediate amounts of variation.

This makes it plausible that a number of lower-level and higher-level traits have evolved to make such cultural learning easier. On the low-level side, these include increased social toleration; a

[3] The present account thus shares some key ideas with that of Ismael (2011), such as the emphasis on centralized informational integration and inference. However, it also differs from the latter, in that—as will be made clearer in the rest of this chapter—the latter are seen to be embedded in a rich cultural and technological feedback loop.

motivation for imitation; strong attention and concentration control, especially when it comes to tracking others' gaze; good facial memory; and a strong disposition to cooperate reciprocally.[4] All of these traits made it easier for humans to learn from others and cooperate with them to thrive in their environments (Heyes, 2018; Sterelny, 2003, 2012a, 2021). Furthermore, as noted in the previous chapter, it is *possible* that humans also evolved some higher-level abilities to aid their cultural learning: a disposition for natural pedagogy (Csibra & Gergely, 2011) or an ability for joint attention (Tomasello, 1999, 2021, 2022). However, as also noted in the previous chapter, since this is still controversial, the present account does not rely on the latter being true—indeed, it is consistent with either outcome here.

Similar points can be made concerning the related cognitive suite of tool use and manufacture. The fact that humans lack major biophysical adaptations to deal with their environment and are obligate social learners makes it adaptive for them to build tools that enhance their biophysical abilities (Boyd & Richerson, 2005; Henrich, 2015; Landy et al., 2014; Muthukrishna & Henrich, 2016; Osiurak & Reynaud, forthcoming; Tennie & Over, 2012; van Schaik & Pradhan, 2003). In turn, this also makes it adaptive for humans to have a number of lower-level cognitive traits, such as fine motor control, attentional and concentrational control, patience, and an enhanced motivation to seek out novelty and play. On top of this, it is possible—but also controversial—that humans have evolved a specific set of high-level expectations concerning technology, including ones concerning causal learning (Papineau, 2003; Tennie & Over, 2012; Vaesen, 2012). Again, this will thus not be central in what follows.

The final building block to be mentioned here concerns language. The fact that humans are obligate social organisms and

[4] For some implications of the human ability to track caregivers, see, e.g., Osmanoglu and Schulz (2019).

cooperators makes it adaptive for them to communicate in a manner that is precise and directed: the ability to express an open-ended set of propositions allows for very specific and fine-grained learning, communication, and cooperation that it is hard to achieve purely through observational learning. In turn, this fact is widely thought to have led to the selection of dedicated mental structures for speaking in a natural language (Carruthers, 2006; Pinker & Bloom, 1990; Sterelny, 2012c). The present account accepts this consensus position.

4.1.3 The Missing Element: Interaction

However, in all of the above, it is important to note that, at their root, these cognitive abilities are not orders of magnitude more advanced than what is found in non-human animals. As noted in Chapter 2, by themselves, the human capacities for increased social toleration, imitation, attention and concentration control, facial memory, reciprocal cooperation, fine motor control, etc. are merely different in degree from what is found in non-human animals. In this way, the present account already differs quite strongly from the modular accounts sketched in the previous chapter: human cognitive uniqueness cannot be fully account for *just* by noting the evolution of some distinctive cognitive capacities. While such an evolution is plausible, it is not sufficient to account for the depth and breadth of human cognitive uniqueness noted in Chapter 2. Something else needs to be added to the evolution of the capacities mentioned in the previous subsections to provide a fully compelling account of the evolution of distinctively human cognition—the preceding just provides the basic building blocks of such an account.

In particular, the preceding discussion still fails to answer any of the questions from the previous chapter:

1. How can we explain both the cultural similarities and cultural differences in cognitive traits across people?

   So far, the building blocks provide the foundations for some human universals in cognition—widely shared representational expectations or low-level abilities. The above also gestures at the possibility of cultural learning but does not spell this out and thus leaves unclear what is culturally specific about human cognition and why.
2. How can we use our evolved sets of representations to make decisions?

   As noted earlier, they might narrow the set of inferences, but they do not reduce it to zero. How do we then actually go on and use these representations? Also noted in the discussion of the symbol processing account of distinctively human cognition, this use is not costless and needs to be explained.
3. Why is it that humans are able to build tools of such complexity, rely on concepts with such complexity, and be cultural learners of such complexity when other organisms do not have these abilities? What explains the divergence here exactly?

   Put differently, how can we explain that humans have extensive abilities in these regards. How do the different elements of the account—low-level abilities, evolved representations, etc.—come together to account for the particular human cognitive traits inventoried in Chapter 2. As also noted in the discussion of the symbol processing–based account of distinctively human cognition, these questions are still open but deserve to be answered.

On top of this, there is one further important question that the above elements, as they stand, do not explain as yet either:

4. Why do humans have extensive abilities in *all* of these dimensions?

> Is it coincidental that humans are strong and creative cultural learners *and* tool users *and* language users (etc.), or is there something unifying these different abilities?

The main contention of the account to be developed here—and in the book as a whole—is that, to answer these questions, the fundamental building blocks sketched above need to be placed in the right kinds of *relationships*. While technology, evolved representations, cultural learning, creativity, etc. have been investigated individually, and while some connections between these have been pointed out earlier (see, e.g., Boyd & Richerson, 2005; Clark, 1997; Fabry, 2017; Fogarty & Creanza, 2017; Heyes, 2012b, 2018; Klein & Edgar, 2002; Legare, 2019; Muthukrishna & Henrich, 2016; Osiurak & Reynaud, forthcoming; Reindl et al., 2018; Schulz, 2018b; Sterelny, 2012a; Tennie & Over, 2012; Tomasello, 1999; van Schaik & Pradhan, 2003), there is a set of mutually reinforcing interactions among technology, evolved representations, and cultural learning that has not yet been clearly documented or which has only be hinted at in previous treatments.

As I show in what follows, though, laying out this set of mutually reinforcing interactions is key for getting at cogent answers to each of the above four questions. The key idea is that the different elements of distinctively human cognition mutually amplify the existing selection pressures on these elements, creating a positive feedback loop that ultimately leads to the kind of distinctively human cognition that can be observed now (and which has been inventoried in Chapter 2). In this way, the present account goes beyond the existing literature by combining the best of the previous perspectives in a novel package: human cognition is characterized by the *interaction* of cultural learning with innate, evolved representations and technology. To bring this out, I begin by considering some of the major dynamic relations just between our set of evolved representations and technology. Once that is done, I bring cultural learning into the picture.

4.2 Technologically Enhanced Representational Decision-Making

The availability of technology makes it much easier to rely on complex and abstract mental representations (whether these are innately, culturally, or individually acquired).[5] Two different types of this sort of technology can be distinguished.

First, technology might be available that allows an organism to store, for the long-term, some or all of the information it needs to rely on to use its evolved or learned store of mental representations to make decisions about what to do in a given situation. This will be called technology of *type S* (for storage) in what follows.

To see the benefits of technology of type S, assume an organism makes foraging decisions by relying on a decision rule involving a number of abstract and complex mental representations, such as "compare the food recovery rate at the current locale with the average rate for the surrounding area." Assume also that the organism had technology available that allowed it to store this average rate for long-term usage once it had it estimated. Then, instead of needing to estimate the latter rate over and over again, the organism just needs to do it once and can then refer back to it whenever needed. This can significantly lower the *average* costs of making the relevant representational inferences. After the initial setup, these costs can drop to near zero. Importantly, given the complexity of some representational inferences, reducing the need for them can (though of course need not) come with major improvements in cognitive

[5] It is sometimes suggested that certain forms of technology allow organisms to use the environment directly to make a decision about what to do and thus make it *less* adaptively valuable to rely on mental representations (Beer, 1990; Brooks, 1991; Silberstein & Chemero, 2012; Smith & Thelen, 1994). A lot can be said about this dynamical systems perspective on these issues; however, for now, it just needs to be noted that it does not exhaust the relationship between technology and mental representations: as the rest of this section makes clearer, there are also adaptively positive connections between the latter two. See also Schulz (2018b, chap. 7), which focuses especially on the use of other minds as enhancers of representational decision-making.

efficiency. Of course, this assumes that the organism in fact needs to make the same type of representational inference multiple times. However, while not universally plausible, this is at least sometimes plausible: streamlining repeated decision-making is precisely one of the drivers of the evolution of representational decision-making in the first place (Millikan, 2002; Schulz, 2018b; Whiten, 1995).[6]

The second type of technology that can significantly enhance an organism's ability to rely on complex and abstract mental representations is one that itself makes some or all of the needed representational inferences for the organism. This will be called technology of *type I* (for inference) in what follows.

To see the benefits of this kind of technology, note that if an organism can access technology—some form of calculator, say—that is able to estimate the average food recovery rate of the area, then the organism *never* has to make the relevant inferences. In fact, the organism does not even need *to be able to* make the relevant inferences (or at least not within ecologically realistic timescales) and could still rely on complex and abstract mental representations to interact with its environment. In this way, the availability of technology that makes some of the relevant representational inferences for the organism can also make the use of abstract and complex mental representations much faster and less concentration- and attention-hungry.

Technology of types S and I are especially noteworthy in this context as it is plausible that the complexity—and thus the costliness—of a representational inference is, *ceteris paribus*, related to the abstractness and complexity of the concepts involved (Ramsey, 2014; Schulz, 2018b). The closer a concept is to being perceptual, the easier it tends to be to relate it to the state of the world. Highly abstract concepts do not have a straightforward empirical signature and thus tend to require more work to connect to the environment of the organism (Fodor, 1983, 1990; Margolis &

[6] See also Note 14.

Laurence, 2015; Prinz, 2002).[7] Similarly, more complex concepts—that is, concepts with more parts—generally require more in the way of tracking the different parts in representational inferences than do concepts with fewer parts (Fodor, 1983, 1990; Margolis & Laurence, 2015; Prinz, 2002). Two further points are worthwhile to note about technology of types S and I.[8]

First, some of the most famous examples of human technology are of types S and I. So, as far as technology S is concerned, humans have long found ways of creating symbols that can be stored, transported, and manipulated (Kelly, 2015; Muthukrishna & Henrich, 2016). Indeed, technology of this type is quite old: on a conservative estimate, symbolic cave art, figurines, and musical instruments appear in the material record about 40,000–50,000 years ago (Bednarik, 2008; Kelly, 2015; Klein & Edgar, 2002; Lawson, 2012; Mellars, 1989; Morley, 2013; Pike et al., 2012; Renfrew & Scarre, 1998; Shea, 2017).[9] Technology of type I—that is, broadly inferential technology—is newer, but even that goes back several thousand years. For example, the "Senkereh Tablet" is a Babylonian calculating device about 5,000 years old (Sugden, 1981), and the first sundials date from about 3,000 years ago (King,

[7] This does not mean that this tradeoff is linear or one-sided. Some kinds of reasoning with abstract concepts can be quite easy (If I know that I have a hammer, and if I know that this a bone, then I can infer that I can break the bone with the hammer to obtain marrow). Still, the key point here is that the cognitive labor involved in the reliance on complex and abstract mental representations is *generally* an increasing function of the abstractness and number of abstract concepts employed: although determining how to use a [HAMMER], while non-trivial, might be relatively easily done, it remains true that determining whether a [PROPERTY] [ARRANGEMENT] is [JUST] is harder.

[8] The distinction between technologies S and I need not be sharp. For example, an organism might use tallying sticks first as a calculating device and then as a mobile storage device for the results of such a calculation.

[9] This also fits well to Kelly's (2015) argument that external devices of various kinds (such as the building complexes in the Chaco Canyon of New Mexico) have been constructed as memory aids. While Kelly focuses on using technology as an aid for remembering, rather than as information storage itself, this difference is not so important for present purposes. What matters here is that technology of this type enables humans to reduce the costs of making various kinds of representational inferences and decisions.

1955); needless to say, recent human history has seen an explosion of such computational tools.

Second and crucially, it is plausible that precisely the existence of this kind of technology is a key factor underlying the human ability to rely on many highly complex and abstract concepts in many of their representational decisions.[10] Because they can outsource key aspects of the associated cognitive labor, they can rely extensively on concepts like [CAUSE], [JUST], [KNOWS], or [NUMBER] (Landy et al., 2014; Muthukrishna & Henrich, 2016; Stout & Chaminade, 2012).

In some cases, the use of these concepts themselves is underwritten by the use of technology. For an obvious example, much of science (in a broad sense) is and has been conducted with the aid of written symbols—including (especially) mathematical ones—and computational devices (Fabry, 2017; Hutto & Myin, 2012; Menary, 2007; Russo, 2023). However, many other examples can be cited as well, from making representationally difficult investment decisions (Benbasat & Dexter, 1982; Todd & Benbasat, 1992) (including in Babylonian times—Sugden, 1981) to determining where a ship is located (Pacey, 1992). Interacting with the world based on concepts such as "the procession of the perihelion of Mercury," "maximize revenue streams in risky environments," or "is at longitude 156.3319°" is only made possible through the use of technology like mathematical written calculations, the abacus, or the marine chronometer (Benbasat & Dexter, 1982; Henrich, 2015, p. 230; Pacey, 1992; Schliesser & Smith, 1996; Smith, 2005).

In other cases, it is the fact that technology exists that allows *other* decisions to be made quickly and efficiently which ensures that the above abstract and complex concepts can be relied on.

[10] Note that the claim is just that cognitive technology underwrites the reliance on *many highly complex and abstract concepts*—not that *all* (representational) cognition depends on external aids, as it is often argued in the literature on embedded or situated cognition (Clark, 2013; Haugeland, 1999; Hutto & Myin, 2012; Menary, 2007).

Humans can spend time assessing what is *just*, for example, because they have tools available that allow them to track quickly and easily exactly who *owns* what: written records in a natural language. It is a lot harder to determine whether an inherited property arrangement is *just* if it is not clear what the property arrangement is or how it came about—not to mention if our cognitive and other resources are already extremely taxed by the needs to organize the basics of survival (Gilbert, 2018; Landy et al., 2014; Mithen, 1990, 1999; Wynn & Coolidge, 2011). Technology of type S or I can help make this kind of determination.

Note that the point here is not that *all* uses of complex or abstract concepts in humans must directly or indirectly rely on this kind of technology or that this is the only thing this use depends on. In particular, as noted earlier, this use also depends on various innate capacities: both low-level ones like flexibility in directing attention and an increased capacity for working memory (Coolidge & Wynn, 2009; Mithen, 1999) and high-level ones like language and innate sets of representations such as [KNOWS], [BELIEVES], [CAUSES], [IS A NUMBER] (Carey & Spelke, 1996; Carruthers, 2006; Cosmides & Tooby, 1992; de Hevia et al., 2014; Margolis & Laurence, 2015) but see also Cowie, 2003; Heyes, 2018; Sterelny, 2003). Indeed, it is also plausible that humans can use *other humans* as aids in streamlining their representational decision-making (Schulz, 2018b, chap. 7).

Rather, the point here is that the reliance on technology of types S and I generally is one of the major components that makes it possible to often rely on many especially complex and abstract mental representations. Whatever else is needed, without technology of type S or I it would generally be too time-consuming or take too much concentration and attention to use *many different*, *highly abstract*, or *complex* concepts like [KNOWS] or [IS JUST] for *many different* decisions—which is precisely what it characteristic of distinctively human cognition. In other words, this kind of technology allows us to *use* the narrowed set of representations that have

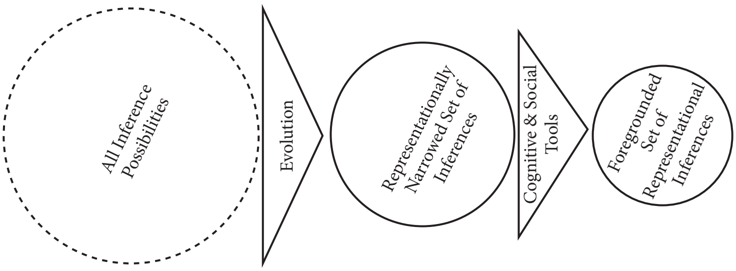

Figure 4.3 Sociotechnological narrowing of evolved mindreading expectations.

evolved to enable us to deal with certain specific environmental circumstances. Figure 4.3 makes this clearer.[11]

At this point, it is useful to make explicit that there is also an important *abstract* version of technologies S and I: social institutions. While the focus of the present account is often on the design and use of physical tools—artifacts—social tools are also an important part of the human technological sphere. (Indeed, given that artifacts are embedded in social practices, the difference between artifacts and social institutions is not clear to begin with: see also Preston, 1998, 2013; Winner, 1980). In general, social institutions are rules that help structure many human interactions and which set out the kinds of behaviors that, in a given type of situation, members of the society are expected to—and expect others to—engage in (North, 1990; Parsons, 1951). Famous examples of these institutions are governmental systems such as representative democracy, market-based systems of exchange, marriage protocols (such as polyandry), and religious practices (e.g., collective prayer).

There are many reasons why social institutions arise and spread in a given human population (Bigelow, 1998; Elster, 1979; Kincaid, 1990; Pettit, 1996; Rosenberg, 2012; Schulz, 2022): for example,

[11] By narrowing down the inferential space left open by our innate representational expectations, the use of technology of types S and I can thus help solve the "challenge of (cognitive) plasticity" in human thought: see also Machery (2018).

they might increase social coordination (Andrews et al., 2024; Westra & Andrews, 2022) or enhance existing human psychological tendencies in various ways (Henrich, 2015, 2020). Importantly, though, some such institutions also function as technologies of types S and I: they can be and are used to store important information and to help make (representational) inferences easier. For example, prices in well-functioning markets can be used to easily infer the demand for a particular good, social role expectations can be used to help guide inferences about other's motivations, and religious rituals like the Lord's Prayer can function as mnemonic devices that facilitate the storage of important information (much like sayings such as "every good boy does fine" for music). (The remaining chapters of the book will present some of these examples as well as others in more detail, too.)

Now, it is furthermore important to note that not all social institutions can function in these ways; only some can. Markets that are distorted do not signal supply and demand well, social roles with a large and non-unified set of behavioral and motivational expectations (such as being a "celebrity") might not make mindreading easier, and some religious rituals may not work well as storage of information. Only some institutions can play these roles—a point to which I return momentarily. Indeed, in general, the availability of technologies S and I of any type cannot be taken for granted.

For example, while many features of the organism's environment can be used to *temporarily* store information in some form (Clark, 1997, 2008; Griffiths & Stotz, 2000; Rowlands, 2010), the long-term storage or computational enrichment of information is unlikely to be easily obtainable. So, when it comes to physical technology of type I, it is just not generally the case that features of the environment themselves perform appropriate representational inferences.[12] Only deliberately generated tools are likely to be able

[12] Of course, some features might perform some kinds of calculations (Reed, 1996). The point is just that this will not be so for many of the representational inferences organisms might need to make.

to do this. Furthermore, building such tools is not straightforward. To be usable as an inferential aid, a tool needs to actually be able to perform the needed inferences, and it needs to be able to do so sufficiently efficiently. This implies that finding or building tools that can play this role is unlikely to be easy (Muthukrishna & Henrich, 2016; Osiurak & Reynaud, forthcoming; Reindl et al., 2018; Tennie et al., 2009; van Schaik & Pradhan, 2003).

Much the same holds for physical technology of type S. To be useful in enhancing the frequent use of many complex and abstract mental representations, such technology needs to be able to store information in a way that buffers it from environmental contingencies—both inorganic and organic. It also needs to be able to store this information in a way that makes it easily and reliably accessible. This rules out many features of the environment, as these two demands pull in opposite directions: the need to buffer information from external influences favors storage that is not easily and reliably accessible, and the need for easily and reliably accessible information favors storage that is open to external influences.

To overcome this, organisms are likely to need to manipulate the environment in some form. One of the major ways to do so is by devising or finding *mobile* sources of long-term information storage (Kelly, 2015; Shea, 2017). With mobile long-term storage, external influences can be minimized, and the information remains easily accessible. However, devising suitable mobile information storages is not trivial. To be mobile, such storage needs to be light, but, to remain stable, sufficiently robust. Apart from some classic human examples, few such mobile information stores are known (Bednarik, 2008; Fragaszy et al., 2013; Morley, 2013; Renfrew & Scarre, 1998; Shea, 2017; Shumaker et al., 2011). Another way of solving this problem is by relying on technology that stores the information in places that are *only* accessible to the relevant organisms but which *are* easily accessible to the latter. Such places are rare and need to be constructed. Generally, if one organism can access a given locale, so can at least several other organisms—not to mention the weather and so forth—and, among the places that are only accessible to a given organism, few are *easily* accessible to this organism. Apart from human examples (such as paintings in strategically located caves: Miyagawa et al., 2018; see also Kelly, 2015), it is therefore unsurprising that few examples of such forms of long-term storage have been found (Shea, 2017).[13]

In sum, technology that strongly enhances the frequent reliance on many complex and abstract mental representations—including those that are innate, those that are acquired by cultural learning, and those that are individually derived—is rare and can only be constructed with difficulty. So, given that (as just noted) humans have been able to manufacture and use technology of type S or I, how did they manage to do this? The answer is, unsurprisingly, cultural learning.

4.3 A Positive Feedback Loop

Cultural learning interacts with the two aspects of human cognitive uniqueness just mentioned (evolved sets of representations and technology), and it does so in both directions. Indeed, this is the heart of the present account and the core of the solution to the questions posed at the end of Section 4.1. What is missing from just looking at the evolved components of distinctively human cognition in isolation is their interrelations; indeed, their evolution cannot be properly understood in isolation from each other. A proper understanding of their evolution—and thus, of distinctively human cognition more generally—requires paying attention to the dynamic relations between these elements. To see this, consider the three pairs of relationships (cultural learning–technology,

[13] It is possible that humans made paintings in many different places but that these paintings only survived in caves. In that case, these paintings would not be an example of technology S. However, the most common view of cave paintings is that they were, indeed, deliberately created in caves. See also Kelly (2015).

technology–frequent use of complex and abstract mental representations in decision-making, and cultural learning–frequent use of complex and abstract mental representations in decision-making) separately at first and then combine them to get an overview of their overall interplay.

4.3.1 The Cultural Learning–Technology Nexus

First and most straightforwardly, as noted earlier, cultural learning allows for the piecemeal, cumulative manufacture and refinement of technology of type S and I (Boyd & Richerson, 2005; Henrich, 2015; Heyes, 2018; Sterelny, 2012a; Tennie et al., 2009; Tennie & Over, 2012; Tomasello, 1999; van Schaik & Burkart, 2011; van Schaik & Pradhan, 2003). In this way, the answer to the question of how humans managed to build this kind of technology despite the difficulties that come with designing, manufacturing, and using it becomes easy to see. No individual human needs to be able to fully grasp the details of the representational inferences they are seeking to outsource. Rather, the appropriate kind of long-term storage and the external representational inference machines can be built slowly and over time. Again, it is important to emphasize that this is not to say that technology does not also get enhanced by other things—including, perhaps, by humans' innate technological competence and their intrinsic curiosity and motivation for innovation (Kidd & Hayden, 2015; Osiurak & Reynaud, forthcoming; Reindl et al., 2018; Tennie et al., 2009). The point is just that cultural learning is a key element in the development of sophisticated technology. As noted above, this point is quite widely noted (Sterelny, 2012a, 2016, 2017, 2018; Tennie & Over, 2012; Tomasello, 1999).

Importantly, this also holds for social institutions: while, as noted earlier, social institutions arise and spread for many different reasons—and are relatively rarely explicitly "designed"—the step-by-step refinement of institutions so that they can better play the role of technology of type S or I is common. Efficient democracies, useful accounting conventions, well-functioning markets, and memorable religious rituals do not need to be found in one instant, but can be refined over generations (Henrich, 2015, 2020; Nichols, 2004).

However, there is also a set of reverse influences from complex technology to more sophisticated cultural learning. This set of influences is less widely recognized, but it is very important still. As organisms become able to rely on tools that help them make some or all of their representational inferences, their ability to culturally learn from others is expanded, too.

On the one hand, they can now learn more *efficiently* from others. Given the fact that learners have technology (including social institutional technology) available that allows them to make some of the needed representational inferences, more information can be transmitted to them in a given time period and with a given level of effort. Models or teachers can just provide outlines of the needed information and let the learners fill in the details as needed on their own. In this way, the *effective* (though not the actual) bandwidth of the transmission channel is being increased (see also Sterelny, 2012a). In turn, such increases in the effective bandwidth of the cultural learning channel mean that more information, and more complex information, can be obtained from others.

On the other hand, cognitive and social technology can increase the set of possible sources of cultural learning. Scrolls, books, and other instances of technology of type S can be passed on to others, preserved over time, and carried across mountains. Cave paintings can be found by future generations or strangers traveling through a given area. Technology of type I may provide organisms with ways of inferring what those only distantly related to them (in temporal, spatial, social, or epistemic position) are likely to think: ruins of a previous building can be used as the basis of a representational inference about where and how early generations thought it would be good to live. Clothing worn by travelers can be used as the basis for of a representational inference about how those in other places live and what environmental conditions they face. However achieved, an increase in the number of possible sources of cultural information is important as it makes the institution of cultural learning more resilient (Sterelny, 2012a). As there are more models, the probability is lessened that cultural information is being lost. This is of major adaptive importance for a cultural species like the human one (Henrich, 2015; Heyes, 2018). (It is also of crucial importance in the context of emerging technologies involving artificial intelligence (AI), as will be made clearer in Chapter 8.)

4.3.2 The Technology–Representational Decision-Making Nexus

The idea that technology can enhance complex and abstract representational decision-making was the topic of Section 4.2 and thus does not need to be restated here. However, what is important to note is that there is also a reverse impact of the move from the reliance on more complex and abstract mental representations to more sophisticated technology. In particular, as organisms are able to rely more often on more complex or abstract concepts like [CAUSES], [IS A FULCRUM], [IS A MARKET CLEARING EQUILIBRIUM], or [IS A PRIME NUMBER] they are able to build more complex kinds of technology (Brisset, 2019; Guala, 2016a; Osiurak & Reynaud, forthcoming; Reindl et al., 2018; Tennie et al., 2009; van Basshuysen, 2023).

As noted earlier, it is controversial to what extent (some of) these representations are innate and required components of the ability to build tools at all. However, what is not controversial is that understanding the causal, epistemic, or mathematical structure of the world is useful for manipulating the world to manufacture ever more complex tools—including tools that aid the understanding of the causal, epistemic, or mathematical structure of the world (Bender & Beller, 2016, 2019; Muthukrishna & Henrich, 2016; Osiurak & Reynaud, forthcoming; Reindl et al., 2018; Tennie et al., 2009). The availability of more complex and abstract mental representations at least *enhances* the technological competences of organisms—and thus enables them to build ever more complex technology. The present account thus combines the perspectives of Tennie and Over (2012), van Schaik and Pradhan (2003), Osiurak and Reynaud (forthcoming), and Vaesen (2012): both individual competence and cultural learning matter to technological competence because these can reinforce each other. While our innate individual technological competences need not be drastically different from what is found in non-human animals, once we are able to rely on more complex and abstract mental representations—through more basic forms of technology—ever more sophisticated forms of technology become accessible to us.

4.3.3 The Cultural Learning–Complex and Abstract Representational Decision-Making Nexus

Finally, there is a direct relationship between cultural learning and the use of complex and abstract mental representations. Given the fact that cultural learning can be cumulative, if the starting place of the cultural learning can be highly abstract and complex thoughts, then cultural learning can make these yet more abstract and complex (Heyes, 2012a, 2018; Tennie et al., 2009; Tomasello, 1999). This thus makes for another explanation for why humans can rely frequently on many complex and abstract mental representations: the ability to start the process of cultural learning with complex and abstract mental representations (e.g., because it has been harnessed by sophisticated technology) enables humans to cumulatively learn increasingly more complex and abstract representations from others.

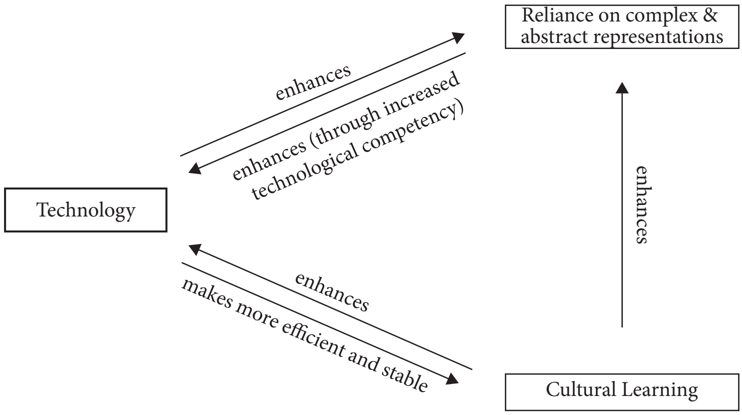

Figure 4.4 A positive feedback loop between cultural learning, technology, and representational decision-making.

Putting these three sets of interrelations together creates an overall positive feedback loop between cultural learning, technology, and complex and abstract mental representations. Cultural learning can make it possible to cumulatively manufacture and use the kinds of complex technology that, directly and indirectly, can be used to make it easier to frequently rely on complex and abstract mental representations, which in turn significantly expands the quantity and stability of cultural learning (Figure 4.4).[14]

It is also possible to put at least a rough timeline on this looping evolutionary trajectory. It is reasonable to see the feedback cycle as coming into existence somewhere around 3.5 million years ago; at this point, cranial capacity (e.g., of *Australopithecus afarensis*) was only slightly larger than that of *Pan troglodytes*, and early hominins are generally thought not to have been too different

[14] In this way, this account shares some affinities with that in Ross (2004, 2007). However, the present account goes beyond the latter's focus on language, narrative tools, and social coordination and also emphasizes other cognitive and social technologies—writing, accounting conventions, markets, banking, etc.—as well as the latter's impact on cognition more directly.

from non-human primates (Shultz et al., 2012). A handful of iterations of the feedback loop then take place until about 1 million to 500,000 years ago##; at this point, complex tools (such as weighted javelins) are being made and human encephalization greatly increases (Anton et al., 2014; Barham, 2013). The feedback process then starts to really bite, and the cycle accelerates by at least 40,000 years ago: at this point, humans create musical instruments and other complex tools (including social institutions) that enhance and which are enhanced by the reliance on complex and abstract mental representations and cultural learning (Adler, 2009; Killin, 2018; Morley, 2013). Finally, with the advent of writing and sophisticated computational devices about 5,000 years ago (Nissen, 1985), the loop is put into full swing, and advances in the reliance on complex and abstract mental representations, technology, and cultural learning greatly accelerate each other.

4.4 An Interactionist Explanation of Distinctively Human Cognition

Taking a step back, it is now possible to return to the questions left open at the end of Section 4.I and lay out how the present account can answer them.

1. How can we explain both the cultural similarities and cultural differences in cognitive traits across people?
   Answer: The fact that all humans start out with the same sorts of representations and dispositions creates a shared basis for thought and underlies much of our cultural communalities. However, since this shared basis is differentially expanded though cultural learning and technology in a complex feedback loop, there will be many differences in the way humans think and act. In this way, the present account can account for both cultural communalities and differences—a key benefit over other

accounts. The details of how this works are best illuminated in the context of particular issues. This is thus a point to which I return in the next three chapters.

2. How can we use our evolved sets of representations to make decisions?
   Answer: Technology provides the means to close the gap between the narrowed hypothesis space created by our innate expectations and the actual resulting decisions. It is because we have tools available that allow us to store elements of the relevant representational inferences, as well as potentially even doing these representational inferences themselves, that we are able to employ many even highly abstract and complex representations in our interactions with the environment.

3. Why is it that humans are able to build tools of such complexity, rely on concepts with such complexity, and be cultural learners of such complexity when other organisms do not have these abilities? What explains the divergence here, exactly?
   Answer: It is (partly) because humans are cultural learners that they are able to build the kinds of tools that allow them to rely on highly complex and abstract concepts—and it is (partly) because they are able to build the kinds of tools that allow them to rely on highly complex and abstract concepts that they are cultural learners of such complexity. For other organisms, the positive feedback loop did not get going. They were not sufficiently strong cultural learners to build the kinds of tools that would allow them to rely on highly complex and abstract concepts—and thus, they were not able to become sufficiently strong cultural learners to begin with.[15]

[15] Since positive feedback loops need not be deterministic, this can explain why technological competence can sometimes decrease (Jagher, 2016; Premo & Kuhn, 2010). If one step in the cycle happens to fall below the needed threshold, the remaining steps are more likely to do so as well.

4. Why do humans have extensive abilities in *all* of these dimensions?
   Answer: All of these abilities co-evolve in a positive feedback loop. Advances in one of them are likely to be coupled with advances in others: it is *not* coincidental that these abilities are co-instantiated in humans (Gibson & Ingold, 1993; Tennie & Over, 2012; van Schaik & Burkart, 2011; van Schaik & Pradhan, 2003).

In a bit more detail, consider the cognitive abilities inventoried in Chapter 2. Humans, by virtue of their social nature and lack of physical adaptations to their environment, experience selection for a number of lower-level cognition abilities—like a reduced discount rate, a motivation for reciprocal cooperation, enhanced fine motor control, and social toleration—as well as some higher-level abilities—like expectations about the causal structure of the world and enhanced communicative abilities, as well as various other representational expectations. However, by themselves, the resultant differences are not drastically removed from what is found in non-human animals.

Rather, it is by virtue of the interplay among these capacities that yet further biocultural selection pressures come to be, which in turn create increasingly greater differences in the relevant cognitive traits. So, as cultural learning increases in power, human technological abilities increase—which, in turn, creates further pressures on enhanced motor control (to use the relevant tools), innovation, risk-taking, and patience (to increase the power of the available tools). In turn, the latter makes it possible to rely on ever more complex and abstract representations, further enhancing cultural learning and technological sophistication. The upshot are the distinctively human cognitive capacities for language, tool use, mindreading, moral cognition, reciprocal altruism, innovation, motor control, etc. noted in Chapter 2.[16]

[16] See also Barrett (2018).

At this point, it is important to come back to the point briefly raised earlier as to whether these distinctively human cognitive capacities should be seen to differ in kind or merely in degree from what is found in non-human animals. In particular, some might agree with the picture concerning the feedback processes laid out here but note that such processes can be nonlinear and thus lead to "phase shifts" in the evolution of cognition: the initially small human–non-human differences may create dynamics that do not just build each other up, but actually lead to major transformations in the *kinds* of resultant cognitive traits. For example, Godfrey-Smith argues that the evolution of consciousness is the result of simpler forms of cognitive capacities, such as reflexive sensory responding, that have been *transformed* by the addition of other cognitive capacities, such as sensory integration, a stronger self–other distinction, and a deeper coupling between sensing and acting (see, e.g., Godfrey-Smith, 2016, pp. 92–95; Godfrey-Smith, 2020; see also Ismael, 2011). Similarly, Ross (2007) argues that the human, linguistically mediated ability to creative a "narrative selfhood" (roughly, the creation of ourselves as characters in social "plays") sets up a phase shift that led to the massively increased social coordination and cooperation found in humans.

However, while not ruling out the existence of such phase shifts in the evolution of human cognition, the present theory is not built on them. In the main, this is because the difference between a large difference in degree and a phase-shifted difference in kind is not so clear either empirically or theoretically. That is, it is just not so clear when a (large) difference in cognitive abilities amounts to a genuine difference *in kind*. There is no obvious analogue to different states of matter in the context of human cognition: How do we individuate kinds of thought, such that human mindreading, say, can be shown to be different *in kind* from what is found in non-human animals?[17]

[17] In fact, the classification of states of matter is complex even in physics and chemistry: see, e.g., Hsu et al. (2016).

Again, this is not to say that a phase transition in human cognitive evolution did or could not have happened; the point is just that the present theory does not require or presuppose it.

With this in mind, it is now useful to note how the answers to Questions 1-4 sketched here differ from those provided by the accounts sketched in the previous chapter—as well as one further account that, while not directly an account of distinctively human cognition, still deserves mention here. (Again, more details on particular aspects of these answers will be provided in the next three chapters.)

First, unlike the nativist modular accounts, the present theory does not leave it open about how to explain human cognitive diversity—this follows from the centrality of technological enhancement of representational decision-making and the importance of cultural learning for human cognition. It also does not leave it open about how to explain the ways through which we manage to actually *use* our innate psychological endowment—again, technological enhancement is key here. However, the present account shares with the modular account in emphasizing the importance of innate psychological preparations for dealing with our environment—and thus it shares with the modular account the ability to predict the existence of many cross-cultural human cognitive communalities.

Second, when considering the cultural learning–based accounts, there is again much that these accounts and the present one share. Both give pride of place to cultural learning in explaining human cognitive uniqueness, and they are thus equally able to predict the existence of much cultural variation. However, the present account differs from the cultural learning–based accounts in that, unlike the latter, the present account is not committed to a strong anti-nativist stance about human cognition. This is helpful in explaining the cultural communalities across people: fundamental building blocks of cognition are shared across cultures. Second, the present account notes the impact of the manufacture of physical tools—symbolic mnemonic devices, the abacus, sundials—for making sense of how it is possible for us to be the kinds of cultural learners we are. As noted in the previous chapter, it is not clear that the existing accounts are able to do this. Technology is also crucial for explaining how we can rely on complex mental representations (whether culturally learned or not)—something else the cultural learning accounts leave open. Specifically, the present account adds an explicit treatment of the ways in which (a) cultural learning not only allows for the manufacture of more complex tools, but is also amplified by the existence of complex tools; and (b) complex tools allow for the use of more complex and abstract representations, which in turn allow for the manufacture and use of more complex tools.[18]

Third, the present account agrees with the symbol processing account on the idea that systematic, abstract, and complex representational cognition is central to human cognitive uniqueness. However, it expands on the latter in two dimensions. First, instead of seeing enhanced cultural learning just as the *outcome* of human representational abilities, it also sees it as a *source* of the former—both in terms of content (where the representations come from) and in terms of the use of complex and abstract representations

[18] Osiurak and Reynaud (forthcoming) discuss tool use and tool manufacture, but they do not integrate this into a positive feedback loop with cultural learning and evolved representations. Much the same goes for the niche constructionist stance of Fogarty and Creanza (2017). In particular, the point emphasized in the present account is less about altering the niche in which humans live, but instead concerns specifically the existence of a positive feedback loop among evolved complex and abstract representations, technology, and cultural learning. So, while Fogarty and Creanza (2017) note that agricultural food production can buffer human populations from environmental instability, thus supporting technological innovation even among small populations, they do not consider the ways in which technology can enhance and be enhanced by both cultural learning and the use of complex and abstract representations—as is done here. A related point concerns the (different) niche constructionist view presented in Ross (2007) and mentioned earlier. Similarly, while Sterelny (2012a) briefly notes that changes in technology can prepare the ground for enhanced cultural learning, he does not consider in any kind of detail the kinds of connections among evolved complex and abstract representations, technology, and cultural learning here laid out. Something similar goes for the account of Tomasello (see, e.g., Tomasello, 1999, pp. 208–209), in which he hints at the importance of technology, but does not spell out these hints.

(through culturally acquired cognitive technology). This not only expands the account by providing explanations of where some of these complex and abstract representations are coming from, but it also closes a lacuna in the symbol processing account: through appeal to cognitive technology, it makes clearer how humans are able to use many complex and abstract representations. The second dimension of expansion of the present account vis-à-vis the symbol-processing account is in terms of an explanation of cultural communalities and differences in cognition—something the latter leaves open. By making explicit the existence of innate representational expectations, the present account can fill a lacuna in the symbol processing account. Relatedly, the present account embeds the explanation of low-level features of human cognition in that of high-level features: the former are precursors of the kind of feedback loop sketched in Figure 4.4 and are also further enhanced by the latter. This thus closes another loophole of the symbol processing account.

A final picture of human cognition that, while somewhat different in focus, is worth mentioning here is the work based on the "extended cognition" framework (Clark, 1997, 2008, 2013; Colagè & d'Errico, 2018; Dennett, 1995, 2000; Russo, 2023; Sterelny, 2017). According to the latter, understanding cognition *in general* (in human and non-human animals) cannot be done by seeing it as limited to what is going on inside a brain or even a body—either because cognitive states literally *extend into* the social and non-social environment or because they are so embedded in the social and non-social environment that not including the latter in our theorizing about them would lead us to miss important cognitive phenomena.[19]

[19] This extended cognition perspective partly cross-cuts some of the other accounts just mentioned (especially that falling under the niche constructionist framework: Sterelny, 1999, 2012a, 2017, 2018).

There is no question that the present account of the evolution of human cognitive uniqueness, with its emphasis on cultural learning and technology, has many affinities with the work on extended cognition. However, there are also several important differences to note. First, the present account, unlike some of the major accounts in the extended mind literature (see, e.g., Clark, 1997, 2008), is not metaphysical: no claims are being made here about where cognitive states begin and end. Second and most importantly, it is again the specifics of the account here—namely, the existence of a self-reinforcing enhancement process between the frequent use of complex and abstract representations, cultural learning, and technology—that set it apart from what is found in the extended cognition literature to date. Other accounts do not spell out this process in the way that is done here.

All in all, therefore, the core of what makes the present account stand out from what has been presented in the literature up to now is that it expands the set of relationships that need to be recognized as influencing the evolution of human cognitive uniqueness. While other accounts have also looked at aspects of the frequent use of complex and abstract representations, technology, and cultural learning, they have not looked at these as creating a positive feedback loop. That is (picking up a point made earlier), the goal of the present account is not to downplay the importance of mindreading, language, technological competence, cultural learning, or symbol processing. Rather, the point is that without paying attention to the positive feedback loop among frequent use of complex and abstract representations, technology, and cultural learning, a compelling account of human cognitive uniqueness cannot be provided. While these different elements have other underlying enablers and presuppositions, they also influence each other, and this needs to be kept in mind when making sense of human cognitive uniqueness.

4.5 Summary and Upshot

A plausible inroad into the explanation of the nature and evolution of distinctively human cognition lies in the set of complex and dynamic relationships between cultural learning, technology, and the frequent use of complex and abstract representations. Cultural learning enables humans to build the kinds of tools that allow their decision-making to be based on many highly abstract and complex mental representations, and the ability to rely on many highly abstract and complex mental representations, in combination with sophisticated technology, expands their cultural learning in its possible content and sources. While other animals may also have some of these elements—they can engage in some cultural learning, in some technology use and manufacture, and in the use of some kinds mental representations—none of these elements appears sufficiently far advanced so that it can enhance the remaining elements in a positive feedback loop: in the human lineage, there was just sufficient selection on these elements individually to kickstart their positive, reinforcing interactions. In short, what makes human cognition so different from that of other organisms is that cultural learning, technology, and the frequent use of complex and abstract representations have pushed—and continue to push—each other to new heights.

In this way, we can make progress in explaining distinctively human cognition: we arrive at a novel account that combines the best of the previous perspectives—it places great emphasis on cultural learning but still has room for innate, systematic, complex, and abstract mental representations—and that expands on the latter in important ways. In particular, by adding the development of cognitive technology to the picture, the existence of a positive feedback loop becomes visible that can make inroads into the explanation of human cognitive uniqueness.[20]

[20] Moreover, it can do so in a non-mysterious and still fully representational way—which is not true of some other (extended cognition-focused) accounts.

What is needed now is to leave behind these relatively abstract descriptions of the account and actually apply the latter to some key examples of distinctively human thought. This can bring out the working parts of this account more clearly and also brings out the ways in which this account differs from the existing ones found in the literature. This is the goal of the next three chapters, focusing on mindreading, moral cognition, and trade specifically.

5

The Origins of Distinctively Human Mindreading

A Bio-Social-Technological Coevolutionary Account

Without question, one of the most widely discussed aspect of human cognitive uniqueness is our ability to figure out what others are thinking—that is, to “mindread.”[1] While culturally somewhat relative and not fully reliable, human mindreading is used frequently, in situations of major adaptive importance, and can involve highly complex and abstract mental states. (As noted in Chapter 2, the human ability to mindread may also underlie another human uniqueness: religiosity. This will become important again in Chapter 8. For more on this, see, e.g., Atran, 2002; Boyer, 2001; Henrich, 2020; Sperber, 1975, 1996.) What is less clear, though, is how to explain the origins of this ability. What factors underlie the evolution of distinctively human mindreading (Andrews, 2012; Penn et al., 2008; Tomasello & Herrmann, 2010)?

This chapter uses the general principles behind the account sketched in the previous chapter to develop a novel theory of the evolution and nature of distinctively human mindreading that answers these questions. As also noted in the previous chapter, this new theory goes beyond the existing biologically adaptationist

[1] In this chapter, I use the terms “mindreading” or “mentalizing” to refer to the ability to attribute mental states to others. To what extent this is a theory-like or simulationist process is left open here (which is why the term “theory of mind” is avoided). For more on this, see, e.g., Goldman (2006); Nichols & Stich (2003); Saxe et al. (2006).

It’s Only Human. Armin W. Schulz, Oxford University Press. © Oxford University Press 2025.
DOI: 10.1093/9780197800188.003.0005

(Carey, 2011; Carey & Spelke, 1996; Carruthers, 2006; Nichols & Stich, 2003) and cultural-learning based ones (Heyes, 2018) and argues that, since human mindreading is cognitively very costly, further cognitive and social tools are needed to make it work efficiently (or at all). This implies that human mindreading is embedded in a positive feedback loop with cultural learning and sociotechnological development—the acknowledgment of which brings to view overlooked areas of investigation.

The chapter is structured as follows. In Section 5.1, the nature of human mindreading is laid out in more detail. Section 5.2 makes precise which issues the existing accounts of human mindreading leave unexplained. On this basis, Section 5.3 develops the new account of the origins of distinctively human mindreading and contrasts it with two further accounts of the evolution of human mindreading that also go beyond the biological adaptationist and the cultural learning–based ones. Section 5.4 summarizes the discussion.

5.1 Human Mindreading

What makes human mindreading unique? As noted in Chapter 2, it is not that humans are the only organisms that can do any kind of mindreading. For example, there are good reasons to think that chimps can determine what others can see and that they can empathize with others (Andrews, 2012; Tomasello et al., 2005; Tomasello & Herrmann, 2010), and something similar holds for dolphins, elephants, and magpies (Parker et al., 1994; Suddendorf & Butler, 2013; Toda & Platt, 2015). It is true that these findings are somewhat controversial still (see, e.g., Penn et al., 2008; Povinelli, 2003), but, for present purposes, it is sufficient to note the following two points.

First, it is very plausible that non-human animals can engage in many forms of nonconceptual or nonrepresentational "low-level" mindreading, such as emotion attribution, behavioral mirroring, and other forms of implicit mental state tracking (Apperly & Butterfill, 2009; Butterfill & Apperly, 2013; Edwards & Low, 2017; Goldman, 2006; Heyes, 2018; Nichols & Stich, 2003; Zawidzki, 2011).[2] For example, many non-human animals seem to show empathy or sympathy and can at least track some perceptual and cognitive states (Acebo & Thoman, 1995; Apperly & Butterfill, 2009; Bowlby, 1958; Churchland, 2011; de Waal, 2008; Edwards & Low, 2017; MacLean, 1985; Schulz, 2017).

However, these "submentalizing" (Heyes, 2018) forms of mindreading will not be central in what follows. This is not because they are not significant, or because there are no human–non-human differences whatsoever in them. Rather, it is because they deserve an inquiry of their own: the low-level mindreading system is sufficiently different from the other, high-level one—to be laid out in more detail momentarily—to call for an investigation of its own (Apperly & Butterfill, 2009; Edwards & Low, 2017). Given this, the focus here is solely on the high-level system: that is, the focus here is on the explicit representation of the mental states of others. This is an inherently interesting topic, and it also marks the strongest contrast between human and non-human forms of mindreading (Apperly & Butterfill, 2009; Edwards & Low, 2017; Heyes, 2018).

Second, though, it is further important to note that it is possible to simply accept here that there are some non-human animals that can engage in some forms of this "high-level" (i.e., conceptual and representational) mindreading as well. Apart from the fact that this assumption is empirically not implausible, it does not bias the discussion here. The point in what follows is to show that there is still something uniquely human when it comes to "high-level"

[2] The distinction between low- and high-level mindreading is far from clear (Goldman, 2006; Schulz, 2011a). It is here used only for expository purposes, though, and is not central to the argument.

mindreading; all of the arguments that follow will go through even if it were to turn out that non-human animals cannot engage in "high-level" mindreading after all. In particular, there are four key features of high-level human mindreading that make it stand out from the mindreading abilities of other animals.

5.1.1 Attributive Sophistication

Humans can and do attribute mental states to others in ways that are cognitively sophisticated. There are several dimensions to this cognitive sophistication.

5.1.1.1 High Orders of Intentionality

Humans frequently attribute mental states at the third or even fourth level of intentionality: for example, I might attribute to you the hope that I think that you think the treasure is hidden over there (because I think you are trying to deceive me), or I might attribute to you the fear that I think that you think that I am arrogant (due to a Jane Austen-like misunderstanding, say) (Bennett, 1991; Sterelny, 2003). Indeed, if it turns out that the systematicity of human thought is extensive (i.e., if something like the language of thought hypothesis turns out to be true), then humans can attribute thoughts at arbitrary levels of intentionality (Fodor, 1975; Fodor & Pylyshyn, 1988). By contrast, it is not clear whether non-human animals even attribute second-order thoughts to others (Bennett, 1991).

5.1.1.2 Mental State Concept Complexity

Apart from perceptual states, belief states, and basic emotional states like fear and hunger, humans can *also* attribute more complex mental states that are composed out of other, not necessarily atomic, mental states ("He is just hangry;" "That's just wishful thinking on her part"). By contrast, non-human animals seem to generally attribute less complex mental states only (Andrews, 2012, 2015; Hare et al., 2000; Kanngiesser et al., 2020; Tomasello et al., 2005; Tomasello & Herrmann, 2010).

5.1.1.3 Mental State Concept Abstractness

Humans can and do attribute mental states that are highly abstract and which lack straightforward links to behavior. So, humans can attribute states like *feeling slighted*, *nostalgic*, or *proud*. These are not states that are the outcomes of easily determinable circumstances (like perceptual states) or which are tied to specific behavioral patterns, as is true for the basic emotions (Ekman, 1989; Ekman & Rosenberg, 1997; LeDoux, 2012; Sauter et al., 2010). Non-human animals appear to mostly attribute less abstract mental states (Allen, 1999; Bennett, 1991).

5.1.2 Attribution Frequency and Importance

Humans use mental states attributions very often in their decision-making, and these attributions are often crucial factors determining the outcomes of these decisions. It is part and parcel of the nature of human sociality that we interact based on our intentions and other mental states. We live in groups consisting of kin as well as non-kin and are utterly dependent on each other for our survival. As noted in the previous chapter, this makes it plausible that successful social interaction requires being attuned to quickly shifting social conditions (alliances, etc.). In turn, this kind of social interaction requires or at least rewards an understanding of others' mental states (Boyd & Richerson, 2005; Cosmides & Tooby, 1992; Henrich, 2015; Henrich & McElreath, 2007, 2011; Sterelny, 2003, 2012a; Tomasello, 1999; Warneken et al., 2011).

Of course, this is not to say that mental state attributions are not important for non-human animals, too (Andrews, 2012). On the contrary, what others can see, where they think food is hidden, or even what they think I can see or know about where food is hidden are important elements in animal decision-making, too. However, the key point here is just that mindreading is a particularly central facet of human decision-making (Tomasello & Herrmann, 2010; Warneken & Tomasello, 2006).

5.1.3 Cultural Variability

Human mindreading is not culturally homogeneous. While mindreading seems to be a part of all human cultures (Henrich, 2020, pp. 67, 76, 129; Tomasello, 2021), different cultures still differ somewhat over how to mindread others (Luhrmann et al., 2015), in the emphasis they put on mindreading others (Henrich, 2020), and in the development of mindreading (Mayer & Träuble, 2013; Shahaeian et al., 2011).

So, in individualist ("WEIRD"—Western, educated, industrialized, rich, and democratic) cultures, children seem to be able to attribute beliefs different from their own before they can determine what others know; children from more holistic (non-WEIRD) cultures show the opposite (Shahaeian et al., 2011). Similarly, in some cultures, it is more common to think of others as acting in line with social role expectations rather than stable character states, and in other cultures the opposite is the case (Henrich, 2020). For example, in South Korea, a young person contradicting an elder might be taken to be in an angry state of mind—given its violation of social role expectations about deference to elders—whereas in the United States, someone like that might be taken to be in a relaxed, comfortable, or playful state of mind—given social role expectations of people having open discussions with people they respect, value, and feel comfortable with, regardless of their age. An account of distinctively human mindreading thus needs to provide an explanation of the fact that there are *both* cross-cultural communalities *and* differences in human mindreading.[3]

5.1.4 Attributive Unreliability

Human mindreading should not be taken to be perfectly reliable (Andrews, 2018; Nisbett & Wilson, 1977). This is evidenced by much of human art and history: the latter are not just full of instances of mindreading playing a major role in people's lives, but also of mindreading *failures*. (Consider the polyandrous marriage of Draupadi and the Pandava brothers in the Mahabharata: this rests on a misunderstanding of what the Pandava brothers' mother wanted—which is in turn due to a misunderstanding of what Arjuna was trying to tell her.) Theoretically, this is not surprising since mindreading others is a nontrivial inference problem. As noted earlier, similar kinds of behaviors can result from very different mental states, so that figuring out what someone else is thinking from looking at their behavior contains many unknowns and uncertainties (Lurz, 2018). This also explains why mindreading is cognitively costly and takes significant resources in terms of time, concentration, and attention (Schaller et al., 2007).

This is not to say that we are hardly ever able to form reasonably accurate representations of the thoughts of others. Rather, the point here is just that humans cannot be seen to *always* mindread others in a *completely* error-free manner. At times, we find others' behavior inscrutable, and at other times we misattribute thoughts to them. This less than full reliability in mental state attributions

[3] Note that this issue cross-cuts the question of whether mindreading is simulationist or non-simulationist in nature (Goldman, 2006; Nichols & Stich, 2003). Even if we all used the same decision-making processes, and even if these same processes are key components in mental state attribution, there can be cultural variation in mindreading, as there may be cultural variation in what inputs we use for our simulations, as well as in how we interpret the outputs of that simulation (Rossi, 2014). This point becomes important again in Section 5.2.

needs to be kept in mind when formulating an account of the distinctive nature of human mindreading.[4]

5.2 Explaining the Evolution of Distinctively Human Mindreading: Biological Adaptations versus Cultural Learning

Several influential accounts have been put forward to explain the evolution of distinctively human mindreading. Before laying out these accounts, though, it is important to stress again that these accounts focus on only one aspect of the human mindreading system: the one dedicated to the explicit representation of other's mental states. As noted earlier, this explicit, representational mindreading system should be seen as layered on top of an implicit mindreading system that humans may share with some non-human animals (see, e.g., Apperly & Butterfill, 2009; Edwards & Low, 2017). What follows is thus restricted to the former system only. That is, what follows solely concerns mindreading with the four sets of features sketched in the previous section.

The first influential account of this kind of explicit, high-level, distinctively human mindreading sees it as the result of the evolution of a dedicated neuropsychological mechanism that has been selected for to enable humans to deal with the particular social environments in which they live. Several different versions of such an account can be distinguished.

The major one of these sees human mindreading abilities as the result of selection pressures surrounding the fact that humans needed to navigate a special kind of social environment (Baron-Cohen, 1997; Byrne & Corp, 2004; Humphrey, 1986; Leslie, 1994; Sterelny, 2012a; Whiten & Byrne, 1997). As noted earlier,

[4] This is something mindreading shares with many other mental capacities, from perception to episodic memory.

adaptively accomplishing this navigation requires humans to make decisions that take into account what others want, think, and feel. Importantly, this is true even in relatively "tight" societies in which human action is highly constrained by its social environment (Elster & Gelfand, 2021; Henrich, 2020). There is a myriad of factors that can influence human decisions. Human cultures are complex and allow for many different social roles that can change over time. This means that individual humans need to figure out how to balance the different roles and relationships they have, even if their decisions are mainly the result of the social roles they inhabit and the social relations they have (Hollis, 1994, chap. 8). The fact that person P was a reliable foraging partner yesterday does not mean that P will also be one such today: P may have obtained other social obligations or requirements in the meantime that imply that P will not be highly focused on this particular task. To deal with this kind of environment, it is therefore advantageous to be able to understand and predict the mental states of others. It is these mental states—intentions, emotions, expectations, etc.—that drive human behaviors and which determine how humans manage their different and dynamically changing roles.[5]

According to the most common version of the modular account of the evolution of human mindreading, it is precisely this advantageousness that has led to the evolution of a specific set of psychological expectations. These expectations include representations about (1) when to mindread (e.g., if the focus is the behavior of other people, rather than that of insects); (2) what to consider when mindreading (e.g., where someone looks and what they did in the past), (3) what some of the possible types of states are

[5] It is noteworthy that it is not clear whether early human societies were looser or tighter. On the one hand, it seems that the WEIRD societies we encounter today are derived from tighter ancestral human cultures (Henrich, 2020). However, since cultural institutions themselves culturally evolved, it is not clear what is true about earlier hominin (e.g., Australopithecus) societies. Also, chimp societies seem relatively looser (Brosnan et al., 2008; Byrne & Corp, 2004; Whiten & Byrne, 1997). At any rate, as noted in the text, this does not materially affect the conclusions here reached.

that could drive the behavior in question (e.g., beliefs, motivations, emotions), and (4) some specific candidates of such states to consider (e.g., that the person believed that no two physical objects can occupy the same place at the same time, that parents are motivated to help their kids, or that the person wants to do something jointly with me) (Carey, 2011; Carey & Spelke, 1996; Carruthers, 2006; Tooby et al., 2005; Tomasello, 2021; Tomasello et al., 2005).

This evolved set of psychological expectations makes it possible for us to infer what others are thinking in a way that would not otherwise be possible. Without (1), we might be mindreading at the wrong times or fail to mindread at the right times (e.g., we might try to mindread a thunderstorm or fail to mindread our spouse).[6] Without (2), we might be stuck considering too many possibilities for what to pay attention to: Are my brother's mental states a function of his eye color? Without (3) and (4), we might be unsure about what sort of thing to attribute to someone else: Was their behavior caused by a stack overflow error, a desire for drinking a can of paint, or a mistaken belief that this glass contained water? Flipping this around: given expectations like (1)–(4), we narrow the inferential space when trying to predict and explain the behavior of other people. In turn, this makes us better able to handle the pressures of social living.

A slightly different version of the modular account of the evolution of human mindreading sees the latter abilities as the result of the fact that it was biologically advantageous for humans to be good communicators (Godfrey-Smith, 2002; Skyrms, 2010; Sperber, 1975, 1996). Determining what others are thinking, on this account, is important not so much because it enables us to anticipate their behavior, but because it makes it possible for us to

[6] As suggested earlier—and as will also be discussed in more detail in Chapter 8—mindreading expectations might underlie religious beliefs, and so we might interpret a thunderstorm as an expression of Zeus's wrath. However, it remains true that humans do not mindread indiscriminately.

influence each other's thoughts (and thus their behavior) through communication.

Communication was key in the kinds of social environments humans evolved in because it enables us to learn from each other and cooperate in the ways that we do, involving, for example, complex divisions of labor and trade. Importantly, though, it is not implausible that this kind of communication requires sophisticated mindreading abilities. An easy way to see this is by considering the account of meaning in Grice (1989). On this account (roughly), for person A's utterance of *p* to mean "grass is green," A has to intend for another person B to form the belief that A is, in fact, intending B to form the belief that "grass is green" (and A and B have to know this about each other). To recognize you as a communicator, I need to recognize that your communicative behavior has a certain fixed relationship to the world and that you *want* me to *think* it does.

Now, it needs to be acknowledged that this is not the only account of communication in the literature and that other such accounts do not presuppose that human communicators must be sophisticated mindreaders (Gauker, 2003; Geurts, 2019; Millikan, 1984; Moore, 2018; Skyrms, 2010; see also Brandom, 1994).[7] Fortunately, it is not necessary to settle this debate here. The point here is just that one influential attempt to ground the evolution of sophisticated human mindreading appeals to the needs of human communicators: successful communication is taken to require the mutual recognition of our mental states. If this view is denied, the problems below can be avoided, but only at the cost of finding another route toward the evolution of sophisticated human mindreading—such as the first account sketched above or the cultural learning–based account that will be sketched momentarily.

[7] Note while the account of Skyrms (2010) can be used to spell out the evolution of a Gricean picture of meaning, the former is broader than that and does not require the latter (see also Moore, 2018).

In sum, therefore, what both versions of the standard account of the evolution of human mindreading abilities have in common is the idea that there existed biological selection pressures for these kinds of abilities. The psychological details of the machinery that is hypothesized to have evolved in this way will be made clearer in the next section. For now, what is important to note is that—very much in line with the points made in the previous chapter—this account does well in explaining the first two features of human mindreading noted in the previous section: *attributive sophistication* and *attribution frequency and importance*.

In particular, the hypothesis that there was selection for mindreading, whether for predicting the behaviors of others in one's culture or for communicating with them (or both), predicts that human mindreading abilities would be (a) sophisticated and (b) used frequently and in important situations. Point (a) is a consequence of the very starting place of this kind of account: given that humans need to predict the behavior of and/or communicate with others who do the same thing and who react to the world in otherwise complex ways, too, requires that they are able to attribute mental states of high orders of intentionality, complexity, and abstractness (see also Sperber & Wilson, 1986). For example, we are not just making our behavior dependent on the weather—we are making our behavior dependent on what others think about the weather (or even what others think about our attitudes toward the weather, as when I try to impress you by wearing shorts during a snowfall). Similarly, effective and efficient communication might require the use of metaphor: we might respond to someone saying "Robert is a bulldozer" by attributing to them the thought that Robert has a stubborn and forceful personality—not that he is literally a machine. Point (b) also follows straightforwardly from the present account of the evolution of mindreading: after all, on this account, helping us make frequent adaptively important decisions well is precisely why these abilities are said to have evolved in the first place.

These points gain further support from the facts that (as also noted in Chapter 4), as far as we can tell, all human cultures mentalize in some form or other (Henrich, 2020, pp. 76–77) and that basic mindreading abilities seem to mature early (Baillargeon et al., 2010). These kinds of communalities are exactly what we would expect if human mindreading were the result of biological selection pressures deep in the human lineage—for example, they match what is true about the ability to speak a language or to walk bipedally. That is to say, the assumption that mindreading has been selected for in human evolutionary history predicts that this is a trait deeply embedded in our pan-specific human psychology and that it is therefore culturally universal and quite canalized in its development. The fact that the latter two are indeed the case thus lends credence to the hypothesis of selection for human mindreading (Forster & Sober, 2011; Sober, 2008).

However, this hypothesis also faces some challenges. In particular, in line with modular accounts of human uniqueness more generally, this hypothesis does far less well when it comes to explaining the cultural variability of human mindreading. If human mindreading abilities are a selected response to the pressures of social living, it is not clear why there is variation in human mindreading across cultures. Consider bipedalism: if walking on two legs is a selected-for response to the pressures of long-distance running and hunting (Bramble & Lieberman, 2004), it will be surprising if people in different cultures walk differently—which is not what we in fact find. Why is there this cultural variance when it comes to human mindreading, though?

At first, it would seem that this could be explained by analogy with natural languages. Even though there is considerable variation across cultures over *which* language a human speaks, the ability to speak *some* language plausibly has been selected for in human evolutionary history (Pinker & Bloom, 1990). Could we not say the same thing about mindreading? In particular, could we not see culturally different mindreading practices as different mindreading "languages" that are, however, anchored in a common "universal mindreading grammar" that is the product of natural selection deeper in the human lineage?

However, in fact, this is not a plausible response to the cultural variability of human mindreading. Unlike different mindreading practices, different languages can be equally "truthful": they can be different but equally capable means of communication and representation. That is, from the fact that two humans speak different languages, we are not forced to conclude that they represent the world drastically differently.[8] Mindreading practices, though, are more than just *means* of representation: they have the goal to *accurately* attribute mental states to others. In turn, this implies that the above account of the evolution of human mindreading would need to say that (1) people in different cultures think differently and that our mindreading practices have evolved to be culturally extremely finely tuned to the exact details of how people in different cultures think or that (2) human mindreading is maladaptively culturally variable and that it has not yet come to fixation in the population. Neither of these options is greatly plausible, though.

Response (1) does not match the fact that culturally specific genetic-biological adaptations are rare. As noted in the previous chapter, cultures change too quickly to make it possible for natural selection to lead to the evolution of traits that are specific responses to the features prevailing in a given culture (Boyd, 2018; Boyd & Richerson, 1985, 2005; Henrich, 2015, 2020; Henrich & McElreath, 2011, 2007). While exceptions to this do exist (Boyd & Richerson, 1985; Zeberg & Pääbo, 2021), these are uncommon and do not match the frequency of cultural variation in human mindreading.

[8] There can be differences in the ease with which different languages allow the expression of various truths (Boroditsky, 2001). However, the point here is just the weaker and widely accepted one that languages are, in the first instance, means of communication and expression.

Response (2) fails to be compelling as it effectively gives up on the core explanatory schema of the present account of the evolution of human mindreading. Noting that human mindreading may not yet have come to fixation in the population is of course perfectly coherent; however, it abandons the main idea behind the above account—viz., that mindreading evolved as a response to the pressures of human social living. Put more starkly, response (2) effectively amounts to an acknowledgment that the above account cannot explain the cultural variation in human mindreading.

Given these challenges for the above account of the evolution of human mindreading, it is tempting to look for alternatives. Unsurprisingly, a major such alternative is an account that sees mindreading as the result of cultural learning and not of biological evolution (Heyes, 2012a, 2018; Heyes & Frith, 2014; Tomasello, 1999, 2021, 2022; Tomasello et al., 2005).[9] On this picture, mindreading is still crucially related to the social nature of human living, but in a different way from the sort of accounts sketched above: it is the result of the fact that humans are strong cultural learners. That is, on the cultural learning–based account, it is not the case that human mindreading abilities are themselves biological adaptations; rather, they are learned from others in the culture a person lives in. As Heyes (2018) puts it, mindreading is like book-reading. On this account, therefore, the explanation of human mindreading abilities lies in the fact that humans are standout cultural learners—there was no specific biological selection on human mindreading abilities. These abilities are simply culturally learned, like riding a bicycle.

[9] A quick point of terminology: since cultural learning has also evolved by biological evolution, and since biological and cultural evolution often affect various traits simultaneously, contrasting one with the other can seem somewhat problematic (Boyd & Richerson, 2005; Henrich & McElreath, 2007, 2011). However, here, the point is just terminological: of course, the cultural learning–based account is also biological in a wider sense. However, the point at issue here is whether human mindreading abilities are the result of a canalized biological maturation process akin to beard growth or language acquisition, or whether they are the result of a cultural learning process akin to book-reading—whatever the biological origins of the latter are.

It is important to recall that this alternative account is not distinctive for seeing human mindreading as the product of *some kind of* learning. The standard, biologically adaptationist account can allow for learning, too (Tooby & Cosmides, 1992). For example, the evolved mental structures that biologically adaptationist accounts posit as underlying mindreading could in fact mostly function to scaffold our learning in such a way that we end up acquiring the kinds of mental representations and processes that allow us to successfully mindread others (Carey & Spelke, 1996; Tooby & Cosmides, 1992). In this way, this picture of the evolution of mentalizing can be made consistent with the idea that children learn to mindread by testing and revising various "hypotheses" about how minds work (Gopnik, 1996) or by learning to put themselves in the situation of others and then interpret the results of their own simulated cognitions (Goldman, 2006).[10] In particular, which hypotheses are tested and how children determine what they are thinking can be underwritten with innate mental structures (Nichols & Stich, 2003; Schulz, 2011a).[11]

In contrast to this "scaffolded-learning" picture, the cultural learning–based account of human mindreading sees the latter ability as culturally learned *in the same way* that book-reading or bicycle-riding are learned: through a *domain-general* learning process (Csibra & Gergely, 2011; Henrich, 2015; Heyes, 2016, 2018). That is, on this account, humans learn mindreading abilities by directly or indirectly observing others in their culture (who may even be explicitly engaged in teaching them the relevant skill).

[10] It is possible to interpret either of these views as being based on *purely* on individual learning, without any innate priors (or the like) grounding this learning. In that case, though, it is not clear why there is so much cultural and even cross-cultural homogeneity in human mindreading. This makes this kind of view less compelling than either an innately grounded one or a cultural learning–based one.

[11] Again, the ability to speak a language may be a good example of this sort of picture (Pinker & Bloom, 1990): the biological adaptation underlying this ability is precisely a set of psychological expectations and processes that make the learning of language possible. See also Note 9.

Their learning is not scaffolded by any specific mental structures but rather is the result of a process that underlies all other instances of cultural learning as well.

It is furthermore noteworthy that this account is perfectly consistent with the facts that human mindreading may be susceptible to being selectively damaged (Baron-Cohen, 1997) or that it may be underwritten by a specific set of neural structures (Goldman, 2006; Saxe & Powell, 2006). Cultural learning can lead to the development of specific brain structures that underlie the learned activities and which can thus be selectively damaged. Indeed, the same is true for book reading (Henrich, 2020, pp. 3–7; Heyes, 2018).

It is also clear that this account does well at explaining features (3) and (4) of human mindreading and that it does so better than the biologically adaptationist account does. In the first place, the cultural learning–based account straightforwardly *predicts* the existence of cultural variation in human mindreading (Heyes, 2012a, 2018). On the assumption that mindreading skills are learned from our culture, general cultural differences are likely to translate into mindreading differences as well. As noted in the previous chapter, it is known that cultural learning is adaptive in circumstances where the relevant environments change too quickly (spatially or temporally) to make purely genetic adaptation feasible, but not so quickly that individual learning alone or no learning is required (Boyd & Richerson, 1985, 2005; Henrich & McElreath, 2007, 2011). Since people in different cultures think slightly differently (Henrich, 2020; Nisbett et al., 2001), this could thus make for precisely one of the situations in which it would be adaptive for humans to culturally learn how to mindread (especially since, as noted above, it is not plausible to posit genetic adaptations to quickly changing environmental conditions).

Second, the cultural learning–based account also seems to sit well with the lack of full reliability in human mindreading. On the assumption that mentalizing is a culturally learned trait, many people might not end up being expert mindreaders. This matches what is true about book-reading: not all humans can read equally well—some can plow through *Anna Karenina* in a few days, whereas for others this is hard work taking years. Indeed, many humans cannot read at all. Similarly, since cultural transmission is consistent with the existence of various kinds of biases in what is learned (Boyd, 2018; Boyd & Richerson, 2005), it is possible that humans learn to mindread in a slightly unreliable manner. For example, the cultural transmission of *pedagogical techniques* to teach children to learn to read can be stable despite being in fact unreliable (i.e., not very good at teaching children to read). Mindreading could be like this, too.

However, this does not mean that there are no problems for this alternative account (see also Jacob & Scott-Phillips, 2021; Morin, 2019). In the first place, the account faces the problem noted in Chapter 3 that the power of human cultural learning is just taken for granted here. However, it is not clear that this is plausible: Exactly why is it that humans are such capable cultural learners—keeping in mind that their mindreading abilities may not be so developed, given that the latter are meant to be a result of the former? Without some further explanation, this is a major lacuna at the heart of the account. (The interactionist account defended in this chapter can be used to provide an answer to this question; however, the point is that this then requires adopting precisely this interactionist account.)

Second, the cultural learning–based account seems to *overshoot* its target when it comes to feature (3) of human mindreading. As noted earlier, while there is some cultural variability in human mindreading, it is also true that human mindreading is found in all cultures (Jacob & Scott-Phillips, 2021; Peterson & Wellman, 2009). The contrast to book-reading is actually quite telling here: relatively few human cultures are book-reading literate *at all*, but all cultures are mindreading literate. Compared to book-reading, mindreading appears highly culturally conserved. Indeed, book-reading takes many years of study (something it shares with learning how to make handaxes, bows, poison arrows, goulash, or maps: Sterelny, 2012a). While we may be able to observe some mindreading instruction in some cases—a point that is somewhat controversial (Jacob & Scott-Phillips, 2021)—we do not observe the kind of intensive learning and teaching found in other culturally learned abilities.[12]

All in all, what this means is that both the biologically adaptationist account and the cultural learning–based one struggle to make sense of the full gamut of features surrounding human mindreading—although both accounts also seem to get important aspects of the evolution of human mindreading right. The obvious solution is thus to combine these accounts; the picture sketched in the previous chapter makes exactly this possible.

5.3 The Bio-Social-Technological Evolution of Human Mindreading

At the heart of the present account of the evolution of human mindreading is the mutually reinforcing triad laid out in Chapter 4 that comprises not just biological adaptation and cultural learning, but also sociotechnological development. Applying this triad to the present, specific case of mindreading, the first step consists in noting that for either social interaction–focused or for communication and cooperation–focused reasons (or both), mindreading abilities are highly adaptive. It is therefore plausible that we have evolved psychological expectations—about when to mindread, what sort of factors to pay attention to when doing so, and what some plausible options are for mental states to attribute—that

[12] The account of Tomasello (see, e.g., Tomasello, 1999, 2021, 2022; Tomasello et al., 2005) also faces the additional problem that it sees non-human mindreading as mostly competitive in nature and human mindreading especially cooperative: this is problematic as there is also much competition in human cultures (Whiten & Byrne, 1997) and much cooperation—especially among kin—in non-human animals (see also Boesch, 2005). Indeed, there is some work suggesting that complex forms of mindreading may be particularly helpful in competitive situations (Devaine et al., 2014).

make it easier for us to infer others' mental states from their observable behaviors and other features (Csibra & Gergely, 2011; Luo & Baillargeon, 2007).

However, the present account adds to this fact the further fact that, even given these kinds of expectations, mindreading of the kind shown by humans is quite costly, in terms of both time and the need for cognitive resources like attention and concentration.[13] Even knowing when to mindread, what sort of factors to pay attention to when doing so, and what some plausible options are for mental states to attribute, very many possibilities remain open. Human behavior is too variable to allow for mindreading expectations to be so fine-grained as to narrow the space of theoretical hypotheses to consider when attributing mental states to others to zero. To narrow down these possibilities further—beyond what is provided by our evolved mindreading expectations—the mindreader needs to consult their knowledge about the person in question and any general knowledge about how human minds work in this culture. Moreover, this might need to be done quite quickly if the mindreading inference is a crucial input into a decision.

To accomplish this, significant computational resources are needed. Applying the evolved mindreading expectations to the case at hand and supplementing them with specific information about the person and culture in question takes large stores of personal memory as well as the ability to access the latter. It also requires highly flexible control over attention, concentration, and other executive functions to monitor the target and their interaction with their social, biological, and physical environment, potentially over significant periods of time.

[13] Westra and Nagel (2021) also note that mindreading is cognitively costly and therefore suggest that much of it is instead focused on "factive" states like what others *know* (see also Phillips et al., 2021). However, as the later discussion makes clear, their argument underplays the fact that cognitive technology and social roles can lessen these cognitive costs considerably.

Given that we have only finite amounts of such resources available to us and need to make decisions in real time (Gigerenzer, 2008; Lieder & Griffiths, 2020), this implies that mindreading is not a trivial task for us. In general, relying on abstract and complex mental representations when making decisions is typically computationally and cognitive-resource hungry (Lieder & Griffiths, 2020; Schulz, 2018b). As noted in Chapter 3, this is also why it is not compelling to see human mindreading as simply the result of more sophisticated (combinatorial) representational capacities in general: it may be the case that humans are able to represent more complex and abstract notions than non-human animals—however, the *use* of such representations in mindreading is costly and cannot be taken for granted. Put differently, it needs not just be stated *that* humans have more sophisticated representational capacities (i.e., point (3) of our evolved mindreading expectations) but also *how* they can employ these capacities.[14]

With this in mind, the assumption that humans have evolved mindreading expectations to aid them in the navigation of their social environment—which, as noted earlier, is plausible because it narrows the inferential space—also implies that it is highly adaptively useful for humans to have access to the right kind of social

[14] As also briefly noted in the previous chapter, it may be thought that humans evolved not just the expectations and representations that *guide* mental state attributions—such representations about which situations call for mental state attributions and what features of the situation to pay attention to when making mental state attributions—but also the mental structures that help them *make* these inferences, such as dedicated neural processing space. However, the extent to which this is in fact a coherent view is not clear: mostly, it is thought that the way to help an organism *make* mental state attribution is by providing it with representations that *guide* the inference (Carruthers, 2006; Fodor, 1983; see also the overview of a related debate in linguistics by Chater & Christiansen, 2010). At any rate, the proposal here is just that humans have evolved specialized expectations about how other minds work, but that these expectations still need to rely on the same cognitive and computational resources we use for other cognitive tasks (such as other kinds of representational decision-making). Note also that, to the extent that the alternative, processing-focused proposal is found plausible, it continues to fall prey to the concerns for the biologically adaptationist account noted in the previous section—without the ability to appeal to the kinds of considerations put forward in this section.

or physical tools that help them lessen the costs of needing to frequently rely on such expectations. As noted in the previous chapter, such tools can support the evolved psychological expectations with which humans make complex, frequent, and adaptively important mental state attributions. Two types of these kinds of mindreading tools can be distinguished. (Note that these two forms of technology embody aspects of both types S and I from the previous chapter: as noted in the text, both myths and songs function as storage devices and help make the relevant mindreading inferences.)

First, cognitive technology like books, myths, songs, and pictures can help sort out which of the large stock of evolved expectations concerning mental state attribution are especially important to focus on in the situation at hand. To deal with the complexity of human social living, humans are likely to need to rely on a large number of expectations about how to attribute mental states to others. Where someone is looking tells us something about what they were thinking; what someone did in the past may tell us something about what they are motivated to do now; who someone is related to may tell us something about what they feel like. We may have evolved to be able to make all sorts of mental state attributions; however, the ones that are most useful to do in the case at hand may be highly culturally specific and their selection thus cognitively quite difficult. Stories, myths, and songs (among other things) can help make this selection. They model mental states attributions and tell us what features of the situation to pay particular attention to if we are to single out the mental states of particular importance.

For example, in some cultures (such as a WEIRD one like the United States), ascertaining someone else's character traits (e.g., whether they are conscientious) is useful for navigating social interactions: after all, in these (individualistic) cultures, people are encouraged to act on their character traits. In other culture (such as a less WEIRD one like South Korea), ascertaining someone else's beliefs about their social ties (e.g., what status they have in relation to other members of their group) is more useful for navigating social interactions: after all, in these (less individualistic) cultures, people are encouraged to act in line with these beliefs about social ties, not their character traits per se (Henrich, 2020, pp. 32–33). In both cases, people may rely on the same set of evolved mindreading expectations; it is just that they "foreground" different ones.

Cognitive technology like stories, myths, and songs can help do exactly this. So, we may grow up listening to songs that say "I know that five years is a long time, and that times change; but I think that you'll find: people are basically the same" (Depeche Mode, *See You*), or we may grow up listening to stories like the Mahabharata, where you see that focusing on what Arjuna wants himself is not as helpful as focusing on his duties as a son to predict his actions. In both cases, the relevant cognitive technology aids us in determining the locally best uses of our otherwise similar mindreading capacities.[15] As Henrich (2020, p. 33) puts it, "being consistent across relationships—'being yourself"—pays off more in America [than in Korea]." Songs, stories, and myths can help us figure this out quickly and easily and thus improve our mindreading capabilities (see also Fabry, 2023).[16]

In this way, this kind of technology is crucial for helping humans be mindreaders. However, this is not because this technology is needed to teach humans to *become* mindreaders (as the cultural learning–based account might have it: Heyes, 2018, p. 154). Rather, this technology supports humans in *being* mindreaders. Cognitive technology can help humans make the locally best uses of the same set of evolved mental representations about other minds without breaking the bounds of their limited cognitive resources (see also Davies, 2018).

[15] It is also conceivable that this kind of technology adds some new, culturally specific mindreading expectations to the set of evolved ones (such as being "hangry" or feeling "ennui"). This is not clear, though, so this will not be further considered here. It does not alter the main point in the text.

[16] Apperly (2011) spells out a related point in terms of mental scripts. See also Spaulding (2011, 2018.

Second, complex, frequent, and adaptively important mindreading is furthermore supported by the existence of social institutions that restrict the scope of possibilities for what an agent may be motivated by or thinking about in any given situation. In turn, this narrows down the option set of which mental state attributions to foreground in the case at hand and thus makes mental state attribution much easier.

For example, if I know that you are the *lead hunter*, this gives me clues about what you might be thinking about when there is a discussion about hunting parties being put together. Where you are looking might provide me with information about whether you are hungry, and your history of social interactions with person S might tell me something about your motivations for marriage. However, since I also know that you are the lead hunter, I know that the key motivational state to attribute to you here is to organize a hunting a party with the greatest chance of success.

As with cognitive technology like stories, songs, and myths, the existence of social roles allows humans to quickly tailor their mindreading to their local environment. It further narrows human mindreading expectations beyond their evolved set to the culturally most relevant ones and thus brings the cognitive labor of mentalizing within the confines of what is possible given the limited cognitive resources open to humans.

However, also as with cognitive technology, it is important to understand this point correctly. It is sometimes thought that social role interpretation and prediction is an *alternative* to genuine mindreading (see, e.g., Sterelny, 2003). However, this is not the point made here. It may be that, at times, the appeal to social roles makes mindreading redundant because we can just go straight to the prediction of behavior.[17] However, at other times, the existence

[17] Social roles also have a normative dimension: they tell us what we *ought* to do and thereby provide a standard with which our behavior can be criticized. Indeed, this can extent to a meta-level: they might tell us that our behavior ought to stem from our mental states (see also McGeer, 2007). I return to this point when discussing the work of Zawidski below.

of social institutions can also support human mindreading by making complex and locally adaptive mental state attributions feasible with limited cognitive resources.

Summarizing, then: if humans have evolved expectations that can guide mindreading inferences, then having access to the right kinds of cognitive technology and social institutions becomes adaptively valuable, too. This can further narrow the inferential space for attributing mental states to others—despite its being out of reach of genetic adaptation—and brings it within the realm of feasibility for cognitively limited organisms like humans. This thus leads to a mindreading focused-version of the figure from the previous chapter (Figure 5.1).

How, though, do humans get access to these technological and social tools? It is here where the core idea of the cultural learning–based account of human mindreading comes into its own.

To see this, recall that that it is widely accepted that complex tools and social institutions develop by cultural learning (Boyd & Richerson, 2005; Boyd et al., 2011; Henrich, 2015). Recall also that this point holds even more strongly for highly complex cognitive technology (like writing, dances, songs, myths, or stories) and social institutions: coming up with a stable social institutions like hierarchical stratification and religious authorities—especially ones that can be maintained in large societies—is very difficult (Henrich,

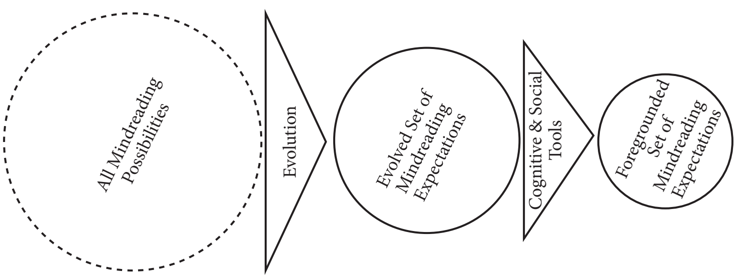

Figure 5.1 Sociotechnological narrowing of evolved mindreading expectations.

2020). In this way, the present account takes on key insights from the cultural learning–based account of human mindreading—but, unlike that account, it does not see mindreading as directly evolving by cultural learning. Rather, it views cultural learning as the process that builds up the tools and institutions that *support* mindreading.

This does not exhaust what needs to be said about the place of cultural learning in the evolution of human mindreading, though. It is not just the case that cultural learning leads to the existence of the kind of physical and social technology that can support mindreading. It is also the case that

(a) Mindreading increases the scope for cultural learning.
(b) Mindreading increases the scope for tool use and social institutional navigation.
(c) Physical and social technology can support cultural learning.

These points are often separately recognized in the literature, but they are rarely put together and synthesized. This, though, is crucial to truly understand the evolution of human mindreading.

As far as point (a) is concerned, to learn complex behaviors from others, it is often necessary to be able to determine what the intention is behind the behavior. Otherwise, the student is unable to figure out which of the myriad of highly specific physical movements the teacher is going through are part of the behavior to be learned and which are not (Csibra & Gergely, 2011; Henrich, 2020, p. 129; Phillips et al., 2021; Sterelny, 2012a; Tomasello, 1999). In other words, determining what the teacher is intending is a key element in many learning processes (viz., those that go beyond cases where simple and direct behavioral copying is possible or useful).

Similarly, as far as point (b) is concerned, an understanding of what a designer intended a tool to be used for can help others use that tool more effectively. This holds for physical tools like screwdrivers (knowing it is intended to screw rather than hammer is helpful for using it effectively) as much as for cognitive and social tools. My understanding of what a story or song is about can be increased by an understanding of the intentions of the author.[18] Knowing that *The Threepenny Opera* is intended as a socialist critique of capitalist society can help decode its lessons. Equally, an understanding of the motives and thought processes of the occupants of various social roles can help me maneuver through my institutional landscape better: to avoid falling foul of a social proscription, it can help to know what the motives are of town officials (say). This is especially important as these institutions get more complex and it gets less clear what they proscribe. It is one thing to know about incest taboos in my culture; it is another to know what is required of foreign merchants in another, geographically distant culture (Henrich, 2020, pp. 318–319). Mindreading can help determine this.

Both (a) and (b) thus make explicit that there is a reverse side to the relationship between cultural learning, cognitive technology, and social institutions, on the one hand, and mindreading, on the other. It is not just the case that the former can guide the latter; it is also the case that mindreading can guide cultural learning and the interpretation and use of cognitive technology and social institutions.

Point (c), finally, is slightly less commonly noted but also very important. Stories, myths, songs, and other cognitive artifacts enable teaching and learning on a scale that is drastically different from what is available otherwise (Henrich, 2020, pp. 436–437; Muthukrishna & Henrich, 2016). By consulting these artifacts, I can learn from many more models—including ones from different

[18] Of course, as is commonly noted in literary theory, the intentions of an author need not be crucial for an understanding of the (or a) meaning of the piece. The point here, though, is merely that there are circumstances where an understanding of the author's intentions improves the understanding of the meaning. This is uncontroversial (e.g., when it comes to instruction manuals).

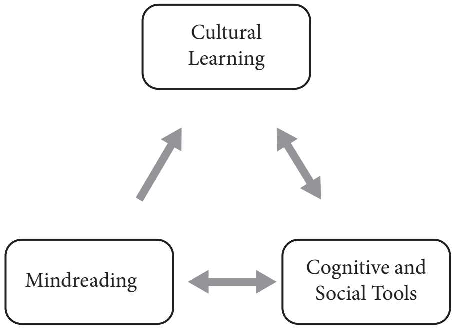

Figure 5.2 The interactionist feedback loop surrounding human mindreading.

times and places—by relying on information stored in long-term, easily accessible media. I can also go over lessons frequently, in my time and at my own pace. Social institutions can further prioritize my learning in important ways. This is obvious with formal schooling, but also applies to the more informal toleration and encouragement of apprenticeship schemes (Sterelny, 2012a). In fact, it also holds for social institutions more generally: certain social institutions can structure society in such a way that it becomes possible to focus learning on different high-status individuals in different circumstances (Boyd & Richerson, 2005). It is not just that we need to look to the village elder for all our instruction; we can look to the expert hunter for guidelines about how to hunt, the expert smith for guidelines about metallurgy, etc.[19]

The upshot of all of these different causal relationships is a complex, interlocking set of feedback loops (Figure 5.2).[20]

The key thing to note about this figure is that it makes clear that cultural learning, technological and social tools, and mindreading

[19] Note that the claim is not that *all* social institutions or all forms of cognitive technology can do this—the point is just that some can. I return to this below.

[20] Henrich (2020, pp. 299, 412) also hints at such a feedback loop but does not spell it out in the way it is done here.

support each other. Cultural learning allows for the development of complex cognitive tools and social institutions, which in turn support innate mindreading representational expectations as well as more efficient and wide-ranging cultural learning. In turn, technologically and socially enhanced innate mindreading expectations support more efficient cultural learning as well as better use of cognitive and social technology. In this way, cultural learning, technological and social tools, and mindreading can build each other up to higher and higher levels of complexity and sophistication.

What makes this account more compelling than either of the other two accounts of mindreading is that it avoids the problems of the other two preceding account while still being able to keep their advantages.

In the first place, it is due to the strength of this feedback loop that human mindreading can involve high levels of intentionality, highly abstract and complex mental state concepts and be used frequently and in major adaptively valuable situations. While doing so would generally be costly and time-consuming even for organisms that had the required representational resources, this use is supported by the existence of cognitive and social technology that helps human mindreading go smoothly.

Importantly, this also explains why non-human animals have not evolved mindreading abilities to rival those of humans despite the fact that they sometimes faced some similar ecologies. Animals may have had rudiments of cultural learning, tool use, and mindreading expectations, but these did not have the strengths and connections to kickstart the feedback loop in Figure 5.1. So there is no groundbreaking, either/or difference between human and non-human animals here; rather, there are sufficiently many differences of degree that a larger difference emerges as a result of the feedback loop in Figure 5.1 (or its absence).[21]

[21] This point is also supported by the differences in the mindreading capacities in domesticated and non-domesticated animals; see also Lurz (2011).

However, this feedback loop also explains the other two features of human mindreading: its cultural variability and lack of full reliability. Starting with the latter: it is precisely because mindreading is hard that it needs to be supported with the right kinds of cognitive tools and social institutions. Importantly, though, while these tools can *ease* our use of our evolved mindreading expectations, they cannot guarantee its success (Boyd & Richerson, 2005; Henrich, 2020). Importantly also, we do not always have access to these tools, and, where they are not available, mindreading efficiency will drop off. This explains why, apart from close social partners, our mindreading abilities are at their best when supported by well-functioning social and physical technology. It is easier to determine what incumbents of social roles in our culture are thinking than otherwise similar random strangers (Henrich, 2020, pp. 318–319). However, the present theory can still make sense of the fact that mindreading is central to human social living: since technology and mindreading capacities culturally and biologically co-evolved, humans did often although—as just noted—not always have access to the right kinds of tools to enable their mindreading capacities.

When it comes to the cultural variability of mindreading, the key thing to note is that Figure 5.1 shows how differences in the details of the cultural environment—the tools it makes available—can lead to differences in the nature of the resultant mindreading. However, at the heart of the account here is still the fact that all cultures mindread and that they do so in ways that are often quite similar. All humans rely on what is a largely the same set of evolved mindreading expectations; it is just that the support system surrounding these mindreading expectations differs, leading to some differences in the way the mindreading is actually accomplished in different cultures. In this way, it is possible to explain how human mindreading practices can be finely tuned to how people in different cultures think (Henrich, 2020; Nisbett et al., 2001) without needing to appeal *just* to genetically driven mindreading adaptations or to *fully* culturally evolved mindreading abilities.

More than that, the present account brings to view novel areas of investigation. In particular, the present account points to the need for a more detailed picture of which sociocognitive tools support which forms of mindreading. Myths emphasizing individual decision-making may favor a character-based rather than a situational form of mindreading, but what about different forms of music and dance? Do more improvisational—and thus interpersonally dependent—forms of music like jazz enhance mindreading abilities more than structured, individualistic forms of music like Western orchestral music? Do visual forms of storytelling (like comic books) lead to different forms mindreading than written forms (like traditional novels)? As yet, we do not know the answers to these questions, but, as the present account makes clear, they can greatly advance the cross-cultural understanding of human mindreading. The present account also aids in the development of tools to aid mutual human understanding. For example, it has become clear that social media does not promote social understanding: it appears to largely lead to the reinforcement of one's own perspective rather than the appreciation of others. However, other forms of technology—such as neutral media sources—may do better here—although exactly which ones is not yet known.

Before concluding, it is finally useful to contrast this account with an alternative account of the evolution of mindreading that also seeks to go beyond the biological-adaptationist and the cultural learning—based perspectives sketched earlier: that of Zawidzki (see especially Fenici & Zawidzki, 2021; Zawidzki, 2011, 2013, 2018). The goal in this is not an in-depth critique of this account, but rather to make clearer the unique aspects of the present account, so as to make explicit what it adds to the existing literature.

According to Zawidzki, at the heart of the human–non-human divergence in mindreading is the human ability to *shape* other minds. That is, we do not just take other minds as given and try to determine what states they are in; rather, we actively *regulate* the thinking of others—through social tools like myths and social norms—to narrow the space of possibilities of what others might be thinking (Godfrey-Smith, 2002). In this way, we overcome the "holism" problem, according to which any particular form of behavior is consistent with indefinitely many mental states as causes. In turn, the existence of these social tools is made possible by a lower-level mindreading system, which is just sufficient to accomplish these more basic communicative and coordinative tasks.

This account shares several key features with the present one—including, most importantly, the emphasis on social tools. However, there are also several important divergences here. First, there is little on physical technology like books, physical models, and images in Zawidzki's account, although, as noted earlier, they play a key role here. Second, the present account does not limit social technology to "mindshaping." Rather, it shows that complex mindreading can be accomplished by using social and cognitive tools that *narrow* an evolved set of mindreading expectations, as in Figure 5.1. That is, the present account does not commit to the claim that these tools are they key elements in leading people in different cultures to think differently, but rather, following the literature in gene–culture coevolutionary theory, allows for a much broader set of reasons for why people in different cultures think slightly differently (from chance to different learning practices) (Boyd & Richerson, 2005; Henrich & McElreath, 2007) and, following the standard evolutionary psychological arguments, accepts the existence of innate mindreading expectations. Finally, the present account points out the existence of the feedback loop in Figure 5.2. This matters because this feedback loop shows that there is another way to get human mindreading off the ground, without the appeal to two mindreading systems (Renner & Zawidzki, 2018): simple forms of cultural learning lead to basic forms of technology (Brown, 2018); in turn, these support more complex mindreading—and thus more complex cultural learning.

In this way, the present account shows that while the picture sketched by Zawidzki is compelling as far as it goes—the kind of "mind-shaping" he sketches does indeed plausibly occur in some cases—it is not *all* that is going on here. Indeed, the present account expands Zawidzki's theory along two dimensions. First, some instances of mindreading are shown not to involve significant mindshaping at all: they are based on a kind of "inferential narrowing" in mental state attributions instead. Second, even where mindshaping occurs, the latter is to be seen as embedded in the kind of ratchetting effect sketched in this chapter: mindshaping influences and is influenced by innate representational expectations, forms of cultural learning, and social and cognitive technology.

5.4 Conclusion

Human mindreading, although lacking in full reliability, is distinctive for its complexity, frequency, and importance. It is found in all cultures but also sees significant cultural variation in the way it is accomplished. This chapter has argued that the two major existing accounts—the biological one and the cultural learning–based one—fail to explain all of these facts. To improve this situation, the chapter then presented a new account, according to which human mindreading is the result of selected-for psychological mechanisms that are, however, supported by a ream of technological and social tools, including images, stories, and social institutions that guide mental state attribution. These tools and institutions are culturally developed. These three components also influence each other, thereby generating a positive feedback loop. A related story—though with many important differences in the details—can be told about distinctively human moral cognition.

6

The Making of Morality

The Bio-Social-Technological Evolution of Human Moral Representations

Especially noteworthy among the highly complex and abstract concepts on which human cognition depends are *moral* ones. Virtually all humans think in moral terms, human moral thought often involves intricate and hard to observe considerations (such as microaggressions or historical injustices), and moral convictions can be the prime determinants of human behavior. While there is some suggestion that proto-moral thought can be found in non-human animals, too, it is clear that human thought is outstanding in the magnitude, complexity, and significance of its moral conceptual aspects (see, e.g., Cosmides et al., 2018; Delton & Krasnow, 2015; Flack & de Waal, 2000; Hamlin, 2015; Kitcher, 1998; Prétôt & Brosnan, 2015).

Of course, it is also widely known that there is massive variation in the ways in which humans think when it comes to morality—at least on a surface level. Are we morally obligated to be vegans? To what extent should current generations be held morally responsible for the behaviors of prior generations (e.g., by paying reparations for past wrongs)? Do we have special moral obligations to close kin (children, parents)? These are fiercely contested issues.[1]

[1] It is possible that underneath these moral differences is a shared, common "universal moral grammar" (see, e.g., Cushman et al., 2006; Mikhail, 2011; Prinz, 2008). However, independently of whether this turns out to be true, it is clear that human moral cognition is not so strongly canalized that we all come to accept the same moral claims. Human moral diversity has to be taken seriously at least initially. See also Note 5 and Section 6.4.

It's Only Human. Armin W. Schulz, Oxford University Press. © Oxford University Press 2025.
DOI: 10.1093/9780197800188.003.0006

However, while the outlines of the ways in which humans rely on moral concepts are thus clear, it is not clear what *explains* this reliance on complex and abstract moral concepts in human thought. Why did humans evolve to think in these ways when other animals do not—especially since they at least sometimes seem to share the presuppositions of moral conceptual thought? Why do humans in different times and places differ in their use and application of moral concepts? It is far from obvious how to answer these questions. Using the account sketched in Chapter 4, it becomes possible to take steps toward answering these questions. In particular, with this account in the background, it becomes clear that we need to see these two questions—Why do humans rely on complex and abstract moral concepts when other species do not? Why do they do it so differently within their own species?—as deeply related. Understanding why humans evolved in such a way as to often think in terms of complex and abstract moral concepts provides us with the key elements for an understanding of what leads human moral concepts to differ across humans and vice-versa.

The chapter is structured similarly to the previous one. Section 6.1 makes the nature of human moral thought clearer. Section 6.2 makes precise exactly which questions the existing accounts of the evolution of human moral thought leave open in this context. Section 6.3 provides the building blocks of a novel account that helps address these concerns. Section 6.4 lays out the biocultural-technological evolutionary dynamics evoked by these building blocks. Section 6.5 shows that these dynamics can explain the inter- and intra-specific diversity in human moral thought. Section 6.6 concludes.

6.1 Human Moral Conceptual Thought

Delineating the nature of human moral thought is far from straightforward. Fortunately, a complete account of all aspects of human moral cognition is not necessary here, as the present focus is narrower (see also Machery & Mallon, 2010). Specifically, at stake here is just to explain (1) why humans, but not non-humans, rely extensively on complex and abstract moral concepts; and (2) why there is much variation in the ways in which humans rely on moral concepts. To make answering these questions possible, it is best to begin by distinguishing these questions from a number of other questions being asked in this context.

First, the issue here is not the explanation of all aspects of human moral cognition, but specifically its sometime reliance on complex and abstract *concepts* like [MORAL OBLIGATION], [JUSTICE], [MORAL DUTY], and [MORAL VIRTUE]. To be sure, human moral cognition may also involve an extended ability to empathize and sympathize with others (Hamlin, 2015; Kumar & Campbell, 2022, chaps. 1–2; but see also Prinz, 2011), to be altruistically motivated toward a wide group of others (Kumar & Campbell, 2022, chap. 1; Schroeder, 2000; Stich et al., 2010; but see also Stich, 2016), and to feel a range of reactive emotions in the appropriate circumstances (Campbell & Kumar, 2012; Strawson, 1962, chap. 2). There are many interesting questions to be asked about why humans evolved in such a way as to have these latter psychological abilities, how they develop, and how they function (Cosmides et al., 2018; Hamlin, 2015; Machery & Mallon, 2010; Piccinini & Schulz, 2019). However, these questions are not our focus here.

Rather, much like the case in the discussion of mindreading, at stake here is specifically the question of why humans rely on moral *concepts* (i.e., on moral *representations*). While nonrepresentational moral emotions and motivations are important elements of human moral cognition, too, it is not possible to make sense of human moral thought without explaining its dependence on moral concepts. On the one hand, these are separate sets of psychological abilities: organisms may have the former without the latter. For example, the ability to empathize may be quite widely instantiated in the biological world, whereas the ability to rely on many highly abstract and complex moral concepts is not (Campbell & Kumar, 2012, chap. 2; Piccinini & Schulz, 2019; Schulz, 2017). On the other hand (and as will also be made clearer below), moral conceptual thought is a unique and important element of many human actions. It is not a minor aspect of human moral behavior, but absolutely central to it. It thus deserves a separate inquiry of its own (see also Kelly, 2018, pp. 347–348; Machery & Mallon, 2010).

Second and relatedly, the issues here also do not concern whether human moral thought is cognitive (belief-like) or conative (desire-like) (Horgan & Timmons, 2006; Schroeder, 2010). The issue here is explaining why and how humans rely on moral concepts. Whether these concepts are employed in cognitive or conative thought is a further issue that will not be addressed in what follows.

The third clarification about the scope of the present inquiry concerns the fact that there is debate about whether and how to distinguish moral cognition (in humans or otherwise) from other forms of cognition.[2] Which concepts should be seen to be moral concepts, and what, exactly, makes it the case that *these* are the moral concepts?

However, it is not necessary here to argue for a precise characterization of moral concepts; indeed, doing so would go counter to the fact that one aspect of human diversity in moral cognition concerns precisely the extent of the set of moral concepts. Is [PATRIOTISM] a moral concept? What about [HONOR]? There is no universal agreement on this point (Haidt & Graham, 2007; Suhler & Churchland, 2011).

For present purposes, it is enough to note that there are several concepts on which there is wide agreement that they are moral concepts and which are important parts of the thought processes

[2] A related question in this context is whether moral thought encompasses prudential thought or whether it is separate from the latter (see, e.g., Dorsey, 2016). This, too, is not central in what follows and will therefore not be further discussed here.

of many humans. Among these concepts are [JUSTICE], [MORAL OBLIGATION], and [VIRTUE]. These are concepts that feature heavily in the thought processes of many humans—a look at a newspaper or many parts of world literature is enough to confirm this. They are also widely recognized to refer to paradigmatically "moral" considerations: these are concepts that set out what we ought to do or be, or what the world ought to be like. It is important to stress, though, that nothing in what follows depends on a precise delineation of the set of moral concepts—or indeed the claim that a precise such delineation even exists (Machery & Mallon, 2010). What matters here is to explain how humans, specifically, evolved to rely on concepts *like* [VIRTUE], [JUSTICE], and [DUTY]; what these concepts have in common with each other (if anything) and exactly which other concepts should be seen as being importantly similar to these can be left open here. Two further points about the nature of "moral" concepts (whatever exactly these turn out to be) are noteworthy here, though.[3]

On the one hand, these concepts are typically *abstract*. [JUSTICE], [MORAL OBLIGATION], and [VIRTUE] famously have no clear correspondence to particular observable circumstances. Whether a particular distribution of items is *just* can depend, among other things, on historical events (what past distributions of these items were like), relational questions (what the distribution of other items is like), and the abilities of the users of these items (how easy it is for someone to use the item).

On the other hand, these concepts are typically *complex*. They are built out of or closely related to numerous component concepts: [VIRTUE] is a [[STABLE] [CHARACTER TRAIT] [TO BE EMULATED]] (or something of the sort) (Doris, 1998;

[3] It is sometimes thought that it is distinctive of moral concepts that their authority is independent of social norms, or at least that it is felt to be thus (Campbell & Kumar, 2012, pp. 97–98; Joyce, 2005; Stanford, 2018). However, the extent to which this is true is controversial (Machery & Mallon, 2010; Sarkissian et al., 2011; Stich, 2018). Hence, this will not be central in what follows.

Nussbaum, 1999). As just noted, the exact nature of moral concepts is not so central here, and there is little reason to think that moral concepts can be given clean necessary and sufficient definitions. Still, what matters here is just that moral concepts are typically deeply embedded in a rich conceptual network so that applying one of these concepts tends to involve many other concepts as well—many of which are also very abstract.

The final clarification about the goals of the present inquiry to be noted here centers on the fact that there is much debate about whether moral cognition is or could be objectively true in a way that is comparable to how the claim that the gravitational constant on Earth is 9.8 m/s/s is true. Without a doubt, the question of the truth and nature of moral realism is important, and it is also possible (though debated) that an evolutionary biological perspective can be useful in resolving them (Joyce, 2005; Kahane, 2011; Kitcher, 2011; Street, 2006). However, these issues, too, are independent from the ones at stake in this chapter. Whatever the nature is of moral thought, and whatever its alethic status turns out to be, questions remain about the inter- and intra-specific variation of moral conceptual cognition. It is these questions that are central here.

In short, at stake in the present context is the question of why and how humans, specifically, evolved in such a way as to frequently rely on a distinctive set of complex and abstract mental representations, such as those of [MORAL OBLIGATION], [VIRTUE], [JUSTICE], [RIGHT], [WELL-BEING], and [DESERT]. To understand this question better, it is useful to contrast the situation of humans with that of non-human animals.

As noted in Chapter 2, non-human animals at least arguably have some proto-moral concepts (Flack & de Waal, 2000; Kitcher, 1998). For example, it is, at the very least, not outrageous to claim that some non-human animals have some kind of concept of [FAIRNESS] (de Waal, 1996; Flack & de Waal, 2000; Monsó &

Andrews, 2022; Prétôt & Brosnan, 2015).[4] However even if the existence of proto-concepts in non-human animals is granted, it is uncontroversially true that humans rely *more frequently* on *more* and *more abstract and complex* moral concepts.

Whether or not chimps (say) can be said to have *some* concept of [FAIRNESS], it is clear that the human concept of [FAIRNESS] is *more abstract*. The human concept of [FAIRNESS] can attach to a wide variety of things, many of which are not straightforwardly observable. In humans, [FAIRNESS] does not just concern whether a conspecific is getting the same food reward for the same amount of physical labor (Brosnan & Beran, 2009; Prétôt & Brosnan, 2015) but can extend to the division of parental effort, access to education and positions of power, and the distribution of nonfungible resources like social capital (see, e.g., Schouten, 2019). The same is true for other moral concepts like [DUTY] or [VIRTUE]: whether person X is *virtuous* is not typically reducible to whether X does an easily observable action A but generally depends on things like the specific set of character traits of X and intricate details of the situation X is in (such as the degree of expected harm it posed to X), etc.

Human moral concepts can also involve a larger number of other concepts. As noted earlier, whether something falls under the concept of [FAIRNESS] might depend on whether it falls under a wide set of other concepts, many of which are further composed out of other concepts: [[HISTORICAL] [EFFORT]], [CHARACTER TRAIT], [[BEHAVIORAL] [EXPECTATION]], to name but a few (Rawls, 1999; Schouten, 2019). By contrast, non-human moral concepts seem simpler and are composed of fewer and simpler concepts (Brosnan et al., 2007; DeAngelo & Brosnan, 2013)

Relatedly, it is also clear that, among humans, *more* moral concepts are used. It may be the case that chimps have a basic

[4] However, some of these findings may also be interpreted in a more deflationary manner as speaking to learned lower-level dispositions only (Machery & Mallon, 2010). As noted in the text, though, settling this is not necessary in the present context.

concept of [FAIRNESS], and it is also plausible that they have a basic concept of [WELL-BEING] (Brosnan & Beran, 2009; Piccinini & Schulz, 2019). However, as far as we know, it does not look like they have concepts such as [RIGHT], [DUTY], [LIBERTY], [VIRTUE], and [MORAL RESPONSIBILITY]. These concepts are a standard part of the human moral vocabulary, though (at least in several cultures) (Henrich, 2020).

Finally, humans rely on moral concepts more frequently than do non-human animals. Chimps may get frustrated if they get unfairly rewarded for their work (Brosnan & Beran, 2009), and they may be motivated to do things that are good for their offspring's well-being (Piccinini & Schulz, 2019). However, their representational moral thought does not extend much beyond this. In the rest of their lives, these organisms do not seem to rely on moral concepts much (if at all). However, humans go far beyond this. Nelson Mandela, Martin Luther King Jr., Mahatma Gandhi, Rosa Parks, Mother Theresa, and Sophie Scholl—to mention just a few—organized virtually their whole lives around moral concepts. They assessed everything around themselves in terms of whether it is *right*, *fair, just, deserved*, what their *moral responsibilities* and *duties* are, etc. They then made these assessments the foundation of their actions. More generally, too, much of human thought and action plays out in a nexus of moral concepts—what we ought to do or be is a constant element of our practical deliberations.

However, while it is thus clear that humans are standouts as far as their reliance on abstract and complex moral concepts is concerned, it also needs to be noted that the details of how humans rely on moral concepts differs across cultures. So, Haidt and Graham (2007) have argued that there are moral differences among political subcultures in the United States, with some people having a wider set of moral concepts than others (e.g., comprising also such concepts [PATRIOTISM] and [HONOR]). More generally, some cultures appeal to moral concepts like [LIBERTY] and [RIGHT] frequently, while others emphasize [DUTY], [VIRTUE], and [OBLIGATION] (Henrich, 2020; Nisbett et al., 2001.

On top of this, there are differences in the nature of how these concepts are used. In some cultures, [DUTY] extends to animals, plants, or ancestors; in others, it is restricted to living humans. Indeed, much intercultural moral conflict centers around questions such as who has which *rights*, *duties*, and *responsibilities* (Darby, 2003, 2004; Donnelly, 1999). This moral divergence needs to somehow be taken into account when making sense of human moral conceptual thought.[5]

Summing up: at stake here is to explain why humans frequently rely on a large set of very abstract and complex concepts, including ones like [VIRTUE], [DESERT], [JUSTICE], [DUTY]. This is something that marks them as different from non-human animals, but is also something that shows considerable intraspecific variation.

6.2 Explaining the Evolution of Human Moral Conceptual Thought: Some Background Considerations

There are several major ways to explain the evolution of human moral conceptual thought and its absence in non-human animals. These span the spectrum from fully nonselective to fully selective accounts. A brief overview of these accounts is useful here to get a better understanding of the theoretical lay of the land. This is especially so since the issues here are slightly different from the dialectic as laid out in Chapters 4 and 5: since the subject matter

[5] As noted earlier (see Note 1), some authors have—controversially—argued that underneath this moral divergence is a shared moral core (Cushman et al., 2006; Mikhail, 2011). However, as also noted earlier, even if this does turn out to be true, the account below is only minimally affected because what drives the superficial differences in the *expression* of the shared moral core would still need to be explained.

is moral cognition—where, as just noted, questions of truth and precision do not apply in quite the same way as in other contexts, such as causal or social cognition—the debate here has not turned around the poles of nativism, empiricism, and symbol processing in quite the same ways as was true, for example, when it comes to mindreading. (However, as will become clearer momentarily, there are many shared concerns here still.)

First, on the fully nonselective side, it is possible that conceptual moral thought is nonadaptive and that it just *happens* to have evolved in humans. That is, there may not be any biological advantage in employing moral concepts when making decisions about what to do; however, for one reason or another, humans ended up doing so anyway. For example, humans may have gone through a period where they had a very low effective population size; population bottlenecks like this make chance a significant determinant of the evolutionary trajectory of the lineage (Amos & Hoffman, 2010; Hawks et al., 2000). Because of this, it is possible that a variety of accidental factors—phenotypic drift, genetic drift, founder effects—led to the evolution of conceptual moral thought in our organismic lineage when it did not in other, even closely related ones (like chimps).

However, this is not in fact a greatly plausible hypothesis. The chance-based explanation of the frequent tokening of *one* moral concept may not be entirely unreasonable. However, the chance-based explanation of the frequent reliance on *many complex* and *abstract* moral concepts is not. It is quite unlikely that, purely by chance, humans ended up frequently relying on a concept of [VIRTUE] *and* on one of [JUSTICE] *and* on one of [DUTY], and so on. To put this in a different way: human moral conceptual thought is best seen as a complex trait with many different interacting elements. It is a web of interlocking mental representations. This matters, because it is a widely accepted principle that purely chance-based explanations are not plausible for traits like this (Godfrey-Smith, 2001; Schulz, 2018b; Sterelny, 2003, 2021). The odds that all the parts of an interconnected set of traits independently evolved are not high.

A more plausible hypothesis therefore is in the idea that the reliance on the rich set of moral concepts of humans was specifically selected for in that lineage. Several different versions of this kind of hypothesis have been put forward in the literature (for overviews, see, e.g., Cosmides et al., 2018; de Waal, 1996; Machery & Mallon, 2010), but a particularly well spelled-out one is in Kumar and Campbell (2022). According to the latter, human moral conceptual thought—what they call "moral reasoning"—is needed to deal with the fact that the application of social norms always leaves ambiguities that need to be resolved. (It is not fully clear where they think these moral concepts orginate or to what extent they think that social norms might themselves focus on moral concepts, though they do consider the evolution of moral norms: Kumar & Campbell, 2022, chap. 3. These are issues to which I return in the next section.)[6]

In the background of this account is the fact that humans have evolved a norm psychology (Boyd & Richerson, 2005; Boyd et al., 2011; Cosmides et al., 2018; Cosmides & Tooby, 1992; Machery & Mallon, 2010; Sripada & Stich, 2006). We have a deeply seated propensity to learn about, adopt, and enforce the norms that are prevalent in our culture and to follow these norms (generally). Note that this point cross-cuts the modular and cultural learning–based accounts from Chapter 3: it is possible that this norm psychology is a selected-for module (Carruthers, 2006; Cosmides et al., 2018), but it could also be just an aspect of processes of cultural learning (Heyes, 2018; Sterelny, 2003). Fortunately, settling this is not so important here (see also Andrews et al., 2024; Westra & Andrews,

[6] Street (2006) also argues that there was selection for moralistic behavioral dispositions. However, since her account is less cognitivist and more behavioral, it does not match the focus of this book as well as that of Kumar and Campbell (2022). Her account also does not target specifically the evolution of moral concepts. For these reasons, the account of Kumar and Campbell (2022) is a better starting place here.

2022). What matters at the moment is just that it is plausible to see humans as having—somehow—acquired the propensity to follow norms: for now, the existence of the propensity is central, but its exact nature and origin can be left open.

What does matter is that following norms is helpful for stabilizing cooperative and other pro-social behaviors that would otherwise be unlikely to be maintained because they counter an individual's narrow best interest (see also Andrews et al., 2024). Importantly, norms can do this with a high degree of resolution, both temporally and spatially: as noted throughout this book, they can respond to the locally prevailing conditions without waiting for a purely "biological" adaptation to occur (which would frequently be impossible).[7]

For example, assume a certain culture is subject to a tragedy of the commons: it is dependent on a limited fish stock that regularly needs to be given time to replenish. In such a case, a norm that says "stick to your fishing allotment and punish those who go beyond theirs" might be culturally adaptive—even though individuals might prefer to exceed their allotment (as long as others do not exceed theirs!). Those cultural groups that follow this norm are more likely to survive and grow than those that do not. Note that this need not be the case in other cultures: humans dependent on practically unlimited stocks of land or ones who always need to hunt *together* might not face such a tragedy of the commons and thus might not find the above norm culturally adaptive (Sterelny, 2021).

However, reliance on such norms does not, by itself, fully suffice to help people navigate their social environments. These norms will sometimes need to be applied to new situations, and different such norms can be in conflict with each other (Kumar & Campbell, 2022, pp. 115–118). For example, in a culture with the above of

[7] As noted in the previous chapter as well, cultural phenomena are of course also part of biology. The contrast here, though, is just meant to be between purely genetic evolution and gene–culture coevolution.

norm of sticking to one's fishing allotment and punishing those who do not, what is one to do where people exceed their allotment of *farmland*? Is farming sufficiently like fishing to have the norm apply? What about when it comes to those spending more than their allotted time *swimming for leisure*? Similarly, assume that this culture also has a norm to help members of one's matrilineal family line. In that case, what happens if members of one's matrilineal family line start to exceed their allotment of fish? Which norm gets priority here?

To resolve these kinds of conflicts, Kumar and Campbell (2022, chap. 5) suggest we need to engage in moral reasoning. To put this in the framework of this chapter, the frequent reliance on abstract and complex moral concepts like [JUSTICE], [EQUALITY], and [DUTY] could be seen to be the result of the fact that it enables us to resolve conflicts among social norms and extend these norms to new cases.[8] By relying on such concepts, we can generalize from specific cases, derive abstract principles underlying different social norms, and then rank-order these principles. In turn, this allows us to extend these norms to new situations and resolve conflicts among them (Kumar & Campbell, 2022). In this way, Kumar and Campbell's picture combines the existence of—innate or culturally learned—moral representations and behavioral norms in an inferential, reasoning-based framework.

Now, Kumar and Campbell (2022) are not explicit in noting exactly what benefits "moral reasoning" has compared to other ways of extending norms or resolving conflicts among them. This will become important again later. However, for now, this point can be left open. What matters here is just that the basic idea behind their account seems to be that, by engaging in "moral reasoning" (i.e., by

[8] To be clear, this is an interpretation of the argument in Kumar and Campbell (2022)—they do not phrase the issues in these terms. However, it appears to be a reasonable interpretation. At any rate, as will be made clearer below, nothing hangs on the exact details of their argument here.

employing moral concepts) we can determine, for example, that underlying the norm "stick to your fishing allotment and punish those that go beyond theirs" is the principle that *[FAIRNESS] requires that everyone (in the group) gets [EQUAL] access to highly limited and slowly replenishing goods*. Hence, we are able to determine that we should punish those who farm more land than they were given but not those who swim for longer than they were allowed. We might also note that norms concerning access to highly limited and slowly replenishing goods need to be given priority to a norm of helping members of one's matrilineal family line because, otherwise, the purpose underlying the former norm is undermined. If everybody made an exception for their family members, it would not be possible to provide everyone (in the group) with [EQUAL] access to highly limited and slowly replenishing goods (see also O'Neill, 2004).[9]

In this way, the frequent reliance on the set of abstract and complex moral concepts of humans is argued to be the effect of the (bio)cultural advantage that this reliance brings with it. This reliance interacts with the (culturally or genetically) evolved norm psychology of humans and enables humans to extend their social norms to new situations and to resolve conflicts among these norms. This is important because without these extensions and norm-conflict adjudication methods, a rich normative psychology cannot get off the ground. To truly be able to enforce pro-social behaviors, norms have to be extendable and rank-ordered; otherwise, their application would be impossible. No two situations are exactly alike, and there is always more than one norm that is potentially applicable to any situation.

Now, there is no question that this account of the evolution of moral conceptual thought in humans has much to recommend

[9] Incidentally, this is an idea that has been much more explored by the classical Western political philosophers: see, e.g., Kant, 1795 [1996]; Locke, 1689 [1983]).

it. Of particular importance here is the fact that it is grounded in two widely accepted facts: the importance of cultural (or gene–cultural) (co-)evolution and that of normative thinking in human evolution (Boyd, 2018; Chudek et al., 2013; Henrich, 2015; Sterelny, 2021; Tomasello, 1999). This gives the account of Kumar and Campbell (2022) a solid theoretical foundation and ensures it goes beyond pure speculation and "just-so" storytelling (Machery & Mallon, 2010). However, that said, the account also has three major shortcomings.

First, as noted earlier, there are several gaps in the account. Exactly what is the relationship between moral reasoning and social norms? Exactly what benefits does "moral reasoning" bring that could not be obtained otherwise? Where are the required moral concepts coming from? Why, exactly, do humans have such an extensive norm psychology? Without answers to these questions, the account of Kumar and Campbell (2022) lacks cogency.

Second, the account underplays the situation of non-human animals. Exactly why did the factors identified by Kumar and Campbell (2022) as underlying the evolution of moral conceptual thought in humans not apply to non-human animals as well? What explains the fact that non-human animals did not find the reliance on many complex and abstract moral concepts adaptive? As acknowledged by Kumar and Campbell (2022, p. 266, note 56), this question cannot be brushed aside by claiming that non-humans are not norm-driven. While the interpretation of the behavioral data here is not yet fully clear (Schlingloff & Moore, 2017), it cannot be ruled out that at least some non-human animals display at least a proto-normativity (Andrews et al., 2024; Monsó & Andrews, 2022; Westra & Andrews, 2022). If so, though, then these organisms, too, would seem to profit from rich moral concepts that allow them to extend norms to new situations and resolve conflicts among norms: if norms stabilize cooperation in humans, and if the reliance on moral concepts is needed for norms to play this role, then it is not clear why norm-using non-human animals did not evolve a reliance on these concepts as well.[10] Until it has been established that normative thinking in non-human animals does not exist, this point thus needs to be taken seriously.[11]

In response, it may be noted that it is not unreasonable to think that non-human animals do rely on moral concepts, too (Andrews, 2015; Andrews et al., 2024; Monsó & Andrews, 2022; Prétôt & Brosnan, 2015; Westra & Andrews, 2022). However, this response does not help here: as noted earlier, the issue at stake in the present context is explaining the evolution of specifically human moral conceptual thought (Delton & Krasnow, 2015). If the logic of the above account of the evolution of moral conceptual thought applies equally to human and non-human animals, it fails in isolating what is unique about humans—which is exactly what matters here. Hence, appealing to the existence of (proto-)moral thought in non-human animals is not in fact helpful in this context.

A better answer to the question of why non-human animals did not evolve a frequent reliance on abstract and complex moral concepts may therefore be thought to lie in the fact that non-human animals have more limited representational abilities *in general*. For example, non-human animals generally do not rely on complex and abstract *scientific and mathematical* concepts (such as [SQUARE ROOT], [NORMAL DISTRIBUTION], [COMMON CAUSE]) either. So perhaps the reason why non-human animals did not evolve a frequent reliance on abstract and complex moral concepts just lies in the fact that their representational abilities were too limited to make this feasible.

However, while this response does get at something important, it, too, is not as straightforward as it might at first seem. As also noted in Chapters 2 and 3, the abstractness and complexity of the

[10] Of course, the answer may also be that this is purely due to chance. However, this would then reduce the present purely selectionist account back to the above, purely chancy nonselectionist one—with all the problems that comes with it.

[11] Even if it has been established, the account of Kumar and Campbell (2022) still fails to be fully compelling due to the first and third problems noted below.

mental representations that non-human animals rely on are not necessarily always highly limited. In some regards, non-human concepts can be more complex than human concepts (Andrews, 2015). Consider for example a tiger's pheromonally based set of representations of its environment: this is temporally and spatially complex in ways that may well match or exceed many human concepts. Now, as will be made clearer in the next section, there often *are* reasons to think that human representational cognition is less limited than that of non-human animals. However, the key point to note at this point is that this is something that needs to be *argued for* on a case-by-case basis. It cannot just be taken for granted across the board.

The third problem with the above account of the evolution of moral conceptual cognition is—the by now hopefully familiar one—that it is not made clear exactly how *humans* are able to employ moral concepts to extend norms and resolve conflicts among them. As noted in Chapters 4 and 5, the frequent reliance on many abstract and complex concepts—moral or otherwise—is no easy task because it comes with major resource requirements in terms of time, attention, concentration, and memory (Lieder & Griffiths, 2020; Sterelny, 2021). The use of concepts like [VIRTUE], [DUTY], and [JUSTICE] to extend existing moral norms and rank order them thus is far from straightforward. There are many ways to do so, and the cogency and consistency of these extensions and rank orders are far from obvious (see, e.g., Rawls, 1999, for a classic example of the difficulty of this kind of project). Hence, to rely on such concepts, and to do so frequently, requires much concentration, attention, and memory: we may need to keep track of historical contingencies, distributional factors, and many other features of the environment to successfully determine whether a given situation truly instantiates a general principle like "ensure everyone has equal opportunities to own needed goods and services." For even moderately large societies, this is highly nontrivial (Sterelny, 2021). Finally, the frequent reliance on abstract and complex moral concepts also comes with great computational requirements. To determine whether a given *property arrangement* is *just*, we may need to combine information about the *historical distribution of property* with its *transmission over time* and its *current value*. This sort of inference can be very difficult and require the integration of a large number of disparate pieces of information.

Given this, any account of the evolution of moral conceptual thought—that of Kumar and Campbell (2022) included—needs to provide an answer to the question of how it is that humans were able to rely on such concepts in first place. Otherwise, it falls short of truly explaining this evolution: it leaves a key element of the explanation open.

A promising way to resolve these shortcomings is to complicate the dimensionality of Kumar and Campbell's (2022) account by adding sociotechnological development to its core dynamic. This introduces some nonselectionist elements as well, but without the brute randomness of the pure nonselectionist account. In this way, it becomes possible provide a more compelling picture of the evolution of human moral conceptual thought, its absence in non-human animals, and the variance within human moral conceptual thought. The next two sections make clearer how this can be done.

6.3 The Bio-Social-Technological Evolution of Human Moral Concepts

To understand the evolution, in the human lineage, of the frequent reliance on many highly abstract and complex moral concepts, it is best to start where Kumar and Campbell (2022) also start: the particular socionormative environment in which humans evolved. However, it is now important to fill in the lacunae their account leaves open. Specifically, it needs to be made clear exactly why the employment of many highly abstract and complex moral concepts is so useful in this kind of environment, where these concepts are coming from, and why human normative thinking is so extensive.

As noted throughout the book, it is highly plausible that human cognitive evolution was deeply influenced by the fact that humans live in complex cultural and social environments. Humans are not like the famous Hawaiian *humuhumunukunukuapuaa*: we do not enter the world ready to make our own way through it. Rather, we are dependent on others. However, unlike some other social species (honeybees, say), we live in groups that involve both kin and non-kin and many shifting alliances. In such a scenario, it is useful to use norms to guide our behaviors: this can stabilize cooperative behaviors in the face of a temptation to defect, clarify individual responsibilities, and thus make available high-fitness equilibria that would otherwise be out of reach (Boyd, 2018; Chudek et al., 2013; Henrich, 2015; Kitcher, 1998; Sterelny, 2021, pp. 67–68, 82–87; Tomasello, 1999).[12]

Importantly, to be functional in these ways, these norms may need to contain moral concepts such as [VIRTUE] or [DESERT]. For example, due to the complexity of the relevant decision-making process, stabilizing cooperation might require punishing non-cooperators, but still make some allowances for accidental non-cooperation (see, e.g., Alexander, 2007; Skyrms, 1996, 2004). If so, it may be adaptive to rely on norms like *punish vicious non-cooperators* or *help the deserving*. Now, whether this is so will depend on the details of the situation: how complex the decision-making processes are and how much leeway non-cooperators should (adaptively) be given. This thus does not provide a universal or general account of why humans came to frequently rely on abstract and complex moral concepts. Still, it points to *one reason* why this reliance evolved (see also Fraser, 2018).

[12] Curry (2016) moves directly from the adaptiveness of cooperation to (proto-) morality, as do Tooby and Cosmides (2010) and Flack and de Waal (2000). However, as noted in the text, the exact relationship between morality and normativity is left open here.

A more general such reason lies in the fact—correctly noted by Kumar and Campbell (2022, chap. 5)—that, as humans interact with their environment, they will encounter situations to which existing social norms do not apply or where different norms have conflicting applications. As noted earlier, from this, Kumar and Campbell (2022) jump to the conclusion that moral reasoning is (bio)culturally adaptive, but they fail to make clear exactly why that is. It is now time to provide this explanation.

To do this, it is important to begin by noting that norms change, come into existence, and go out of existence for many different reasons (Nichols, 2004; Sterelny, 2021, pp. 84–85). However, there is no reason to think that a situationally appropriate norm (i.e., one that encompasses the present situation in a bioculturally adaptive manner) will arise quickly or at all whenever needed. Hence, if norm extension and resolution is just based on this kind of cultural evolutionary process, humans mighty need to wait a long time to be able to re-obtain the high fitness equilibrium normative social living makes possible.

However, this is exactly where the use of abstract and complex moral concepts becomes useful. By using concepts like [FAIR], [JUST], [DUTY], and [VIRTUE], humans are able to adjust to a new situation *immediately*. Indeed, it is precisely in virtue of their abstractness and complexity that these concepts can subsume many different, more specific behavioral norms under one heading and extend norms to new situations. These concepts are not tied to specific observable circumstances—they are abstract—and they are built out of separable components—they are complex. This allows these concepts to extend and rank existing norms of behavior and thus provides humans the means to tailor their behavior to novel situations or to situations with conflicting norms more quickly than waiting for the undirected appearance of a novel norm.[13]

[13] DeScioli and Kurzban (2013) also provide an account of morality as solving a coordination problem, but they do not focus directly on moral concepts and the details of their account are also quite different from the argument here presented.

For example, in the above case where a group of humans faces the question of how to regulate access to farmland, humans without the ability to use abstract and complex moral concepts need to wait for a new norm to arise that replaces the existing norm of "stick to your fishing allotment and punish those who go beyond theirs," such as one that says "stick to your fishing and farmland allotment and punish those who go beyond theirs." This may take months or years and thus deprive the members of that group of many bioculturally adaptive cooperation opportunities during that time (Sterelny, 2021). The same goes for a situation in which a clan leader dies in novel circumstances and where the succession is ambiguous due to lack of an appropriate norm: maybe the leader had children with equal rights to the position and had also failed to designate an heir apparent at the time of their death. In such cases, the society may be left in danger of internal conflict and short of its cooperative potential for a considerable period of time—namely, until an appropriate norm emerges that can resolve this situation.

However, humans with the ability to use abstract and complex moral concepts can note that *fairness* requires sticking to allotments of finite and slowly replenishing resources and that the farmland case should thus mirror that of fishing. Similarly, such humans can appeal to a concept of [VIRTUE] and note that the goal in cases of a leaderless society to identify a *virtuous person* to be the next leader. They can then use this idea to note that person X should become the new leader since they have the appropriate *character traits* of *courage*, *temperance*, and *wit*. In this way, the use of abstract and complex moral concepts can, in turn, be seen to be favored by gene–cultural evolutionary pressures: it makes norm adjustment faster and more efficient and thus enables humans to more reliable achieve high-fitness equilibria in social interactions.[14]

[14] Birch (2021) argues that the reliance on evaluative "normative" concepts may be the result of the importance of craft knowledge and skilled toolmaking. If this right, then this provides an "origin explanation" (Godfrey-Smith, 2009) of these kinds of concepts that is consistent with the account here provided (although Birch does not develop his

Of course, there is generally more than one way of extending a norm and not necessarily one “correct” one. Moral conceptual reasoning need not necessarily land on the moral truth, even assuming there is such a truth. (As noted earlier, the question of the *correctness* of moral claims is left open here.) This means that it is entirely possible for different people to use abstract and complex moral concepts differently and thus infer different resolutions of norm conflicts or different extensions of existing norms. While this may seem to introduce a similar impasse as the initial situation—with a society lacking a cohesive strategy to ensure mutually beneficial cooperative outcomes—this new, moral conceptual conflict is in fact quite different in nature.

In particular, the latter is a straightforward question of the adoption of one out of a set of proposed solutions. It is not a matter of waiting for the *emergence* of an appropriate solution; rather, it is a matter of *deciding* which solution to adopt. To be sure, such a decision can still be fraught with conflict and practical difficulties. However, the point here is just that moral concepts make the quick and reliable extension and ranking of social norms *possible*: without them, norm extension and resolution might only occur slowly or not all (if no suitable norm emerges before the culture collapses). It is true that moral concepts do not guarantee that a particular norm extension or resolution will be adopted and that the use of concepts might need to be further supported to yield sufficiently wide acceptance; this is an important issue to which I return momentarily. For now, though, the key point to note is that abstract and complex moral concepts pave the ground for quick and reliable norm extensions and resolutions in ways that nonconceptual solutions do not: they can function as *coordinating devices* for adjustments to (perhaps quite simple)

account in the way that is done here): it shows where these concepts first came from. As noted below, the present account makes clear why normative concepts are bio-culturally adaptive once they became available to be used.

normative rules (Andrews, 2020; Andrews et al., 2024; Westra & Andrews, 2022).

Where, though, do these abstract and complex moral concepts come from? There are several different answers that are still on the table. Some of the concepts may be innate (see, e.g., Mikhail, 2011). It is highly plausible, though, that many or most of these concepts are culturally learned: for example, we may learn from others not just that *justice requires keeping one's promises*—but also the very concepts of [JUSTICE] and [PROMISE]. In either case, though, it is reasonable to think that humans can acquire a set of moral concepts that enable them to organize and revise behavioral norms.

So far, though, this is only half the story of the evolution of human moral cognition. It is now necessary to address the worry raised earlier that the frequent employment of many highly abstract and complex moral concepts is cognitively very demanding. As noted in the previous section, it needs to be explained how humans managed to defray these costs—otherwise, the fact that the reliance on abstract and complex moral concepts would be adaptive is not sufficient to provide a reason for its evolution (Orzack & Sober, 1994; Sober & Wilson, 1998). The key to this explanation lies, once again, in the fact that humans evolved in environments that not only favored some kind of normative thought—something that is shared with non-human animals as well—but that also provided the resources to *use* many highly abstract and complex concepts. These resources consist in cognitive and social tools.[15]

More specifically, it is not just that the richly cultured social environments humans evolved in rewarded sophisticated psychological and physiological abilities to navigate *through* them, such as a norm psychology. It is also the case that these environments made it possible to *expand* the psychological and physiological

[15] The importance of cognitive and social tools to human cognitive evolution is acknowledged by Kumar and Campbell (2022, p. 107, chaps. 3 and 7), but they do not develop this acknowledgment further.

abilities humans have. The major mechanism underlying this expansion is again cultural learning: humans can learn from others about locally adaptive ways of acting (Boyd & Richerson, 2005; Csibra & Gergely, 2011; Henrich, 2015; Heyes, 2018). In turn, this makes it possible to develop sophisticated technology that it would be impossible for humans to build from scratch in one generation. The cultural evolution of this kind of *cognitive technology* and of *social institutions* shares some aspects with that in the case of mindreading, but there are some differences as well.

As before, *cognitive technology* such as quipus, sundials, calendars, abacuses, petroglyphs, as well as algebra, accounting conventions, and a written language allows humans to dispel the costs of the frequent use of many highly abstract and complex concepts—and thus make this use possible (see also Sterelny, 2021, pp. 45–46). For example, although the details of their use remain unknown, it is clear that quipus allow people to track who acquired which goods when (Urton, 2018). These are thus examples of technology of type S from Chapter 4. There are also examples of technology of type I: sundials and calendars allow us track time (Henrich, 2020, pp. 360–367). Abacuses make many kinds of inferences—all those that can be formalized in arithmetic—easier (Basu & Waymire, 2006). Petroglyphs and written languages allow the storage of much information for long periods of time as well as aiding some representational inferences (Kelly, 2015; Sterelny, 2012a). In turn, this matters here because exactly such inferences are needed to extend norms to new situations and adjudicate conflicts among them (see also Sijilmassi et al., forthcoming).

Cultural learning also allows for the successive refinement of *social institutions*. As noted in the previous chapter, social institutions are important in many different ways in human evolution and cognition (Henrich, 2020; Sterelny, 2021).[16] However, in the present

[16] Note that while Kumar and Campbell (2022, chap. 7) also emphasize the importance of social institutions for the evolution of human morality, they do not consider or discuss the issues raised here.

context, the key aspect of them that needs to be emphasized is that they, too, can lead to a lowering of the cognitive costs of frequently using highly abstract and complex concepts. For example, the existence of an institution governing fair exchanges of property (such as a well-functioning market) in addition to one concerning the fair creation of novel property (such as a well-functioning patent system)—two issues this book returns to in Chapters 7 and 8, respectively—can make it much easier to determine if an existing property assignment is fair. If I know that the existing assignment is the result of these kinds of institutions, then I don't need to track the starting state of the assignment (who owned what at first) or its intermediate states (who gave what to who when). Instead, knowing that the initial assignments and transition dynamics—whatever exactly they were—followed the appropriate rules is enough. With a well-functioning process, I can let *it* do the work, and I do not need to track its operation.

It is important to emphasize that not all forms of cognitive technology or all institutions can play these roles. For example, *outside chalkboards* might not make sufficiently good transaction ledgers—and thus may be unable to lower the memory and computational costs of using concepts like [FAIRNESS] and [EQUALITY]—if they are subject to frequent erasing due to rain or other external influences. Similarly, if markets do not always yield legitimate property transfers (e.g., due to exploitation or power differentials [Liberto, 2014]) or if property generation is not fair (e.g., due to differential starting places of different people [Otsuka, 2003]), then such an institution may not suffice to lessen the costs of using abstract and complex moral concepts like [FAIRNESS] and [EQUALITY].

The point here is just that, through cultural learning, it is possible for humans to refine cognitive and social tools over time to create ones that can, in fact, make the use of abstract and complex moral concepts feasible. The more such tools are available, and the more functional they are, the more feasible the frequent use of abstract and complex moral concepts becomes. To be able to often rely on complex and abstract moral concepts, it is crucial to have the ability to lower some of the costs associated with this. This kind of reliance requires much concentration, attention, memory, and computational power. The fact that humans can culturally learn to build on and improve existing cognitive technology and social institutions is thus crucial as it can make exactly this possible—though this is not guaranteed. Importantly, it is also important to note that cognitive technology and social institutions are very old and have, in some form, been part of human evolutionary history since its inception (Sterelny, 2017, 2021).

This also returns to the point made earlier that moral conceptual reasoning need not yield one universally accepted outcome: even within the same culture, different people can employ the same moral concepts (with slightly different contents perhaps) to arrive at different norm resolutions and extensions. For example, some people might think that *virtue* entails being witty, others might deny this. As noted earlier, this can yield conflict over how to resolve existing norm conflicts or how to extend existing norms to new situations. The existence of the right kinds of cognitive technology and social institutions can help with this kind of conflict, too. So, market solutions can also be applied to decisions over which norm extensions or resolutions to adopt: betting markets or other voting mechanisms allow not just for the determination of whether a given norm is *fair*, but they can also be used to determine which of a set of competing assessments of the [FAIRNESS] of a norm we should adopt. Similarly, written language and abacuses can make the steps of moral arguments clearer for others to follow and thus allow others to be convinced more easily of a given position. In this way, cognitive technology and social institutions—assuming they are of the right kind—not only make the *application* of many complex and abstract moral concepts possible, but they also make it possible to decide which such applications to *adopt*.

6.4 The Biocultural-Technological Feedback Loop Surrounding the Evolution of Human Moral Concepts

Importantly, though, this again does not exhaust the biocultural-technological dynamics surrounding the evolution of human moral conceptual thought. It is not just the case that the cultural environment that humans develop in allows for the construction of the kind of cognitive and other social tools that make the frequent use of many highly abstract and complex moral concepts possible. It is also the case that the frequent use of many highly abstract and complex moral concepts impacts the nature of cultural learning and normative thinking—and thus, technological and institutional development. This creates a positive feedback loop that further deepens human moral conceptual thought.[17]

In particular, the use of moral concepts like [[VIRTUOUS] [AGENT]] allows us to fine-tune who to learn from. We are not restricted to learning from neighbors, parents, or all peers. We have many new ways to bestow certain individuals the status of "learning models." Instead of learning from all elders, for example, we can designate certain elders—the *virtuous*, say—as particularly good sources of information. This can have drastic effects on the speed of cultural evolution (Boyd & Richerson, 2005; Henrich & McElreath, 2011; Muthukrishna & Henrich, 2016; Sterelny, 2021). By making it possible to focus the learning on those most advantageous to learn from, the learning becomes more efficient and less likely get stuck in local, suboptimal equilibria. In turn, this can make cultural evolution faster and more efficient (Henrich, 2020; Muthukrishna & Henrich, 2016). Again, the claim is not that this

[17] Kumar and Campbell (2022, pp. 129, 137) also hint at the existence of a feedback loop when it comes to the evolution of human morality. However, they do not consider cognitive technology at all, and they do not develop this hint in much detail. Sterelny (2021, pp. 26–27) notes the importance of feedback loops for human cultural evolution more generally but also does not develop this point in detail in the way it is done here.

is necessarily so—learning can also end up just elevating an elite (Sterelny, 2021)—but that it *may* do so.[18]

Similarly, it is not just the case that cognitive technology and social institutions allow us to frequently use many highly abstract and complex moral concepts—the latter can also impact the former. Our moral concepts can lead us invent cognitive technologies that we would otherwise not have invented. For example, a commitment to *social equality* led to the development of new ways of paying (Hayashi, 2016; Toh, 2021), a commitment to the *morality of protecting the environment* led to the development of renewable energy generators, and a commitment to *equity* and *fairness* might cause us to refine our market institutions (Jaffe & Lerner, 2004).[19] In turn, the new technologies and social institutions can then further help us use many highly abstract and complex moral concepts as described earlier: the development of financial technologies expands the reach and functionality of markets and can thus allow the latter to underwrite more uses of concepts such as [FAIRNESS], and the development of renewable energy generators can make it easier to determine whether a given consumption pattern is [VIRTUOUS].

Putting this all together thus leads to the kind of feedback loop seen in Figure 6.1:

As the figure shows, cultural learning allows the construction of cognitive and social tools that can support the frequent use of many abstract and complex moral concepts. In turn, the use of these concepts can change the nature of the available cognitive and social tools directly and indirectly, through their impact on the processes

[18] Sterelny (2021, p. 28) emphasizes the usefulness of abstract concepts in general for supporting individual and social learning.

[19] Many of our deadliest weapons—nuclear bombs being an obvious example—were the result of researchers trying to build something that is so destructive that it would not need to be used and would therefore end conflicts just through its existence (Carnesale et al., 1983). (It goes without saying that this did not turn out to be true.) However, as the text notes, other examples of morally driven technological and social innovation exist, too.

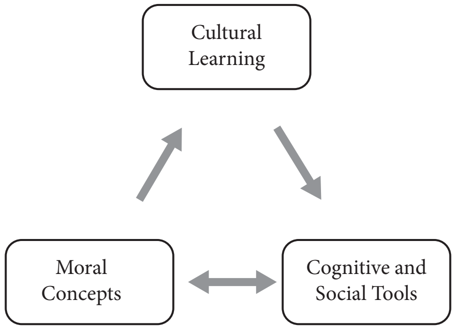

Figure 6.1 An interactionist feedback loop surrounding human moral conceptual thought.

of cultural learning. This generates a mutually reinforcing feedback loop that can lead to increasingly more abstract and complex moral concepts.

6.5 Explaining Intra- and Inter-Specific Variation in Human Moral Conceptual Thought

All of this is important here because it explains the two major questions surrounding human moral conceptual thought. First, there is the question of why the frequent use of many highly abstract and complex moral concepts evolved in the human lineage when it did not among non-human animals.

The answer can now be seen to lie in the nature of the positive feedback loop depicted in Figure 6.1. While there may be relatively few major initial differences between human and non-human animals in their sociality, ability to use tools, normative psychology, and representational abilities, in the human case, these differences crossed a threshold that led to them ratcheting up each other into increasingly bigger differences. Humans were slightly more sophisticated in their cultural learning, which allowed them to build slightly more advanced cognitive and social technology; in turn, this then led to a slightly more extensive norm psychology involving slightly more complex and abstract moral concepts. This kicked off a feedback loop leading to ever greater divergence between human and non-human animals.

Put differently, while the reliance on many highly complex and abstract moral concepts may have been *adaptive* for at least some non-human animals as well, it may not in fact have been *available* to be selected (see also Sober & Wilson, 1998, chap. 10). This introduces a nonselective element into the present account of the evolution of human moral conceptual thought. However, this nonselective element is not of the purely chancy variety that was shown to be implausible in the last section. Rather, the large (maybe even in-kind) difference between human and non-human animals in their moral representational abilities can now be seen as the result of a number of smaller (in-degree) differences that greatly strengthen each other. While non-human animals might have also profited from the frequent reliance on abstract and complex concepts, this reliance was just not feasible for these animals: in the latter case, the above feedback loop did not get started, and they thus were greatly restricted in their moral conceptual cognition.

The feedback loop also answers the second question surrounding human moral conceptual thought: why it is not culturally homogeneous. As noted earlier, both the extent of the set of moral concepts and their extensions differ, often quite drastically, between cultures. Some cultures see [HONOR] as an important moral concept, whereas others do not (Nisbett et al., 2001). Some cultures see [PATRIOTISM] as an important moral concept, whereas others do not (Haidt & Graham, 2007). Some cultures have a concept of [MORAL DUTY] that assigns moral valence to actions depending on the caste, religion, or gender of the agents and patients involved, whereas others do not (Henrich, 2020).

These points also apply historically. As has been noted by various authors, many cultures have seen a widening of the moral circle over time (Kumar & Campbell, 2022, chaps. 8–10; Singer, 2011): at first, only some humans were seen to have moral standing, then all humans, then (some) animals, and perhaps soon artificial intelligences (AIs) as well. Also, it is plausible that older cultures tended to use fewer and simpler moral concepts than their later descendants: medieval Europe was characterized by the widespread use of cruel punishments and tortures, and moral concepts such as [[INDIVIDUAL] [DIGNITY]], [[BODILY] [INTEGRITY]], and [CONSENT] were not used much or at all. This has changed now.

Importantly, the account of the sociocultural-technological evolution of human moral conceptual thought laid out in the previous section can be used to make progress in making sense of these cultural differences. Consider these two cases—historical and cross-cultural differences—in reverse order.

It is a straightforward consequence of the above account that the complexity, abstractness, number, and frequency of use of moral concepts are an increasing function of the power and availability of the right kinds of cognitive technology and social institutions. *Corruption* is hard to make sense of without (a) devices—quipus, algebra, accounting conventions—that allow us to track ownership and (b) the right kinds of social institutions that favor tracking transactions as independent of the nature of the transactant (i.e., as being "impersonal"; Henrich, 2020). Since, as noted in the previous section, it takes time to culturally evolve this kind of technology and these kinds of social institutions, this thus directly predicts an increase in the complexity, abstractness, number, and frequency of use of moral concepts over time.

Importantly, though, this is not guaranteed: as noted earlier, not all social institutions and cognitive technologies are of the right kind to support moral conceptual thought, and it is entirely possible that cultures evolve *away* from the use of such institutions and technologies. Furthermore, since moral conceptual thought can also be self-directed, cultures can *reason* their way to and from moral concepts in complex ways. The point here is just that the account above predicts a *tendency* for moral conceptual thought to increase in complexity, abstractness, number, and frequency of use over time. The historical record is at least consistent with the existence of such a tendency.

On the cross-cultural side, the fact that the frequent use of many complex and abstract moral concepts is the result of a complex feedback loop involving cultural learning and the deployment of cognitive technology and social institutions directly predicts the existence of cultural variation in moral conceptual thought. Importantly, this is not just due to the fact that different cultures may exert different cultural selection pressures on moral norms (Kumar & Campbell, 2022, chap. 3)—though this may be true as well. Rather, the point is again that the kind of feedback loop sketched in Figure 6.1 has no uniform direction in which it moves (see also Sterelny, 2021, pp. 83–84). Even given similar starting points, different end states can be reached. In particular, even if all humans lived in culturally quite similar conditions, it could happen that some cultures developed technologies or institutions that favor the reliance on certain moral concepts rather than on others. In turn, these small differences in moral concepts can then impact the nature of cultural learning in the relevant cultures, leading to ever greater divergences in cognitive technology and social institutions—and thus moral concepts. A good example of this difference is again the reliance on market institutions in some cultures but not others—a point to which I return in the next section. In turn, this makes concepts like [CORRUPTION] or individualized conceptions of [GUILT] prominent in the former but not in the latter cases (Henrich, 2020).[20]

In this way, the account laid out here can help us make progress in our understanding of the intercultural variation of human moral

[20] Haidt and Joseph (2004) argue that underneath this moral diversity is a shared moral intuitive/affective core. As noted earlier, whether this is true is left open here. See Suhler and Churchland (2011) for an alternative view. See also Notes 1 and 5.

conceptual thought. An increasing complexity in the nature of human moral conceptual thought over the time in which humans evolved is (at least partly) accounted for by increases in the technological and social sophistication of human living. Furthermore, since the feedback loop between human moral conceptual thought, cultural learning, technology, and social institutions is nonlinear, the account predicts straightforwardly that there will be much cultural variation in moral concepts.

6.6 Conclusion

Human moral cognition is distinctive in a number of different ways, but a major one of these ways is its frequent reliance on many highly abstract and complex moral concepts. Moreover, this reliance is not culturally homogenous, but instead highly variable both in terms of the concepts that are being employed and in the ways in which a given concept is employed. These facts need an explanation. One way to do so is to appeal to the fact that humans evolved in a special social-normative environment that rewards the frequent use of these kinds of moral representations to resolve conflicts among norms and extend them to new situations (Kumar & Campbell, 2022). However, while on the right track, this view falls short of providing a fully compelling account of these issues. The social-normative environments of some non-human animals (e.g., chimps) are comparable in some regards to those of humans, but their moral representational cognition is not. It is also not clear why we see the kind of historical and cultural variation in human moral conceptual thought that we in fact do see: while all cultures differ slightly, why is it that moral concepts differ so drastically?

To answer these questions, it is necessary to consider the fact that the frequent reliance on many highly abstract and complex moral concepts is cognitively very complex and thus needs to be underwritten by appropriate cognitive technology and social institutions, the development of which is driven by cultural learning. Importantly, furthermore, the employment of moral concepts can shape processes of cultural learning and thus technological and institutional development. This generates a feedback loop in which cultural learning, cognitive technology, social institutions, and moral conceptual thought reinforce each other.

The fact that human moral conceptual thought needs to be underwritten by appropriate cognitive technology and social institutions goes some way toward explaining the evolutionary/historical variation in human moral conceptual thought, and the fact that the feedback loop does not have a fixed aim but is very sensitive to small changes in its different elements goes some ways toward explaining the cultural variation in human moral thought. While there are still many questions open about the nature and evolution of human moral cognition, it is hoped that a clearer understanding of its evolutionary foundations is useful for answering these questions—and for improving the ways in which humans treat each other. With this in mind, consider an aspect of this account in more detail: the existence of trading institutions.

7

Tools of the Trade

The Bio-Cultural-Technological Evolution of the Human Propensity to Trade

Without a doubt, humans are also standouts in their "propensity to truck, barter and exchange" (Smith, 1776; see also Nowell, 2010; Ofek, 2001, 2013): we are at least sometimes disposed to exchange many different goods and services with many different others for the mutual benefit of all involved parties. The fact that humans are standout traders has been known for a while. So, Smith (1776 I, ii) notes that,

> The division of labour, from which so many advantages are derived, is not originally the effect of any human wisdom, which foresees and intends that general opulence to which it gives occasion. It is the necessary, though very slow and gradual consequence of a certain propensity in human nature which has in view no such extensive utility; the propensity to truck, barter, and exchange one thing for another.
>
> Whether this propensity be one of those original principles in human nature of which no further account can be given; or whether, as seems more probable, it be the necessary consequence of the faculties of reason and speech, it belongs not to our present subject to inquire. It is common to all men, and to be found in no other race of animals, which seem to know neither this nor any other species of contracts. . . . Nobody ever saw a dog make a fair and deliberate exchange of one bone for another with another dog. Nobody ever saw one animal by its gestures

It's Only Human. Armin W. Schulz, Oxford University Press. © Oxford University Press 2025.
DOI: 10.1093/9780197800188.003.0007

> and natural cries signify to another, this is mine, that yours; I am willing to give this for that. . . . But man has almost constant occasion for the help of his brethren, and it is in vain for him to expect it from their benevolence only. He will be more likely to prevail if he can interest their self-love in his favour, and show them that it is for their own advantage to do for him what he requires of them. Whoever offers to another a bargain of any kind, proposes to do this. Give me that which I want, and you shall have this which you want, is the meaning of every such offer; and it is in this manner that we obtain from one another the far greater part of those good offices which we stand in need of. It is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard to their own interest. We address ourselves, not to their humanity but to their self-love, and never talk to them of our own necessities but of their advantages.

For present purposes, the key aspect of this passage is that Smith contrasts the ways human and non-human animals go about getting what they want. Whereas the latter simply *fight* over the possession of a scarce resource, humans decide to *barter* over it. Humans as well as dogs (say) sometimes have opposing interests and are not always altruistically inclined toward each other; however, it is distinctive of humans that they often resolve these conflicts through *trade* rather than fighting.[1]

However, in line with the rest of the book, a number of important questions about this human propensity to trade remain open.[2] First, as it stands, it is not clear exactly what makes this propensity so different in the human case from that of other animals. Exactly how do humans differ from other animals in their cognitive and

[1] Human sociocultural evolution also allowed for a greater diversity of preferences, creating more win-win situations and thus making it easier to resolve conflicts through trade. This is an important point to which the chapter returns in Section 7.3.

[2] For an expanded treatment especially of the evolutionary consequences of the propensity to trade, see Ofek (2001). See also Avsar (2020).

motivational abilities when it comes to trade? Second, it is not clear why humans evolved in this way. Why did other animals not acquire the propensity to trade (or at least not to the extent that humans did)? What unique evolutionary pressures in the human lineage led to this divergence?[3] Third, it is not clear what explains the fact that the extent to which humans engage in trade is culturally highly variable—with some cultures (such as contemporary Western society) seeing a lot of trade, and some (such as some contemporary hunter-gatherer cultures like the Hadza) seeing less trade. Given that all humans *could* engage in extensive trade, why don't they all do so? Answering these questions is the goal of the rest of this chapter.

The chapter follows, in outline, the structure of the previous chapters. However, since the issues are slightly different, there is no section directly contrasting the interactionist picture developed here with others in the literature; rather, this contrast is developed throughout the chapter. So, in Section 7.1, the propensity to trade is clarified by showing it to be divisible into a number of component abilities. In Section 7.2, I then contrast and make more precise the extent to which humans differ from non-human animals in these component abilities. In Section 7.3, I provide an account of the interactionist evolutionary pressures that have led humans to diverge from non-human animals in their propensity to trade. I summarize the discussion in Section 7.4.

7.1 The Propensity to Trade

In the passage quoted earlier, Smith (1776) suggests that the core of the distinction between humans and dogs in their propensity to trade lies in the sophistication of their egoism. While both

[3] This evolutionary perspective was, of course, not as such available to Smith (1776); but see also Schliesser (2011).

individual humans and individual dogs are egoistic in the sense of looking to get the best for themselves without an inherent ("ultimate"— Schulz, 2011b; Sober & Wilson, 1998; Stich, 2007) concern for the interests of others, humans recognize that others *have* their own interests and that mutually beneficial outcomes can be achieved through trade rather than all-out fighting.[4]

This is a useful start toward characterizing the propensity to trade, and it does get at several important aspects of this propensity. (Here and in what follows, I use the terms "propensity to truck, barter, and exchange" and "propensity to trade" interchangeably.) However, it also turns out that this characterization needs to be made more precise in several regards.

Indeed, the propensity to trade should be seen to have three key psychological components that underlie it: (1) a concept of [PROPERTY], (2) a concept of [EXCHANGE VALUE], and (3) a disposition and capacity for reciprocal cooperation (Avsar, 2020; Brosnan, 2011; Brosnan & Beran, 2009; Boyer, 2023; Ofek, 2001).[5] Of course, these three psychological components do not fully *exhaust* what is needed to have a propensity to trade. In particular, trading also requires many of the other psychological capacities discussed in this book: communicative abilities (to make clear to others what is to be traded for what), theory of mind abilities (to determine whether others are willing to engage in trade), and abilities to evaluate different possible states of the world (to assess whether trading for a resource is called for in the situation at hand) (Kabadayi & Osvath, 2017; Mulcahy & Call, 2006).

[4] Note that Smith (1776) does not claim that dogs or humans have no altruistic motivations whatsoever. His point is just that they have egoistic motivations that frequently influence their behavior. See also Schulz (2018b).

[5] A quick word about trade in services. In a general sense, trading in services—that is, the capacity for reciprocal cooperation (3)—is a more foundational ability than a propensity for trade proper. For an organism to truly engage in *trading* over services, it would need to see itself as *owning* its actions (cf. Locke, 1987 [1866]). This also converges with the arguments in Ofek (2001) that the provision of fire was foundational in human evolution: fire is an impure public good (it is excludable, but nontrivial) and thus facilitated a shift from mere reciprocal cooperation to genuine trading.

However, the three psychological capacities of a concept of [PROPERTY], a concept of [EXCHANGE VALUE], and a disposition and capacity for reciprocal cooperation are especially central to the propensity for trade specifically. Organisms with even strongly developed abilities of communication, mindreading, and representational decision-making do not have the psychological underpinnings for being a trader if they lack the concepts of [PROPERTY] and [EXCHANGE VALUE] and the disposition and capacity for reciprocal cooperation. Organisms that do not recognize ownership and transfer of ownership—even in a minimal way—and/or which are not disposed toward reciprocal cooperation, cannot exchange goods and services with each other no matter how good they are in communicating, reading each other's minds, or evaluating different possible future states of the world. (That said, as will be made clearer in Section 7.3, these other abilities may be crucial components in underwriting the biocultural evolution of an *extensive* disposition to trade.) With this in mind, consider the three core components of the propensity to trade in more detail.

First, to recognize that mutually beneficial outcomes can be achieved through trade, an organism needs to have a concept of [PROPERTY]. Trade—at least in the sense that is relevant here—concerns the exchange of owned goods and services. For two organisms to exchange anything, though, they need to be able to recognize something as theirs and other things as belonging to someone else. Common pool resources cannot be exchanged: if I want to drink from the pond you usually drink from, and you want to drink from the one I usually drink from, then we can certainly switch ponds. However, this is really no different from me deciding what to do on my own, without taking into account your situation (and the same goes for you). Similarly, if I come to your aid in defending against an intruder in the hopes that you will do the same for me in the future, this may be mutually beneficial and cooperative, but it is not trade. A trade—in the sense that is relevant here—only happens when I recognize that you have a claim on something and I have a claim on something else and that we can swap these two claims (Ofek, 2001, p. 144).[6]

Now, exactly how to characterize the necessary concept of [PROPERTY] is not entirely obvious or uncontroversial. However, at a minimum, it requires the recognition that others are *excludable* from the use of the entity in question: if something P is the property of some organism O, then O can prevent others from using or accessing P (Boyer, 2023; Hartley, 2019; Scorolli et al., 2018; Sherratt & Mesterton-Gibbons, 2015; Tibble & Carvalho, 2018; Torii, 1974; Tse, 2008). Beyond this minimal requirement, property concepts of different depths and extents can be distinguished.

On the depth side, property concepts are shallower if they see property as anything from which others can be excluded and which can be disposed of at will (without there needing to be a specific justifying reason for the exclusion and use); they gain in depth if property is seen as something from which others can be *rightfully* excluded and whose use others have no *rightful* say over (Brosnan, 2011; Torii, 1974; Tse, 2008; Wilson, 2020).[7] So, if O can simply force others off a piece of land, this land would be seen as O's territory on the former concept but not necessarily on the latter; on the latter concept, it would need to be determined whether the exclusion can be backed up with sufficient *reasons*.[8] (This also implies that theft cannot be distinguished from legitimate property transfer on shallower property concepts but that it can be so distinguished on the deeper ones.)

[6] Ofek (2001, p. 144) suggests that pure private goods are unlikely to make the basis for trade, but it is not clear why: while individual tokens of private goods cannot be divided, they can still be traded. At any rate, this issue is not central here.

[7] Alternatively and equivalently, one can mark this distinction as concerning "temporary" or "permanent" ownership (Scorolli et al., 2018). Note also that it is possible to further distinguish things from which one can exclude others and things of which one can dispose at will (cf. Native title to land in Australia). I thank an anonymous referee for useful discussion of this point.

[8] The depth of the concept can be said to be increasing with the quantity and complexity of the needed reasons—although making this precise is not necessary for present purposes.

As far as *extent* is concerned, property concepts are wider the more different things they allow to be owned. So, if only territory is recognized as ownable, this is a narrower property concept than one that would also allow for foodstuffs and tokens of exchange ("money") to be ownable. I return to these points in the next section, but, for now, it is just important to note that the concept of [PROPERTY] is intertwined with the ability to recognize (more or fewer) things as being *claimed* (with or without an explicitly represented reason).[9]

Second, the propensity to trade must be seen to be underwritten by a concept of [EXCHANGE VALUE] (Avsar, 2020, pp. 63–64). For an organism to exchange something A for something B, these two things have to be conceived of as exchangeable in the first place. This need not mean that A and B need to be seen as of the same value for either one or all parties of the exchange.[10] Rather, the point is that trades range over items that are conceived of as having a comparable "trade identity." This trade identity need not be the same as an organism's individuation schema for what the world is like: organisms are not forced to see grapes and cucumbers as exchangeable—and certainly not as exchangeable one-to-one—even though they might recognize that the world divides into cucumbers and grapes (among other things) (Brosnan et al., 2008; Oberliessen & Kalenscher, 2019). What is needed is that organisms recognize that their claims on various things can be transferred to others—and vice-versa.

[9] The two dimensions of depth and extent are at least logically independent of each other (although they may be correlated with each other for nonlogical reasons). Note also that the distinction between property concepts of different degrees of depth and breadth is sometimes marked by distinguishing concepts of ownership, possession, and territoriality (Blake & Harris, 2009; Kanngiesser et al., 2020). However, this latter phrasing seems to suggest that this distinction is one of kind, whereas it is more plausibly seen as one of degree. At any rate, this is merely a verbal issue.

[10] This also implies that there is somewhat of a continuum between trading and mere taking. A forced trade—an offer that cannot be refused—is not really a trade at all. Less pronounced, but still uneven power relationships make for borderline cases: if a party A has the power to make another party B accept terms of trade that they would otherwise be unwilling to accept, then this is a non-paradigmatic case of trade only.

Accordingly, concepts of exchange value can be *wider* and *narrower*. They are wider if more things can be exchanged for each other, with complete comparability being the extreme case. They are narrower if only some things can be so compared (i.e., if some things do not have an exchange value). Note that organisms may have a relatively narrow concept of [EXCHANGE VALUE] and a relatively wide concept of [PROPERTY]: many things can be owned, but only a few things can be traded. The key is just that the ability to trade requires that things can be seen as *changing owners*.

The third component of the propensity to trade is the one Smith (1776) focused on: the disposition and capacity to cooperate reciprocally (Avsar, 2020, pp. 52, 60–63, 90–95). This capacity has already been laid out in Chapter 2, so the relevant details do not need to be rehearsed here. In the present context, the key thing to note is that the benefits from trade might accrue a long time after the trading itself has taken place. In the short term, I may (and generally do) gain more from just taking whatever it is that I want of yours—and keeping whatever I have. However, trading *now* may make *future* trades possible, with the overall benefit of trading outweighing the short-term benefits of aggression (Cosmides & Tooby, 1992; Ofek, 2001). Indeed, the benefits of trade may not even involve the other trading party at all. If A trades with B—instead of just taking whatever it wants of B's—A may make itself more attractive as a future trading partner of C. If C has resources that A could not obtain otherwise, this can make trade with B indirectly adaptive—even if is not directly so adaptive (Noe & Voelkl, 2013; Nowak & Sigmund, 2005; Hammerstein & Noe, 2016). An extreme form of the latter case is where trading yields a benefit to *everyone* in the organism's in-group (e.g., by keeping infighting down; Sober & Wilson, 1998; Wilson et al., 2013).

Organisms somehow need to be sensitive to these future benefits in order to be traders. If organisms are just motivated by the immediate, direct benefits on offer, they will in general not be motivated to trade.[11] There are many reasons for why this might be so: for example, these organisms may not recognize the potential benefits from trade, or they may recognize them but not be motivated to act on them. Indeed, they may even be motivated by them but be unable to resist the pull of the present benefits of aggression (Avsar, 2020, p. 72; Ofek, 2001, p. 142; Rosati et al., 2006; Stevens & Stephens, 2008). Sections 7.2 and 7.3 return to some of these reasons; for now, the key point to note is just that, as also pointed out by Smith (1776), the propensity to trade needs to be seen to be underwritten by a robust disposition and capacity to cooperate reciprocally since the gains from trade need not be immediately obvious and might take a long time to materialize.

All in all, therefore, the propensity to trade is centrally (though not exclusively) underpinned by the conjoint presence of three different psychological and behavioral traits: (a) a concept of property, (b) a concept of exchange value, and (c) a disposition and capacity to cooperate reciprocally. It is important, though, to be clear on the extent of this claim. The idea here is that organisms with a propensity to trade need to have traits (a)–(c). The idea is *not* that organisms with the abilities (a)–(c) necessarily engage in much trade. Whether the latter claim is true depends on a number of further factors—a point to which Section 7.3 returns. For present purposes, what is important to keep in mind is just that abilities (a)–(c) provide the major raw materials out of which the propensity to trade is built.

[11] What about environments in which seizure is either impossible or carries the risk of serious costs (which may have been true for early human social environments)? It is true that, if a party cannot seize the item in question but there is really nothing else that it wants as much, it may be willing to trade something else it has now for the item in question. However, for the other party to accept the trade, it still requires that this other party is interested in this other thing (i.e., that it is not just motivated by the immediate, direct benefits of the item on offer).

7.2 Human Trading

Are humans extensive traders? In one sense, the answer is clearly "yes." As a species, humans trade vast amounts of stuff: in 2018 (nominal) global trade was valued at $19.48 trillion and accounted for about 60% of global gross domestic product (GDP) (World Bank); while the COVID-19 pandemic lowered these values significantly, they are unquestionably still very high from the perspective of comparative psychology. While other animals also engage in some forms of trading—for example, chimpanzees, gorillas, orangutans, and bonobos have been found to engage in the trade of tokens or foodstuffs (Brosnan et al., 2008; Parrish et al., 2013; Pelé et al., 2009), and cleaner fish and their hosts exchange food for cleaning services (Triki et al., 2019)—there is just no question that humans, as a species, are orders of magnitude more engaged in this activity than are other organisms (Nowell, 2010; Ofek, 2001). That is, as a species, humans are extensive traders, in the sense that some humans trade vast amounts of goods and services with a vast number of other humans.

However, in another sense, the fact that humans are extensive traders is much less clear. First, the extent to which humans engage in trade is culturally highly variable (Apicella et al., 2014; Basu & Waymire, 2006, p. 211; Basu et al., 2009; Hartley, 2019).[12] While

[12] Apicella et al. (2014) also show that the Hadza display the *endowment effect* only in situations where many different things are seen as ownable. This is interesting, as endowment effects have been observed in many different species and for many different goods (Brosnan et al., 2012; Brosnan et al., 2007; Drayton et al., 2013; Flemming et al., 2012; Jaeger et al., 2020; Kanngiesser et al., 2011; Lakshminarayanan et al., 2008). For present purposes, though, it is important to note that endowment effects are not about trade per se, but directly only speak to an organism's reluctance to give up something they have control over. So, Jaeger et al. (2020) and Jones (2001) argue that the endowment effect can be explained as the result of an evolved "time-shifted rationality" (see also Brosnan & Beran, 2009). The downside of giving up items—especially fitness-relevant ones—is greater than the upside of obtaining them: there is always a chance that a promised return will not materialize, whereas giving up something one already has in possession is a guaranteed loss. Given this, it is possible that organisms display endowment effects without engaging in much trade (when they think that they are unlikely to attain control over a currently uncontrolled item) or that they do not display them even

trade is a major aspect of life in contemporary Western society, it is less so in other cultures. So, members of the Hadza (or at least those Hadza who are living in more remote areas) tend to engage in little if any trading.

> Aside from bows and arrows and the clothes on one's back, few capital goods or personal possessions exist. If one person has more than his or her share of a resource, such as two knives, the surplus will be demanded by others who have less. Thus, to a first approximation, Hadza economic life is characterized by nearly 100 percent taxation and redistribution. (Apicella et al., 2014, p. 1796)[13]

Similarly, Henrich et al. (2001) note that,

> The Machiguenga show the lowest cooperation rates in public good games [in the authors' sample], reflecting ethnographic descriptions of Machiguenga life, which report little cooperation, exchange, or sharing beyond the family unit. (Henrich et al., 2001, p. 76)[14]

This variance in reliance on trade also exists in the temporal dimension: while there are good reasons to think that trade has been a crucial component of the biocultural evolution of some

if they do engage in much trade (when they think that they are very likely to obtain control over a currently uncontrolled item). Further discussion of endowment effects is left for a future occasion; for now, though, it is just important to note that the burgeoning literature on endowment effects can be easily combined with the account of the biocultural-technological evolution of the propensity to trade set out here.

[13] The Hadza can, of course, still be seen as reciprocators; they just don't *trade* in the narrow sense relevant here. See also Note 5.

[14] However, see also Henrich (1997), which draws a slightly different picture of life among the Machiguenga (at least to the extent that they are connected to Western trading institutions). However, whatever exactly may be true about the Machiguenga, the general point in the text is not controversial: the extent to which humans engage in trade is culturally highly variable.

human populations (Brooks et al., 2018; Nowell, 2010; Ofek, 2001, pp. 176–178, 226; 2013), *the extent to which* ancient human cultures relied on trade likely differed across time (Avsar, 2020; Hartley, 2019). For example, earlier human groups are likely to have relied less on trade, with extensive trading institutions only appearing with the advent of writing technology some 5,000 years ago (Hartley, 2019; Mullins et al., 2013a; Seabright, 2010).[15]

The second reason why the characterization of humans as extensive traders is more complex than it might at first appear resides in the fact—already mentioned in Chapter 2—that many aspects of the major psychological subsystems underlying their propensity to trade (the disposition and capacity to cooperate reciprocally and the concepts of [PROPERTY] and [EXCHANGE VALUE]) are shared with non-human animals. Given this, seeing humans as standouts for their propensity to trade in fact requires a fine-grained comparative analysis of these subsystems.

As far as the disposition and capacity to cooperate is concerned, the first thing to note is that cooperation is widespread throughout the biological world (Griffin et al., 2004; Sober & Wilson, 1998; Sterelny et al., 2013; West et al., 2011). However, as noted in Chapter 2, this is not to say that there are not differences in the ways that humans and other organisms cooperate. In particular, *reciprocal* interactions (i.e., where individuals cooperate only as long as others do so, too) are rare outside of the human realm and significantly more restricted (Avsar, 2020, p. 16; Cosmides & Tooby, 1992; Hammerstein, 2003; Hammerstein & Noe, 2016). Outside the human sphere, tit-for-tat helping behavior has been observed in vampire bats, chimpanzees, and several other species (Carter & Wilkinson, 2013; Hammerstein, 2003; Hammerstein & Noe, 2016; Ofek, 2001). However, there is no question that the extent to

[15] Seabright (2010), in particular, makes clear both the temporal depth of extensive trading and the social and cognitive conditions underlying the reciprocal trust making it possible (including social institutions like money and banking).

which other species rely on reciprocal cooperation is much lower than what is found in humans, where—as just noted—these kinds of interactions can provide the foundation of much their social behavior (Boyd, 2018; Kaplan et al., 2012).[16]

When it comes to the concept of [PROPERTY], the first thing to note is that, as briefly mentioned in Chapter 2 as well, it turns out that Smith (1776) is wrong about the fact that humans are alone in having such a concept (as is Hume, 1978 [1739], p. 489, for that matter). Many other animals also recognize property, at least in a rudimentary manner (Eswaran & Neary, 2014; Gintis, 2007; Kummer, 1991; Tibble & Carvalho, 2018). For starters, many animals are territorial and thus recognize that some parts of the environment belong to others (for an overview, see, e.g., Hinsch & Komdeur, 2017; Sherratt & Mesterton-Gibbons, 2015). Furthermore, various animals—including, most notably, primates—can recognize that they have a claim on things other than territories, too (Brosnan, 2011; Kummer & Cords, 1991; Lakshminarayanan et al., 2008; Pryor, 1981; Torii, 1974). Still, it is true is that the human concept of property is significantly more sophisticated than that of non-human animals (Boyer, 2023; Kanngiesser et al., 2020; Wilson, 2020).

In particular, human property concepts are *deeper*. Non-human animals do not recognize others as having a *claim* on something

[16] Indeed, Boyd (2018) argues that, in many non-human cases, the conditions making reciprocation mutually profitable seem satisfied yet there is no reciprocal cooperation. It is also important to emphasize that the fact that humans are standouts for their ability to reciprocally cooperate is not equivalent to them being traders. So, if Cosmides and Tooby (1992) (for example) turn out to be right in arguing that humans evolved dedicated psychological machinery for detecting cheaters in situations of social exchanges, then this, by itself, does not explain that humans are also standouts for being traders. Social exchange conditionals and the detection of cheaters is much wider than trading in property: humans may be standouts for reasoning with rules of the form "If you receive the benefit, then you pay the cost," but they need not apply these rules to trade with property. I can be good at telling if someone violates the rule "If you have a tattoo, then you climbed the mountain," but I may not be inclined to trade my excess food supplies for your excess climbing gear. So, the fact that humans are especially well equipped for reasoning with social exchange conditionals is only an aspect of the explanation of their being traders—it is not the full such explanation.

and *therefore* being able to exclude others from it (Tse, 2008, p. 287; Wilson, 2020). Rather, they (a) recognize that others simply have the ability (e.g., physical strength) to exclude others from some things, or (b) they fail to see competition for the resource as worthwhile (Brosnan, 2011; Gintis, 2007; Kummer, 1991; Scorolli et al., 2018).[17] Animals think of things belonging to someone as things that others de facto will or cannot challenge them for—without any normative (i.e., not purely causal) reason for this status being represented. By contrast, humans represent exactly this: they distinguish theft from gifting or borrowing precisely on the grounds of whether the *claim* to the relevant good has changed possessor or not (Tse, 2008, p. 287). This is a point that has been studied extensively also developmentally, where it has been found, for example, that even young children recognize a history of use as a key component in establishing ownership conventions (apart from physical excludability) (Friedman et al., 2011, 2013).

Furthermore, the *breadth* of animal property concepts is limited in relation to human property concepts. Non-human animals recognize a variety of things as potential property—some of them even relatively abstract. They can recognize *territories* as property, they can recognize *foodstuffs* as property, and they can even recognize some *tokens of exchange* as property (Brosnan et al., 2008; Parrish et al., 2013). However, they cannot recognize much more than that as potential property (Brosnan, 2011; Kanngiesser et al., 2011). In the human case, by contrast, these limits are very wide: humans recognize "air rights," shares, art works, pets, etc. as ownable. Indeed, it is arguable that the human concept of [PROPERTY] has no limits at all—people invent new ownable items on a regular basis (cf. the development of mortgage-backed securities).

Very similar remarks hold for the concept of [EXCHANGE VALUE]. Again, at least some non-human animals also have

[17] This is obviously so for nonbiological entities (like fire) or predators, but also extends to members of their own species.

rudiments of such a concept. For example, a number of primates are willing to exert effort to obtain a token that can be exchanged for a piece of cucumber—but only if others do not get a grape for the same effort (Brosnan et al., 2008; Oberliessen & Kalenscher, 2019). Thus, these animals show that they think of cucumbers and grapes as exchangeable (though not on a one-to-one basis).[18]

However, the concept of [EXCHANGE VALUE] is significantly *narrower* in animals than in humans. Humans can conceive of many different things as comparable with each other. Indeed, humans have conceived a universal medium of exchange—money—that allows them to trade anything for anything. While some animals can use tokens (Brosnan et al., 2008; Parrish et al., 2013), this use is more limited: they learn that they can exchange tokens for food, for example, but there is no evidence that they can learn to exchange tokens for toys or tools, say (Brosnan et al., 2008; Parrish et al., 2013).[19] Humans, though, are not so constrained and exchange money for toys and tools on a regular basis—as well as for aesthetic experiences (such as concerts or art exhibitions), services (such as transportation), and many other things.

All in all, therefore, while none of the three component abilities that make up the propensity to trade is exclusive to humans, important contrasts between human and non-human animals still exist. In particular, humans have the ability to cooperate reciprocally more extensively and frequently than do non-human animals, and they have the abilities to rely on deeper and/or more extensive concepts of [PROPERTY] and [EXCHANGE VALUE].

Hence, putting all of this together, it becomes clear that, when it comes to the distinctively human propensity to trade (i.e., the

[18] However, it is not clear if these animals would consider an exchange of a large number of cucumber pieces to be equivalent to one grape. Note also that exchange has not been observed in wild populations and that it is not clear exactly what the primates in these experiments have learned.

[19] Similarly, while cleaner fish exchange cleaning services for food, there is no evidence that they are disposed to exchange cleaning services for anything else—or that they can exchange anything besides cleaning services.

combination of a sophisticated disposition and capacity to cooperate reciprocally, as well as deeper and/or wider concepts of [PROPERTY] and [EXCHANGE VALUE]), two facts need to be explained. (1) Why did humans evolve in such a way that they at least sometimes act on such an extensive propensity to trade? That is, why is it that, in some human societies, a major proportion of social interactions revolve around trade—with people engaging in trade of very many goods with very many others—when this is not the case in non-human societies? (b) Why does the distinctively human extensive propensity to trade, barter, and exchange not always display itself? Why do we not find all humans relying on trade to the same extent? The next section sets out answers to these questions.

7.3 The Bio-Cultural-Technological Evolution of the Human Propensity to Trade

To explain the unique features of the distinctively human extensive propensity to trade—its *existence* as well as its only sometime *display*—I proceed in two steps, following the pattern laid out in Chapter 4. First, I bring out the selective pressures on the psychological underpinnings of an extensive propensity to trade in humans. Given this, I then, second, show that there is a distinctive set of sociocultural-technological interactions that makes the *display* of this propensity particularly likely.

Before beginning this analysis, though, it is important to make the starting point clear. Given that many non-human animals engage in trade—if only in a rudimentary manner—common evolutionary biological theorizing suggests that early hominins traded much like closely related non-human animals (such as other great apes). Turning this around, it is plausible to think that humans have engaged in trading for a very long time (Boyer, 2023; Brooks et al., 2018; Nowell, 2010; Ofek, 2001). Given this, what needs an explanation is why trading underwent a pronounced shift in hominins—from something that is not orders of magnitude different from what we find in non-human primates to the heavy dependence on complex markets in places like New York or Hong Kong. Doing this is the goal of the rest of this section.[20]

At the heart of the first step of this explanation is the fact that *extensive* trade—rather than limited bartering over few goods or services—is highly adaptive where it can be engaged in (see also Cosmides & Tooby, 1992; Ofek, 2001, p. 13). If I can only exchange *territories* with *you*, entering into trading relations with you—rather than fighting over the territories—will be adaptive to the extent that future exchanges of territory with you are adaptive. By contrast, if I can exchange territories for food or weapons with many different people, then the potential gains from trade increase massively and accrue more quickly (Apicella et al., 2014, pp. 1796, 1804; Basu & Waymire, 2006, p. 211; Basu et al., 2009; Cosmides & Tooby, 1992; Hartley, 2019).

To see this in a bit more detail, assume organisms can choose to offer up a *trade* (*T*) for a given resource or to simply *take it* (*A*).[21] Assume also that taking the resource from an organism that is offering a trade is very likely to lead to success in obtaining the resource, that trying to take the resource from an organism that is also trying to take the resource for themselves will only sometimes lead to success, and finally that offering a trade to an organism similarly offering a trade provides some of the benefits that access to the

[20] *Inter alia*, what needs to be explained is how (some) humans went from relatively shallow concepts of [PROPERTY] and [EXCHANGE VALUE] (e.g., as laid out in Gintis, 2007) to the deeper one found in some contemporary cultures.

[21] Note that the arguments to follow differ from the classic hawk/dove/bourgeois accounts of the evolution of property rights: see, e.g., Eswaran and Neary (2014), Guala (2016b), Gintis (2007), Maynard Smith and Parker (1976), and Sherratt and Mesterton-Gibbons (2015) (see also Smead & Forber, forthcoming). The argument here is not whether to respect someone else's property rights, but whether to trade for a resource. Similarly, the issue is not just that of determining when it is beneficial for humans to engage in reciprocal interactions: as in Note 16 above, not all forms of reciprocity concern trade.

resource provides—but, since something else needs to be given up for this access, not as many of these benefits as simply taking it. (I return to the reasonableness of these assumptions below.) In that case, the payoff matrix of this interaction will take the shape of the classic Prisoner's Dilemma (the numbers in Table 7.1 represent fitness or utility values—a point to which I return momentarily):

Consider now a situation where the humans playing this game are aware of both the fact that they are interacting strategically and the nature of the game they are playing (the "rational" form of a strategic interaction). In such a situation, the more goods that are conceived as tradeable, the greater the likelihood that this game will be conceived of as infinite. In turn, this implies that the game is more likely to have equilibria other than *A/A*: as is widely known, finitely repeated instances of this game only have *A/A* as a Nash Equilibrium in all rounds (see, e.g., Binmore, 1998). The more different goods that can be traded with more different others, the less likely it is that any given round of play is the final round. If an organism can trade territory for many other goods, with many different other organisms, the potential number of interactions is large, and there is much less of a clear end of these interactions in sight. This can make many strategies—including *T/T* throughout—equilibrium strategies in the game.

Importantly, the substance of this argument is unchanged even if the assumption of the rational nature of the interaction is dropped (so that this strategic interaction is turned into a "biological" one). So, assume that (a) humans happen to follow some strategy (whether it is *T*, *A*, or some combination thereof) and (b) that they

Table 7.1 A Trading Game

Player 1/Player 2	**Trade (*T*)**	**Take (*A*)**
Trade (*T*)	2, 2	0, 4
Take (*A*)	4, 0	1, 1

adjust this strategy in accordance with how well it does in relation to alternative strategies they encounter. As is also well known, on this biological version of the game, organisms that always *take* will invade a population of traders and go to fixation (Alexander, 2007; Binmore, 1998; Skyrms, 1996). Similarly, it is well known that organisms playing a wide variety of non–purely-taking strategies—such as trade tit-for-tat (*TFT*)—can invade a population of takers in even finitely iterated versions of the game (Alexander, 2007; Axelrod, 1984; Binmore, 1998; Hilbe et al., 2013; Skyrms, 1996).

However, what is key for present purposes is that whether this is possible and how quickly it happens depends on the details of the interaction: the number of rounds in each interaction of the game and the differences in the payoffs.[22] So, for example, the conditions for *TFT* to be an evolutionary stable strategy in this game are that (1) $w(TFT, TFT) > w(A, TFT)$, or (2) $w(TFT, TFT) = w(TFT, A)$ and $w(TFT, A) > w(A, A)$, where $w(x, y)$ is the fitness payoff of playing x against y (Alexander, 2009). Using the above payoff values for illustrative purposes, this implies that if the trading interaction is repeated for two rounds, $w(TFT, TFT) = 4$ and $w(A, TFT) = 5$. Hence, in that case, *TFT* is not an evolutionarily stable strategy (ESS). By contrast, for four rounds, $w(TFT, TFT) = 8$ and $w(A, TFT) = 7$.[23] Hence, increases in the number of rounds of the interaction leads trading-based *TFT* to become an ESS.[24] Importantly, furthermore,

[22] In what follows, I focus on *T*-initiated *TFT*; however, the point generalizes to other non-purely-*A*–based strategies.

[23] For three rounds, the second of the above conditions holds, but the issues are clearer to see if the focus is just on the first condition.

[24] This point can be strengthened by considering the dynamics of the situation. If the games last two rounds, the average fitness of *TFT* and *A* are $w(TFT) = 3p + 1$ and $w(A) = 3p + 2$, where p is the proportion of *TFT* traders in the population. The average fitness of the population is $= 2p + 2$. Assuming, not unreasonably but here largely for convenience (Alexander, 2009; Skyrms, 1996) that the population distributions change according to the replicator dynamics—where strategies change as a function of how their payoff differs from the average in the population—the rate of change of *TFT* traders is $(TFT)/t = (1p - 2)p$, which is negative for all p [0, 1]. By contrast, if the game is played for four rounds, the average fitnesses change to $w(TFT) = 7p + 1$, $w(A) = 3p + 4$, and $= 4p^2 + 4$. The rate of change of *TFT* traders becomes $(TFT)/t = (7p - 4p^2 - 3)p$, which is positive for $p > 0.75$. Again, therefore: if the number of iterations of the game increases, trading

increasing the number of goods that are seen as tradable, by the same reasoning as on the rational version of the game above, is tantamount to changing the time horizon of the interaction. The more goods that can be traded for more other goods, the more opportunities there are in engaging in trading relations and thus the more likely it is that trading with other traders incurs sufficiently large benefits to make it adaptive relative to just taking or fighting over the resources—both of which have benefits as well (depending on the costs of fights or taking).[25]

What, though, explains that humans are able to engage in extensive trade like this (i.e., of many different goods and services with many others) when non-human animals are not? The answer to this question lies, again, in the particular social environment in which humans evolved and the sociotechnological interactions the latter can engender. There are two aspects to this answer.

First, given that different members of this kind of environment likely had different intentions and beliefs, which impacted their likelihood of providing or reciprocating help, keeping track of what others are thinking, what their role in society is, and who is in their inner circle of partners became adaptively highly important (Humphrey, 1986; Schulz, 2018b; Sterelny, 2003, 2012a; Whiten & Byrne, 1997). In an environment like this, a specific set of memory abilities is highly adaptively important. More specifically, given the

changes from a strategy that is driven to extinction to one that will spread and take over the population (assuming there are sufficiently many other traders in the population).

[25] Indeed, this is related to the well-known point that, under some conditions, the Prisoner's Dilemma can be turned into the "Stag Hunt" (in which cooperation is an equilibrium) (Alexander, 2007; Skyrms, 2004). So, in the present case, if the payoffs of the four-round version of the above game are added together—so that this is treated as a one-shot game—the game indeed becomes a Stag Hunt. An alternative way of putting this point is that, with minimal general trading abilities, the interaction gains T/T as an equilibrium focal point (Leeson et al., 2006). A final point worth noting here is that this model applies best to intercommunity trading; within a community, the costs of *taking* plausibly are often so high as to make it an unattractive option compared to some form of cooperation. However, it remains true that trading becomes a *particularly beneficial* form of cooperation only when many different goods are seen as tradeable.

biological importance of cooperation for all humans, and given the fact that not all members of society can be depended on to be kin and thus to have direct biological interests in cooperating, keeping track of who cooperated with whom in the past was very important for humans (Sterelny, 2003, 2012a).

Similar points hold for the kind of mindreading and moral-normative motivational abilities that make reciprocal cooperation possible. As noted in the previous two chapters, determining what others are thinking (whether they are currently hostile or friendly, say) and being able to hold out for large (and possibly uncertain) future benefits at the cost of forgoing smaller (and possibly certain) present benefits ensured continued access to a major element of human adaptive success—the support of others—and thus had major adaptive value (Cosmides & Tooby, 1992; Hrdy, 2011; Klein et al., 2002; see also Rosati et al., 2006; Sterelny, 2003, 2012a; Stevens & Stephens, 2008).

As noted in Chapter 2, there may also be further, specific psychological capacities underlying the human capacity for reciprocal cooperation (Cosmides et al., 2018; Cosmides & Tooby, 1992), although that is somewhat controversial still (see, e.g., Sober & Wilson, 1998; Sterelny et al., 2013; West et al., 2011). However, what matters here is just that the kind of social environments that humans evolved in paved the way for the evolution of some of the further psychological abilities associated with an extensive propensity for reciprocal cooperation (see also Cosmides & Tooby, 1992; Hrdy, 2011; Whiten & Byrne, 1997).[26]

Still, by themselves, these initial selective pressures—and the resultant trait differences—may have been relatively minor. However, this is where the second step of the interactionist account of the evolution of the human propensity for trade comes in. In particular, the specific sociocultural environment that humans evolved in matters for the uniquely human propensity to trade because it

[26] Field (2002) notes that this point is missing in Ofek (2001).

allowed them to bioculturally evolve the kind of *cognitive and social technology* that allows them to rely on deep and wide concepts of (a) [PROPERTY] and (b) [EXCHANGE VALUE] and (c) to engage in widespread reciprocal cooperation.

In particular, as noted in the previous chapters, through cultural learning, humans became able to construct and use advanced mnemonic and inferential tools: quipus, paper, calculating devices like the "Senkereh Tablet" (Sugden, 1981) and sundials (King, 1955), written language (Basu & Waymire, 2006; Basu et al., 2009; Hartley, 2019; Mullins et al., 2013a; Tse, 2008), myths, songs, dances, and algebraic conventions. As also noted in the previous two chapters, these devices allowed humans to store and process more and more complex information more efficiently and for longer periods and thus vastly enhanced their cognitive abilities in certain dimensions.

Noting this is key for present purposes as the concepts of [PROPERTY] and [EXCHANGE VALUE]—especially in their deeper and wider forms—are abstract and complex in ways that make the use of such technology highly adaptively valuable for beings like us. These concepts are built out of other concepts like [CLAIM], [EXCLUDABLE], [MEANS], [EXCHANGE], and they do not have a straightforward perceptual signature: many different states of the world can constitute A's owning something. Because of this, reliance on such concepts is highly time-, concentration-, and attention-consuming (Fodor, 1983, 1990; Margolis & Laurence, 2015; Prinz, 2002; Schulz, 2018b). Without cognitive tools, the frequent use of concepts like [PROPERTY], [EXCHANGE VALUE], [CLAIM], [EXCLUDABLE], [MEANS], and [EXCHANGE] would thus generally be impossible or not worthwhile.[27]

Again, the point is not that concepts like [PROPERTY] or [EXCHANGE VALUE] do not have an innate foundation as well

[27] In this way, the present picture agrees with Wilson (2020) that human property cognition is underwritten by social arrangements, but—for reasons laid out in Chapter 3—it diverges from the latter in placing less emphasis on symbolic cognition.

(Carey & Spelke, 1996; de Hevia et al., 2014; Gopnik et al., 2004). As noted in Chapter 4, it is plausible that they do. Rather, the claim is that, without the uses of cognitive technology, reliance on these innate representational expectations is unlikely to be possible. The storage of important, frequently used trading information (e.g., who owns what, wants to trade what for what, etc.) using quipus, writing, or other suchlike tools, frees concentration and attention: organisms do not have to use scarce cognitive resources (including time) to recall that information but can simply look it up. Complex representational inferences about how much of good A makes for a beneficial trade for how much of good B are made easier if intermediate steps of the inference can be written down, if parts of it can be automated (e.g., using computational devices like sundials), and if balances can be tracked and analyzed with cognitive tools like written records and a number system that allows for the calculation of averages and differences (Schulz, 2018b; see also Kelly, 2015, for more on the amounts of information that can be stored and tracked in nonliterate societies).

This is important as—for reasons made clearer in Chapter 4—it implies that, in the absence of tools that help relate these concepts to the actual conditions prevailing in the world, frequently relying on such concepts is unlikely to even be possible. It is only through cognitive tools like quipus, petroglyphs, myths, songs, writing, or algebra that it becomes possible for humans to recognize highly abstract things as belonging to someone, and they can come to recognize the excludability of an owned entity as resting on abstract notions like [DESERT], [HISTORICAL CLAIMS], or [MIXING OF LABOR] (Scorolli et al., 2018; see also Hartley, 2019; Mullins et al., 2013a; Ofek, 2001, pp. 180–187). Recognizing that person A owns object P because A, decades ago, mixed their labor in the appropriate manner with P (say), and that therefore, A deserves to decide what to do with P even if A is not in direct physical contact with P is very difficult without tools that allow us to *store* what A did decades ago, to *compare* this to what A needed to do to acquire a claim on P, to *transmit* what A intends to do with P even where P is physically far from A, etc. Cognitive and social technology is needed to tie these highly abstract and complex concepts to the actual conditions prevailing in the environment at any given point in time.[28]

Importantly, though, since humans are cultural learners, they can obtain exactly these kinds of sophisticated cognitive and social tools. Hence, it is in particular humans that are able to rely on deeper and/or wider concepts of property and exchange value in their interactions with each other.

The second aspect of the importance of cognitive technology is that it supports more sophisticated reciprocal cooperation. Keeping track of who helped who, when, and how is easier—and may even be made possible in the first place—when this information can be stored externally and recalled when needed. Doing this frees up cognitive resources like concentration, attention, and memory (Mullins et al., 2013a). In the absence of this technology, sophisticated reciprocal cooperation may not be practically feasible: with many different goods that can be traded with many different others, the mnemonic and processing demands of tracking trades quickly become so vast as to make the latter practically impossible without relying on cognitive technology like writing, algebra, or quipus. This thus expanded the capacities for reciprocal cooperation—which, as just noted, may not have been drastically different from what is found in non-human animals.[29]

Putting all of this together: in order for a species to be able to be trade with many different goods—and thus to harness the adaptive benefits of extensive trade—cognitive and social technology is

[28] Again, this is not to say that underlying these concepts are not further innate abilities, for example, concerning mindreading or future planning (Kabadayi & Osvath, 2017; Leslie et al., 2004). The point is just that, whatever else is needed to use these abilities, cognitive tools are necessary, too.

[29] In this way, the present perspective expands on that in Ross (2004, 2007), which focuses on linguistic and narrative competences to account for our enhanced abilities for social cooperation.

crucial. Once many things are seen as tradable with many others, it becomes vital to have ways of easily checking who currently owns what and who got what from whom. Cognitive and social technology (quipus, petroglyphs, written language, algebra, computational devices, etc.) allows exactly this. In turn, to acquire such a technology, though, cultural learning is crucial. Such technology is highly unlikely to be discovered by individual organisms but needs to be refined over generations (Heyes, 2018; Sterelny, 2012a; Tomasello, 1999). This, thus, explains why humans—unlike other organisms—evolved the ability to be extensive traders: humans, more than other animals, evolved in a particular, complex social and cultural environment that provided the ground for the acquisition of the cognitive technology needed to support a sophisticated disposition and capacity for reciprocity as well as deep and/or wide concepts of [PROPERTY] and [EXCHANGE VALUE]—and thus, which can underwrite large-scale trading relations.

As before, though, this then starts a feedback process: cultural learning makes cognitive and social technology available, which, in turn, makes extensive trade possible—which can then lead to a division of labor that further enhances cultural learning and cognitive and social technological development. It is a classic insight of economics that with trade and a division of labor, research and development of tools become possible that is orders of magnitude greater than would be possible without it (Ofek, 2001, 2013; Seabright, 2010; Smith, 1776; see also Chapter 8 for more on this). More cognitive technological development, in turn, can then further extensive trade and so on. The upshot is the process laid out in Figure 7.1.

This, though, answers only one of the two questions that are open concerning the extensive propensity to trade found in humans. We also need to explain why this extensive propensity does not always display itself. Why do not all human populations engage in extensive trade? What accounts for the variability in this respect?

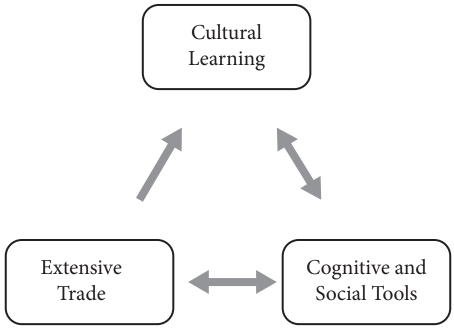

Figure 7.1 An interactionist feedback loop surrounding the human propensity to trade.

Given the preceding discussion, though, the outline of the answer to this question can now be easily seen. In particular, in light with the arguments of Chapters 4–6, this answer turns on the fact that the extensive propensity to trade comes to its own only if it is supported by the right external arrangements. Specifically, humans can rely on deeper and/or wider concepts of [PROPERTY] and [EXCHANGE VALUE] to the extent that the relevant cognitive and social technology is available. Where humans do not have access to this kind of technology, they are also less likely to rely on deeper and/or wider concepts of [PROPERTY] and [EXCHANGE VALUE] and are less likely to be able to engage in the trade of many different goods.[30]

Importantly, though, the appropriate kind of cultural-technological evolution need not occur in all cultures. Not all cultures have evolved the same kinds of cognitive and social technology—for example, fairly few cultures have a powerful cognitive technology like a written language, and quipus, too, are

[30] Again, Seabright (2010) underscores this point well, noting that extensive human cooperation needs to be underwritten by social institutions (like money) that scaffold trusting relations among strangers.

not found in every culture. Cultural evolution is not a deterministic system, but comes with much variability: differences in who learns from whom and in what manner can lead to very different cultural outcomes (Henrich, 2015, 2020; Henrich & McElreath, 2007;Heyes, 2018; Tomasello, 1999). Therefore, not all cultures may have the technological underpinnings to allow for sophisticated reciprocal cooperation and especially deep and/or wide concepts of property and exchange value. Ipso facto, not all cultures will display the propensity to trade to the same extent (or at all).

Turning this around: where cultural learning created the cognitive and social technology that supports relatively strong reciprocal cooperation and relatively deep and/or extensive concepts of [PROPERTY] and [EXCHANGE VALUE], it was possible to trigger an extensive propensity to trade—that is, one that allowed for trade among a wide number of goods with a wide number of people. In turn, the latter can then kickstart the development of more advanced cognitive and social technology, further supporting an extensive propensity to trade. In this way, it becomes possible to see why it is both true that extensive trade is an important part of the human lineage (see, e.g., Ofek, 2001, pp. 136, 172, 217), and that trade is culturally and temporally specific (see, e.g., Apicella et al., 2014; Avsar, 2020, pp. 16–17; see also Eswaran & Neary, 2014).

To understand this better, it is important to forestall some possible misconceptions. First of all, the claim here is *not* that trade in the human lineage came into existence with the development of cognitive and social technology. In fact, there is every reason to think that trading has been around longer than the kinds of cognitive and social technology (algebraic conventions, a written language, etc.) just mentioned (Nowell, 2010; Ofek, 2001). Rather, the argument here is that *the particularly extensive* propensity to trade found in *some* humans is associated with the advent of this kind of cognitive and social technology. Undoubtedly, as noted earlier, some humans traded before they culturally evolved this kind of technology; however, a uniquely strong propensity to trade came later, once humans had culturally evolved this kind of technology.

Importantly, this is also a key prediction that the present account makes: we should find a correlation between the cultural evolution of the above kinds of cognitive tools and the emergence of extensive trading relations. As it turns out, although further corroboration of this prediction would be useful, there is some initial support for this prediction. For example, extensive trading institutions tend to be seen only where a culture has developed writing technology, such as the Mayan, Babylonian, ancient Egyptian, and ancient Chinese ones (Hartley, 2019; Mullins et al., 2013a).[31] At this stage, though, this first prediction is at least to be seen as credible.

Second and relatedly, the claim is also not that humans did not cooperate extensively before they developed this kind of cognitive and social technology. Again, there is every reason to think that this is false (Cosmides & Tooby, 1992; Sterelny, 2003, 2012a). Rather, the point is that the extensive reciprocal cooperation that is characteristic of trade with many different goods and many different people is tied to the cultural emergence of cognitive and social technology.

In this way, the two questions from the previous section can be answered. (1) The reason that humans evolved in such a way that they sometimes display an extensive propensity to trade lies in the fact that they sometimes lived in the kinds of sociocultural environments that provided conditions for the biocultural evolution of the cognitive and social tools needed to underwrite a sophisticated disposition and capacity to cooperate and deeper and/or wider concepts of property and exchange value. (2) This human

[31] Note also that it is not clear that other distinguishing features of these cultures—such as their large range—are a plausible explanation for them being based on extensive trade. It's a classic economic insight that, even in small-scale cultures, there is much to gain from specialization and division of labor (as seen in the classic works of Ricardo and Smith). Hence, the fact that these cultures span different geographical ranges is, if anything, better seen as the outcome of their containing extensive trading and cognitive technology rather than a presupposition of the latter.

propensity to trade does not always display itself as it requires the right institutional settings to be triggered—namely, the existence of cognitive and social technology like writing that can support many repeated reciprocal interactions, as well as deep and/or wide concepts of property and exchange value.

7.4 Conclusion

How do humans differ from other animals in their propensity to trade? Why does this propensity not always display itself even among humans? This chapter developed an integrated set of answers to these two questions. The propensity to trade breaks down into three component psychological traits: the ability to cooperate, and concepts of [PROPERTY] and [EXCHANGE VALUE]. At the core of the human–non-human differences in these three traits is the fact that, due the fact that humans evolved in a specific kind of sociocultural environment, they evolved a more extensive ability to interact reciprocally, as well as the ability to rely on especially deep and/or wide concepts of [PROPERTY] and [EXCHANGE VALUE]. In turn, these differences allowed humans to display an extensive propensity to trade in the right institutional settings: namely, ones where the appropriate cognitive technology culturally evolved that allowed many different things to be traded for many other things with many other people.

It is now possible to bring together some of the key upshots from the arguments of the previous five chapters and apply them to some outstanding current issues. This is the goal of the next chapter.

8

Artificial Intelligence, Religion, and Patents

Some Applications

It is now possible to explore some applications of the present account of the evolution of distinctively human cognition to some issues of contemporary importance. While there are many possible such issues to choose from, the ones focused on here concern artificial intelligence (AI), religion, and patents. There are three reasons why it is specifically these issues that are at center stage here.

First, they undoubtedly hold great contemporary importance. With the rapid advances in AI, major questions about the future place of human cognition on the planet are becoming increasingly more pressing. The importance of religion to human life has not lessened over time; it continues to be a major aspect of human thought and action, and getting clearer on exactly how it is related to the latter is thus very important. Similarly, determining the factors influencing economic growth is one of the major practical questions of economics, making the investigation of patent regimes as engines of innovation a much-discussed issue. In short, there is no question that these three issues deserve a closer look in and of themselves.[1]

Second, these three applications are useful for clarifying various aspects of the present account of distinctively human

[1] This is, of course, not to say that some other issues would not have this intrinsic importance, too—just that these three do.

It's Only Human. Armin W. Schulz, Oxford University Press. © Oxford University Press 2025.
DOI: 10.1093/9780197800188.003.0008

cognition. As also noted throughout the book, religion is often seen as a major aspect of distinctively human cognition (Atran, 2002; Boyer, 2001; Henrich, 2020; Pyysiainen & Hauser, 2010); it thus deserves a closer look for that reason alone. More than that, though: all of the applications here reveal new aspects of the theory of human cognition developed in this book; for example, the discussion of AI reveals important facts about the human ability to create and enjoy music, too. These applications thus also have instrumental importance in this context: they can clarify the overall theory of the evolution and the nature of human cognition laid out in the book.

Finally, the applications were chosen to be useful paradigms. They span a wide spectrum and focus on the different major aspects of the account here developed: technology, complex and abstract representational decision-making, and cognitive and social institutions. In this way, other applications of the present account (e.g., to climate change or human conflict) can be modeled on the ones sketched here. The applications discussed in this chapter can therefore also be seen as models for other such applications.

Before beginning the consideration of these applications, though, it is important to emphasize that a detailed treatment of each of them would call for a book of its own—for each of them. Fortunately, such a detailed treatment is not necessary here. The goal in what follows is narrower—it is just to (a) *illustrate* the practical implications of the account here developed and (b) develop this account further *through* the consideration of these applications. For this, a brief treatment of the issues raised here is sufficient.

The chapter is structured as follows. In Section 8.1, I briefly summarize key elements of the picture that emerges from the discussion of the previous five chapters to provide the necessary foundation for the discussion of the applications to follow. In Section 8.2, on this basis, I apply this picture to debates about AI. In Section 8.3, I consider the implications of the account developed in the book for the role of religion in human thought. In 8.4, I consider the implications of the present account for the defensibility of extensive patent regimes. Section 8.5 summarizes the discussion.

8.1 It's Only Human

To provide the expository foundations for the applications of the rest of the chapter, it is now important to note two abstract implications from the theory developed in this book. First, a major overarching lesson from the account here developed is that distinctively human cognition is deeply interwoven with technology and social institutions (including ones concerning cultural learning). Both in its actual execution and its evolution and development, human cognition depends on the availability of appropriate cognitive technology and social institutions. In order to be sophisticated mindreaders, humans require access to cognitive technology like myths to narrow down their innate mindreading expectations. In order to employ abstract and complex moral concepts, they need to be supported by storage devices and social institutions for evaluating and carrying out resolutions among conflicting norms. In order to be extensive traders, they need to be supported by the right technology and social institutions to make the efficient tracking of claims possible. In short, in order to understand how humans think in the ways that they do, we need to know about the social and technological support structures available to them.

Second, though, another overarching lesson here is that this social and technological dependence of human cognition is not mysterious or some deep metaphysical claim about the nature of thought.[2] Rather, it just rests on the need for humans to lessen the costs of relying on abstract and complex representations (in terms of time, attention, and concentration: Lieder & Griffiths,

[2] For more on the latter, see, e.g., Chemero (2009); Clark (1997); Haugeland (1999); Rietveld (2008); Rowlands (2010); Varela et al. (1991).

2020; Schulz, 2018b), and it involves concrete and quite specific processes, such as the storage of information needed for many different representational inferences or the expansion of who can learn from whom. The social and technological dependence of human cognition is also not uniform: depending on the details of the case, some humans rely on it differently from others (e.g., extensive trading is not part of all human cultures), and different forms of distinctively human cognition depend on technological and social support differently (e.g., some forms of mindreading, cultural learning, and communication may have fewer technological and social prerequisites than extensive trading, say).

The key upshot of these two points in the present context is that applying this account of distinctively human cognition is about tracing out the specific social and technological dependencies of different forms of uniquely human thought. It is in these dependencies where the novel implications of the account come out particularly clearly. With this in mind, the next section considers the impact of new forms of technology—specifically, recent versions of AI like ChatGPT—on the shape of distinctively human cognition.

8.2 Artificial Intelligence and Human Cognition

AI—which is here understood as comprising all forms of technology that are designed to solve some or all of the same cognitive problems that humans can solve (Haugeland, 1985; Russell & Norvig, 2009)—is advancing in leaps and bounds. A recent example is the launching of advanced large language models (LLMs) like ChatGPT. These are complex, deep-learning architectures (structured neural networks) trained on a large dataset of linguistic information. They are generally made available as freely accessible webtools or chatbots and can generate texts of different lengths in response to seemingly arbitrary questions. While not radically new in either design or function, the fact that LLMs like ChatGPT have the abilities they have suggests that the frontier of what computers are able to accomplish is reaching new levels of sophistication. The production of grammatically correct text alone was until recently difficult to achieve with computers—not to mention doing so in response to arbitrary queries and in ways that are also semantically coherent. Contemporary AI has far surpassed these boundaries, and there is every indication it will continue to surpass existing boundaries.

These kinds of developments in AI have led scholars to raise a number of questions (see, e.g., Eden et al., 2013; Kurzweil, 2006; Moravec, 1999; Russo, 2023). Is the presence of ever more sophisticated AI fundamentally changing the place of distinctively human cognition on Earth? Will AI ultimately replace human thought—or human living altogether? Will it be conscious, does it have a moral status, and what will its morality be like (Wallach & Allen, 2010)?

Now, there is no doubt that these are fascinating and important questions to consider. However, given the account developed in this book, there is reason to think that the excitement and fascination with answering them—and surrounding "the singularity" (the point at which AI exceeds human intelligence) more generally—masks an important set of more immediate questions. To appreciate this point, it is useful to go back to Searle's (1980) distinction between "strong" and "weak" AI. Strong AI concerns (the attempt to build) technologies that can, in fact, think like humans do; weak AI concerns (the attempt to build) technologies that can accomplish some or all of the tasks humans can accomplish, without, though, necessarily doing so by way of genuine thought. In the philosophical literature on AI, it is typically the strong version that is thought to be the more controversial topic. As Bringsjord and Govindarajulu (2022) put it (emphasis in original): "After all, what *philosophical* reason stands in the way of AI producing artifacts that *appear* to be animals or even humans?"

In fact, though, there are many philosophically interesting issues raised even by contemporary, pre-singularity, weak AI. Indeed, even such restricted forms of AI as ChatGPT can have major implications for how humans think and act and thus for the place of human cognition on the planet. The account of human cognition developed in this book brings this out clearly—and, importantly, it can also help determine which forms of AI are beneficial and which detrimental.

In the first place, given the close interactions between human cognition and technology, it is a direct implication of the account developed in this book that, as our cognitive and social technology changes, this is bound to impact the kinds of thoughts we can have and act on. As made clear in the preceding four chapters, technology changes both the ways in which we can apply our thoughts to the world and the kinds of thoughts we have through our processes of cultural and individual learning. This matters here because LLMs like ChatGPT—and other forms of AI like it—can have a profound impact on exactly this: they make it possible for us to learn things in ways that we could not before, to make inferences that we could not beforehand, and thus to develop technologies in ways that we could not until now.

For example, if I want a summary of what people have thought about any particular topic, it is now easier to obtain this. This makes existing cognitive technology like myths and stories and forms of writing even more powerful: these become easier to use by more people. Computers can help me translate foreign languages, comb through archives of thousands of pages in seconds, and find pertinent information in a wide variety of formats. This has the potential to fundamentally alter the kinds of thoughts we can have: what a person is able to respond to is much less limited now. I can now learn important insights from different cultures much more easily than before. This changes my dispositions and how I interact with others: my mindreading abilities, even of those in other cultures, are increased, cooperative trade relations can be deepened, international moral conflicts have more tools of resolution, and novel scientific and technological advancements become possible (see also Ford et al., 1997; Hoffman et al., 2001). In short, since even weak AI (e.g., in the form of advanced LLMs like ChatGPT) can be a technology of types S and I (from Chapter 4), and since such technology has a profound effect on the ways in which humans think and act, AI has the potential to affect many parts of human living.

However, it is important to understand this point correctly. First, as also noted in the previous section, the claim here is not that technology is so deeply intertwined with human thoughts that the development of even contemporary, weak AI would amount to a fundamental change in the nature of human cognition (as is true of some forms of the extended mind thesis; see, e.g., Chemero, 2009; Clark, 1997; Haugeland, 1999; Rietveld, 2008; Rowlands, 2010; Varela et al., 1991). Rather, the claim here is the more limited one that technology can affect how we express our thoughts and what thoughts we can acquire. To be sure, the latter is an important point, but it need not be given a major metaphysical emphasis.

Second, the claim is not that the current version of ChatGPT (or other systems like it) is *guaranteed* to have all of these positive effects. Such systems often are trained on a given set of data and are therefore more likely to give correct answers to questions concerning data relevantly similar to the training data (Narayanan & Kapoor, 2023). As of the time of this writing, for example, ChatGPT does reasonably well at questions asked about the main character of a novel, but seems to do far less well with recognizing aggressive language in a novel (Kocoń et al., 2023). It is thus not being claimed here that the most common, current weak AI systems must always be able to enhance our thought processes. Rather, the point here is that, within its appropriate domain, there is every reason to think that advanced LLMs like ChatGPT are *overall* able to make a positive contribution to human cognition, given its aims and structure.

Third and relatedly, it is important to note also that the claim here is not that AI will function as a cognitive enhancer for human cognition *across the board*. For example, even future versions of LLMs like ChatGPT, for all their strengths, might not greatly help us adjudicate conflicts among moral norms (which, as noted in Chapter 6, is a major part of distinctively human cognition). Such systems may be able to tell us what others have thought about some particular conflict, what they might think about it, or what benefits and drawbacks of various resolution schemas are. In this sense, they may well be helpful to consult in cases of normative conflict. However, even future versions of LLMs like ChatGPT may not be able to tell us what the best resolution *actually is* or be able to help us collectively *adopt* one particular norm resolution (in the face of the existence of several possible such resolutions). For this, these systems would need to be able to consider the details of exactly the normative conflict we are facing—including our own personal and cultural histories, preferences, and opinions. The best resolution of a normative conflict for us—given our unique perspective on the world—may not be the best resolution for others. Note that this does not imply that there is one "correct" such resolution for every person or culture. As noted in Chapter 6, this is far from an obviously true claim. Rather, the point is just that there is a difference between laying out potential norm resolutions and actually adopting one such. The latter task is not obviously part of the ability set of even advanced future versions of LLMs.

An analogy of these points is in another classic form of human uniqueness: musical appreciation. Music is used in virtually all cultures around the world. As with the other cognitive suites inventoried in Chapter 2, this suite is the result of the interaction of innate high- and low-level psychological capacities for enjoying sounds of certain types (Pinker, 1997) with motivations for cooperating with others as well as culturally developed tools, such as rituals and instruments (Henrich, 2020).[3] However, human music-making has not stayed the same over time but has seen much change, both in structure (canonizing harmonies) and the

[3] For more on the nature of music, see, e.g., Scruton (1983).

tools used to create it (the creation of the piano, say) (Killin, 2018; Mithen, 2005).

A key innovation in this context was the invention of drum computers and synthesizers (see, e.g., Burgess, 2014): in the past 50 years, it has become possible to create complex, arbitrarily precise drumbeats in a matter of a seconds and create sounds that have never been heard before. This has fundamentally changed the kind of music that exists: tonalities can be created and heard that, before the advent of these technologies, were beyond the capabilities of humans (e.g., perfect synchronization across vast numbers of instruments has become possible, and music can be created out of raw waveforms).

However, these innovations have left many other parts of music-making intact—even ones closely related to the technology. So, while it is true that drum computers have made it easy to automate the tasks of keeping the beat for the other musicians or making a beat pronounced enough so that people have an easy time dancing to it, other aspects of the job of a human drummer are beyond the capabilities even of contemporary drum computers and synthesizers. Drummers often function as the real-time arrangers, producers, and conductors of a group of musicians. If aspects of a musical performance are sought to be changed in the moment—for example, in response to environmental circumstances (as when a solo is prolonged due to the crowd's reaction)—it is often the drummer who leads these changes from the back. As the instrument that keeps the other instruments together, drummers can be like sheepdogs that corral the herd in new directions as circumstances require. Drum computers are not helpful for this. Of course, this is not to say that it will always be beyond the capacities of computers to do this: beyond the singularity, it is likely that AI systems can improvise better than humans can. The point here is just that the fact that existing drum computers can do some of what a drummer used to do does not entail that they can do all of it.

The general lesson behind these points is that the fact that even though contemporary, weak forms of AI can change human cognition drastically in some regards does not mean that it affects the latter in *all* regards. Indeed, something stronger holds here: even pre-singularity, AI need not be a cognitive enhancer at all but can actually *hinder* human cognition. This is the second key implication of the theory of distinctively human cognition developed in this book for the contemporary debates surrounding the role of AI in the future of humanity. More specifically, the theory developed in this book can make clearer which forms of contemporary AI have a net *positive* impact on distinctively human cognition and those that have a net *negative* such impact. This distinction turns on the extent to which AI can function as technology of types S and I.

In particular, helpful forms of contemporary, weak AI are those that enhance cultural learning, mindreading, and moral cognition (to name a few). Music is again a great example: due to wide availability of instructional content on the internet, more people are able to become proficient musicians. This thus massively increases the creative potential of humankind. Forms of AI—interestingly, this is likely to include advanced LLMs like ChatGPT, warts and all (Narayanan & Kapoor, 2023)—likely fall into this category, too. They make it easier for more people to learn about more things. Now, as noted above, as such, these forms of AI do not help people figure out what to think about something or what questions are worth pursuing. However, they are at least able to be highly efficient stores of information that can be easily searched and accessed. In this way, these forms of AI are useful for expanding the cognitive horizons of humans.

By contrast, less helpful forms of AI are those that *limit* cultural learning, mindreading, and moral cognition (to name a few). High on this list are deep learning algorithms designed to determine our preferences about some domain (politics, say, or movies) and then provide us with more of that type of thing (more ads from politically likeminded candidates, say, or more movies of the type I have watched before). This is negative because, instead of enhancing our information store, it shrinks it: the AI is used not to reach information sources I could not reach before, but to close off information stores I used to access. Instead of helping us to appreciate perspectives different from our own, it entrenches us into our own views of the world.

The important point about these implications is that, while they are often appreciated in passing, they are now placed in a systematic framework. The present account of distinctively human cognition can be used as a principled way to assess AI systems for their usefulness. The issue comes down to whether the technology enhances our capabilities for cultural learning, mindreading, and moral cognition or hinders them. While even contemporary ("weak") AI has the potential to have a positive impact here, it also has the potential to have a negative such impact. How the balance comes out will differ on a case-by-case basis, but the key point to note is that, given the nature and evolution of distinctively human cognition, we do not need to await the "singularity" for AI to have important implications for human thought—it does so already. Importantly also, the present account provides a principled structure for classifying which of these implications are positive and which negative.

8.3 Religion and Human Cognition

Religion is a key part of human living. As noted in Chapter 2, it is another cognitive and behavioral suite that is often considered a key example of something that is distinctively human.[4] Given its importance, it is thus not surprising that much work has gone

[4] For example, in archeology, it is used as key evidence for when a transition to behavioral modernity has taken place: Mithen (1999), Sterelny (2012a); see also Trinkhaus & Shipman (1993).

into explaining exactly what it means to think and act in religious ways, how and why humans do it, and what some of its social consequences are (see, e.g., Atran, 2002; Boyer, 2001; Dunbar, 2022; Henrich, 2020; Wilson, 2002; see also the seminal Weber, 1958). The work in this area is sprawling and multifaceted; however, for present purposes, the key thing to note is that its emphasis is typically on explaining why and how humans think in religious terms—which may include paying attention to the socionormative consequences of doing so (for a good overview, see, e.g., Bulbulia, 2004; see also Dunbar, 2022; Johnson et al., 2015). What the account of distinctively human cognition developed in this book shows, though, is that religious behavior interacts with a wide variety of aspects of human thought and action. By paying attention to these interactions, overlooked phenomena and interventions become visible.

To see this, note that one key approach in the naturalistic explanation of religion is based on the idea that many aspects of religious thinking can be seen as the result of innate representations of agency that are being "triggered" in circumstances for which they have not evolved (Atran, 2002; Boyer, 2001). This might lead us to interpret thunderstorms as the expression of the anger of a deity (Zeus, say). The details of this kind of explanation differ across the different theories in this area; what matters in the present context, though, is just its general outlines.

Now, it may at first seem that this account of religion based on the erroneous triggering of innate expectations of agency is inconsistent with the picture laid out in this book. After all, it may seem that, on the account of Chapters 4 and 5, mindreading expectations are said to have evolved precisely to avoid erroneous triggering of the kind that is meant to underlie religious cognition. Put differently: Is it not the case that the very reason why it is plausible to posit innate mindreading representations is that it is only through these representations that efficient interactions with our environment become possible? If so, though, it may be very puzzling how these expectations can be erroneously triggered by a thunderstorm (say).

In response, two points can be noted. On the one hand—and this is the point emphasized by Boyer (2001)—innate expectations need not be 100% correct all the time. In general, human minds are flawed, and it would not be unexpected if, on occasion, our mindreading expectations, say, got triggered in the (evolutionarily) wrong circumstances. On the other hand—and this is a point much emphasized in the literature on gene–culture coevolutionary theory (Boyd & Richerson, 2005; Henrich & McElreath, 2007)—cultural learning can override biological biases. We may have an innate motivation to have sex, but, in the right circumstances, celibacy can be a culturally stable institution. The same could be true here: while we may initially not think it compelling to interpret a thunderstorm as an expression of Zeus's rage, cultural pressures can make us override this aversion. For these reasons, there is no deep inconsistency between the present account of distinctively human cognition and the "erroneous triggering of innate expectations" account of religion.

What is more noteworthy is that the traditional picture of the evolutionary psychology of religions is limited in its focus. It explains religion by noting its *triggering* of innate expectations. Other accounts of religion, while differing in focus, are similarly limited: they explain religious thoughts and actions as culturally adaptive—for example, for enabling decreases in in-group competition or increases in in-group cooperation (Henrich, 2020; Johnson et al., 2015; Wilson, 2002).[5] However, the account of the evolution of distinctively human cognition of this book implies that there are several further connections between religion and human cognition that deserve to be noticed. In particular, since religion is a store of myths and stories, and thus a form of cognitive and social technology—a point widely accepted in the literature (Atran, 2002;

[5] These accounts are clearly consistent: see, e.g., Atran and Henrich (2010).

Henrich, 2020; Sterelny, 2021)—it can affect many other aspects of human representational cognition as well. Three such are particularly noteworthy here.

First, it is not just the case that religion is the *result* of our mindreading expectations. Rather, as also briefly hinted at in Chapter 5, religion can also impact how we use our mindreading expectations to make decisions: as illustrated by Figure 5.3, religious stories and institutions aid and guide ordinary mindreading inferences. Importantly, this is true not just in the space of religious thought, but also more generally as well. So, while some scholars have noted that people in different cultures "hear voices" differently, depending on the mindreading expectations prevalent in their culture (Luhrmann et al., 2015), it is also the case that people's expectations about what a divinity wants or thinks (say) can impact how they use their mindreading expectations to infer what other humans are thinking. By being told to ask, "What would Jesus do?" we are taught to ascribe stable character traits to people. By being told that "All things, conditioned or unconditioned, are void of self and soul" (Buddha), we are taught not to always attribute praise and blame to stable character traits, rather than features of the situation.[6] Note that this is not to say that these sorts of religious prescriptions are the only things that guide our mindreading inferences or that this is all that these religious prescriptions do. The point here is just that this is one of the things that they do.

The second key implication of the present account is that religion can affect our causal and technological cognition. In a much discussed and replicated finding of cognitive science, it has been shown that understanding the theory of evolution and accepting it are uncorrelated: people can understand the theory and still fail to accept it (Lombrozo et al., 2006, 2008). However, this is in fact to be

[6] Buddhist moral thinking is complex and also includes virtue ethical elements. The point, though, is that it is generally less character-focused than some other religions. See, e.g., Harvey (2000).

predicted on the present account: what mental representations we have available is one thing; which ones we use with is another. In particular, as made clear by Figure 4.3, social and cognitive tools—which, importantly, includes religion—can narrow down the set of representational inferences in certain ways. This explains why a more thorough teaching of the evolutionary theory need not lead to its wider acceptance: the issue is not having access to a representation; the issue is using it. This is something that is predicted and explained by the present account.

The third noteworthy interaction between religion and human representational cognition concerns property cognition. As noted in Chapter 7, the extent to which humans are inclined to view others as having property claims on something depends on the extent to which there are sociocognitive tools available with which to underwrite such attributions. One of the classic disputes in this context concerns the extent to which "mixing one's labor" with a natural resource makes that resource one's property (Blake & Harris, 2009; Locke, 1987 [1866]; Otsuka, 2003). Given the account of distinctively human cognition laid out in this book, it can be predicted that religion can influence a person's view on property: the myths and models provided by religion can highlight the circumstances in which property can be acquired. So, a religion emphasizing that God "owns" his creation can lead to different property claims than one that emphasizes that creating does not lead to ownership (Waldron, 1990). Importantly, the claim is not just that religion can affect people's view on who *deserves* to own what—though this may be true, too, as noted in Chapter 6. Rather, the claim is that it can influence who *in fact* owns what. While the latter claim has not, in fact, been empirically established yet, the present account at least brings its interest and theoretical plausibility clearly into view.

This is indeed the overarching lesson from this section. The interactionist account of human cognition defended in this book unifies phenomena that have been before thought to be very disparate (mindreading, scientific thought, property cognition), and it brings phenomena into view that would otherwise have been missed.

8.4 Patents and Human Cognition

In economics, it is still quite common to see extensive patent rights as a necessary motivator of innovation. That is, it is often thought that, without patent protections, economic actors lack the means to profit from their innovative activity, are thus not incentivized to engage in this innovative activity, and will therefore not in fact engage in this activity (Aghion et al., 2001, 2005; Aghion & Howitt, 1992; Karbowski & Prokop, 2013; Kesan, 2015). However, as the account of distinctively human cognition defended in this book makes clear, there are some compelling reasons for thinking that this material interest–based view of innovative activity is unconvincing.[7]

Patent protections are monopoly rights to sell a given type of good or service—typically for a fixed time period (Horowitz & Lai, 1996; Klemperer, 1990; Nordhaus, 1969). They are initially awarded to the inventors or discoverers of the good or service in question, but they are increasingly treated as commodities to be traded (an important point to which I return momentarily) (Kesan, 2015; Sichelman, 2010; Teece, 1986). Since their inception and development, a number of issues surrounding the issuance of patent protections have been raised.

The first of these issues is that there is no question that patent protections have downsides. On the one hand and most directly, in virtue of the fact that they create monopolies, patent protections

[7] The nature of economic motivation in general has, of course, long been a point of contention (Angner, 2018, 2016; Fehr & Camerer, 2007; Fehr & Fischbacher, 2003; Fehr & Gaechter, 2000; Fehr & Schmidt, 1999; Hausman, 1992, 2012; Henrich et al., 2005; Skyrms, 2004). However, the question of what motivates *innovation* raises its own set of issues; these will be in focus here.

create inefficiencies (Mas-Colell et al., 1995). By definition, such protections reduce the supply of a given good or service and thus raise its price. In turn, this implies that (a) some economic actors that would otherwise have access to the good or service are now prevented from doing so, and (b) the economic actors that are still getting access to the good or service now have to pay a higher price for doing so. All in all, this creates a "deadweight loss": while some of the losses in (a) and (b) are compensated for by gains in the profits of the patent holders, not all of them are. Hence, possible gains from trade are lost through patent protections (Boldrin & Levine, 2013; Gifford, 2004; Klemperer, 1990; Nordhaus, 1969). On the other hand, patent protections—especially if they are very broad—can make it too difficult for new inventors to enter into a given product market and thus prevent beneficial new inventions from getting made (Aghion et al., 2001, 2005; Aghion & Howitt, 1992; Horowitz & Lai, 1996; Muthukrishna & Henrich, 2016; Nelson et al., 2018; Nordhaus, 1969).[8]

Given these downsides of patent protections, why then are they thought to be useful at all? Several different—not necessarily mutually inconsistent—answers have been given to this question (Karbowski & Prokop, 2013). Some have suggested that patent protections are a consequence of the fact that discoverers or inventors have natural (intellectual) property rights over their discoveries or inventions (Shiffrin, 2001; Shrader-Frechette, 2006; for discussion, see, e.g., Damstedt, 2002; Otsuka, 2003). (Note that this ties this point to the discussion of religion in the previous section.) Others have noted that patents can lead to an economy functioning more efficiently as inventions become more

[8] This was part of what led Schumpeter (1942) to worry about patents leading to the collapse of the entrepreneurial system. Of course, as will also be made clearer below, it is also the case that patents are thought to spur innovation (Aghion et al., 2005; Schumpeter, 1959); the point in the text is just that, if patents prevent an innovator from producing anything at all like a given product, it may make the creation of incremental improvements of this product overly difficult.

easily tradable or commercializable (Kesan, 2015; Penin, 2005; Sichelman, 2010; Teece, 1986).

However, the central justification of the need for extensive patent regimes—and the one that is key in this context—is that patents are often seen as a major motivational tool. Specifically, patent protections are typically thought to be necessary to incentivize innovation (see, e.g., Karbowski & Prokop, 2013; Kesan, 2015; Nordhaus, 1969; Torrance & Tomlinson, 2009). The term "innovation" here includes all kinds of R&D activity: anything from the discovery of new medicines to the invention of a new process for building some of the components that go into the assembly of a particular kind of machinery. The core idea behind this argument for patent rights is that R&D activities can be very expensive to conduct and tend to be uncertain in terms of their payoff: a company might need to expend significant resources in the development of a new drug even though it is quite likely that this drug ends up not getting regulatory approval or failing to come to market for other reasons. If it then were the case that companies would face the possibility that even successful innovations would almost immediately face direct competition, thus driving their price down to just cover their marginal costs (which are also often near-zero), they would have no way to recoup the costs of failed innovations. In turn, this would mean that they have no profit incentive to engage in the innovation activities to begin with.

Now, it needs to be noted immediately that a number of challenges to this motivation-based argument exist. For example, it has been noted that changes in a country's patent regime often do not lead to concomitant increases in the country's growth or innovation rate (Branstetter & Sakakibara, 2001; Qian, 2007). However, these historical challenges (so far) failed to dethrone the motivational argument for patent regimes from its central position in the literature. On the one hand, this is due to the fact that the historical situation is in fact quite ambiguous: for example, while some studies fail to find a positive effect of patent regimes on innovation rates, other studies do find such positive effects (Hu & Png, 2013; Moser, 2005). On the other hand, these studies can be hard to interpret as changes in the patent regimes of a given country often come with other changes as well (such as changes in the tax regimes or broader legal environment). In turn, this makes it hard to disentangle whether and to what extent changes in the country's growth rate (or a lack thereof) are due to the changes in the patent regimes or to some of the other changes that occurred.

This point is strengthened by the fact that a more charitable reading of the traditional argument for patent protections requires acknowledging that not *all* the motivational sources of innovative activities need to lie in a profit motive. The claim made by defenders of the classic argument for patent protections should really just be seen to be that a *significant* proportion of the motivation for innovation has its source in a search for material profits. After all, it can be taken for granted that humans (like most other organisms) are motivated by *both* external, physical rewards—such as monetary gains—and by other considerations—such as social status or curiosity.[9] The question really is how important each of these is. Defenders of the traditional argument for patent protections think the answer is: the profit motive is crucial. Without it, only insignificant amounts of R&D would be completed—humans are *mostly* driven by the promise of direct or indirect external rewards.

In this context, what the account of uniquely human cognition laid out in this book shows is that it is plausible that extensive patent protections are *not* needed to motivate even significant

[9] Social status, for example, is widely accepted to be a major human motivator, too (Barkow, 1992; Hayden, 1998; Henrich, 2015; Skyrms, 2010; Sterelny, 2012a; Sugiyama & Sugiyama, 2003). This is especially noteworthy here since significant aspects of the institution of contemporary science are based on the fact that people are keenly driven by social approval (Strevens, 2003; Zollman, 2018). (That said, the issues here are complex because scientists can also obtain patents on some of their inventions and might choose to become university researchers rather than industry ones for external, monetary reasons—such as a job opening near family, a better benefits package, a better work–life balance, etc.; see also Nelson et al., 2018; Torrance & Tomlinson, 2009.)

amounts of R&D. In particular, as noted earlier, curiosity was crucially important in human cognitive evolution: it drove increased cultural learning, technological advancement, and risk-taking—and was in turn enabled by the latter (Bar et al., 2023; Boyd et al., 2011; Henrich, 2015; Henrich & McElreath, 2011, 2007; Heyes, 2018). This makes it very plausible that curiosity is a *frequent and significant human motivator*. Given this, there is good reason to think that it would continue to drive significant amounts of R&D activity even in the absence of guaranteed major external rewards for successful such activity. Of course, this is not to say that external considerations are not also very motivating for humans. However, since cultural learning is universally regarded as very important to human evolution, and since curiosity does seem to drive much behavior of contemporary human infants and adults, there is reason to think that a *strong and frequent* drive for innovation is likely to be a significant part of human psychology.

This matters because it suggests that the traditional argument for the plausibility of extensive patent regimes is built on unconvincing motivational assumptions. Given that humans plausibly are very curious by nature—that is, given that humans have evolved to be internally driven to frequently engage in information-seeking behavior—innovative activity does *not* need to be incentivized by patents and would still remain about as abundant and creative as before. In fact, more than that is true. Given that there is no question that humans are also motivated by external considerations, such as fame and fortune, it is plausible that the stronger these external motivations are (e.g., due to the existence of extensive patent protections) the more likely it is that they will direct innovative R&D activities in the externally rewarded direction. In the latter kinds of cases (i.e., in situations where certain innovations are given extensive patent protections), the distinctively human drive toward innovation and curiosity is coupled to external awards, leading to especially strong motivations to innovate in this particular manner. This, though, can be problematic: it can direct innovative activity away from other targets of innovation that might be more beneficial to human living but which are not associated with external rewards (such as extensive patents) and therefore do not get explored much.

A concrete example of this kind of case is the evolution of antibiotic resistance in bacteria. The question of how to treat these "superbugs" has major public health consequences: there is a danger that ever greater numbers of people will fall ill due to an infection with bacteria that do not respond to any known antibiotic (Ventola, 2015; Yates, 1999). One potential answer to this question centers of bacteriophages: these are viruses that can infect bacteria and kill them (Burrowes et al., 2011; Łusiak-Szelachowska et al., 2022). As of now, the efficacy and safety of treatments based on bacteriophages is still unknown, but there is some anecdotal evidence that they have potential (Burrowes et al., 2011; Łusiak-Szelachowska et al., 2022). However, investigating this potential further is currently institutionally difficult: as naturally occurring parts of the world, bacteriophages cannot be patented (Pathak-Vaidya et al., 2021).[10] This means that researchers have little *external* incentive to focus their work on them: it is largely their curiosity that drives the research here.[11] Instead, medical researchers are directed toward increasingly more complex synthetic treatments of antibiotic resistance. While the latter are certainly also deserving of further work, the point is that the existence of extensive patent regimes skews the incentives here: even though increasingly more complex antibiotics may be less effective or safe than treatments based on bacteriophages, most research focuses on the former and not the latter.[12]

[10] Note that, in turn, this restriction is justified in terms of the principles underlying ownership.

[11] This is "largely" so, as there may well be some external rewards here: fame, say, or the potential to market and profit from ancillary treatments. Still, the point is that these external rewards are significantly lower than in the case of innovations that can be patented—such as standard drug-based treatments.

[12] Another example of this kind of phenomenon is the fact that medical research is skewed away from considering the health benefits of common foodstuffs, which also cannot be patented.

What the account of distinctively human cognition defended in this book suggests, therefore, is that less extensive patent regimes may have the potential to lead to a more even distribution of innovative activity. Far from it being the case that, without extensive patents, R&D would come to a standstill: without extensive patents, innovative activity may become more variegated and thus potentially effective. This does not mean that patents have no role to play in research and development: they might still be important to make the manufacture and production of costly inventions possible (Kesan, 2015; Penin, 2005; Sichelman, 2010; Teece, 1986). It is just that, instead of seeing them as the prime motivator of the latter, they should just be seen as *one* element in the social structures put in place to make innovative activity easier. They can be scaled back and human innovation would not only continue to exist—but continue to thrive.

These arguments are not just important for the narrow question of the cogency of extensive patent regimes. They are also useful to note as they bring out the extent to which the account of distinctively human cognition laid out here has some direct policy implications. In general, the account defended in this book implies that social institutions can help or hinder distinctively human cognition and action. More than that: as was true for the case of AI and technology, the account can be used as a framework with which to assess *which* social institutions enhance human thought and action and which do not.

8.5 The Overall Upshot

All in all, here is where we have ended up. Distinctively human cognition is the result of the combination of a set of innate representations and low-level capacities that are enhanced through culturally learning and sociocognitive technology, with all of these elements interacting in complex ways. The upshot of this is a manner of interacting with the world that, while sharing many these elements with other organisms, is still highly unique. In this way, we can account for both the culturally specific and culturally universal aspects of human cognition. Importantly, this account also has important upshots for many contemporary questions.

For example, it can help clarify which forms of AI are beneficial and which not: those that enhance distinctively human cognition fall in the former category, those that hinder it fall in the latter. Further, makes clearer the relationships between religion and other aspects of human representational decision-making and can help elucidate some of the challenges of interfaith dialogue. Finally, it can show that extensive patent protections are not necessary to motivate significant amounts of R&D activity, contrary to the widely held and influential material interest–based view of innovative activity: the plausibility of deep and strong human curiosity supports and explains the economic claim that such activity would persist even in the absence of guaranteed monopoly profits.

All in all, therefore, this shows that the present account of distinctively human cognition can help guide technological development, point to overlooked sociopsychological phenomena, and aid in the formulation of appropriate social policies. What is left is to summarize the main points of this account; this is the task of the final chapter.

9
Conclusion

What does it mean to think like a human—and why and how did humans end up thinking like this? The theory laid out in this book develops an interactionist answer to these questions that expands and synthesizes the existing work on this topic. It also provides a methodological blueprint through which further inquiries in this area become more easily possible. The following five points are the key takeaways from the development of this theory.

First, human cognition, while not necessarily differing in kind from what is found among non-human animals, is still unique and distinctive in many ways. It can be broken down into a number of different elements: low-level traits (such as fine motor control and high motivation for reciprocal cooperation), high-level traits (expectations for how to mindread others and abstract and complex causal, moral, and property concepts), and suites combining different low- and high-level traits (language, cultural learning, tool use, and innovation). These traits need not be culturally universal in all respects, but they still mark humans as a distinctive type of thinker.

Second, the most compelling account of the evolution and development of distinctively human cognition accepts the arguments from the modular, nativist accounts that efficient interaction with our environments requires the reliable, canalized development of psychological expectations that narrow down the kind of representational inferences we are called upon to make. The account also accepts the idea that making sense of distinctively human cognition cannot be done without acknowledging the major impact that cultural learning plays in human thought and action, as well

It's Only Human. Armin W. Schulz, Oxford University Press. © Oxford University Press 2025.
DOI: 10.1093/9780197800188.003.0009

as the importance of abstract and complex, systematically organized mental representations. However, what is key in setting out a compelling account of the evolution and development of distinctively human cognition is noting that the independent selection pressures on these different elements were sufficiently strong to create a positive feedback loop among them. In turn, this has added further force to these independent selection pressures.

Third, at the heart of this interactionist feedback loop is the fact that cumulative cultural learning can lead to the development of cognitive and social technology that aids in representational decision-making and cultural learning. More sophisticated representational decision-making can then lead to more powerful cultural learning (including a wider set of models and pedagogical techniques) and more complex technology (including storage tools like written language and inferential tools like abacuses and algebra). Through this kind of process, humans can deploy abstract and complex mental state concepts (such as "wishful thinking") in their mindreading, abstract and complex moral concepts (such as "social justice") in their norm-driven social lives, and abstract and complex property concepts (such as "air rights") in their cooperative interactions. All of these concepts can then further enhance human cultural learning and technological development.

Fourth, while the present inquiry focused on certain specific human cognitive traits—mindreading, moral conceptual thought, and trade—its basic approach is easily extended to other areas: music, for example, or humor, art, and religion. Methodologically, the key is to combine what we know about the psychological foundations needed to engage in the relevant tasks—for example, the kinds of representations likely to have evolved to make the relevant cognitions efficiently possible or the low-level traits needed to underwrite the task—with what we know about the sociotechnological enhancements needed to make the reliance on these cognitive traits possible. These sociotechnological enhancements can differ in kind from each other, and they may not even always be equally important; however, it is likely that they have a major impact on human cognition as well as on our practices of cultural learning. By tracing out the resulting feedback loop we can discover new phenomena that we might have otherwise overlooked and have a unifying framework within which otherwise disparate phenomena can be placed.

The final, related point to emphasize about the present account of distinctively human cognition is the fact that it has some important implications for matters of contemporary importance. The first among these are guidelines about which technological innovations are beneficial for humans (e.g., artificial intelligence such as ChatGPT that can make cultural learning easier, as opposed to echo chamber–generating machine learning algorithms designed to elicit our preferences to give us more of what we already want). Another key application concerns the deep impact that religion has on human cognition —an impact that needs to be acknowledged when we want to change people's views (e.g., with regards to accepting the theory of evolution). Finally, the account has implications for the design of social institutions like patent regimes: given the strong inherent drive for innovation in humans, such innovation need not be externally incentivized in a way that imposes major social costs. For these reasons, the theory of distinctively human cognition laid out in this book is not only inherently interesting, but it also has instrumental contemporary importance.

In short, we are only human. However, by understanding what this means, we not only have the chance to get to know ourselves better, but we also get the chance to improve life on this planet—for us and everyone else.

References

Abarbanell, L., & Hauser, M. D. (2010). Mayan morality: An exploration of permissible harms. *Cognition*, *115*(2), 207–224. https://doi.org/10.1016/j.cognition.2009.12.007

Acebo, C., & Thoman, E. B. (1995). Role of infant crying in the early mother-infant dialogue. *Physiology and Behavior*, *57*, 541–547.

Adler, D. (2009). The earliest musical tradition. *Nature*, *460*, 695–696.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., & Howitt, P. (2005). Competition and innovation: An inverted-U relationship. *Quarterly Journal of Economics*, *120*(2), 701–728.

Aghion, P., Harris, C., Howitt, P., & Vickers, J. (2001). Competition, imitation and growth with step-by-step innovation. *Review of Economic Studies*, *68*(3), 467–492.

Aghion, P., & Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, *60*(2), 323–351.

Alexander, J. (2007). *The structural evolution of morality*. Cambridge: Cambridge University Press.

Alexander, J. (2009). Evolutionary game theory. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. http://plato.stanford.edu/archives/fall2009/entries/game-evolutionary/

Allen, C. (1992). Mental content. *British Journal for the Philosophy of Science*, *43*(4), 537–553. https://www.journals.uchicago.edu/doi/10.1093/bjps/43.4.537

Allen, C. (1999). Animal concepts revisited: The use of self-monitoring as an empirical approach. *Erkenntnis*, *51*(1), 537–544.

Allen, C. (2014). Models, mechanisms, and animal minds. *Southern Journal of Philosophy*, *52*, 75–97. https://doi.org/10.1111/sjp.12072

Allen, C. (2018). Associative Learning. In K. Andrews & J. Beck (Eds.), *The Philosophy of Animal Minds* (pp. 401–408). London: Routledge.

Allen, C., & Bekoff, M. (1997). *Species of mind: The philosophy and biology of cognitive ethology*. Cambridge, MA: MIT Press. https://books.google.com/books?id=H97oavF8NO4C

Allen, C., & Hauser, M. (1991). Concept attribution in nonhuman animals: Theoretical and methodological problems in ascribing complex mental processes. *Philosophy of Science*, *58*, 221–240.

Amos, W., & Hoffman, J. I. (2010). Evidence that two main bottleneck events shaped modern human genetic diversity. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1678), 131–137. https://doi.org/doi:10.1098/rspb.2009.1473

Andrews, K. (2012). *Do apes read minds? Toward a new folk psychology*. Cambridge, MA: MIT Press.

Andrews, K. (2015). *The animal mind: An introduction to the philosophy of animal cognition*. London: Routledge.

Andrews, K. (2018). Do chimpanzees reason about belief? In K. Andrews & J. Beck (Eds.), *The Routledge handbook of philosophy of animal minds* (pp. 258–268). London: Routledge.

Andrews, K. (2020). Naïve normativity: The social foundation of moral cognition. *Journal of the American Philosophical Association*, *6*(1), 36–56. https://doi.org/10.1017/apa.2019.30

Andrews, K., Fitzpatrick, S., & Westra, E. (2024). Human and nonhuman norms: A dimensional framework. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *379*(1897), 20230026. https://doi.org/doi:10.1098/rstb.2023.0026

Angner, E. (2016). *A course in behavioral economics* (2nd ed.). London: Palgrave Macmillan.

Angner, E. (2018). What preferences really are. *Philosophy of Science*, *85*(4), 660–681. https://doi.org/10.1086/699193

Anton, S., Potts, R., & Aiello, L. (2014). Evolution of early Homo: An integrated biological perspective. *Science*, *345*(6192), 1236828.

Apicella, C. L., Azevedo, E. M., Christakis, N. A., & Fowler, J. H. (2014). Evolutionary origins of the endowment effect: Evidence from hunter-gatherers. *American Economic Review*, *104*(6), 1793–1805. https://doi.org/http://dx.doi.org/10.1257/aer.104.6.1793

Apperly, I. A. (2011). *Mindreaders: The cognitive basis of "theory of mind."* London: Routledge.

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychology Review*, *116*(4), 953–970. https://doi.org/10.1037/a0016923

Atran, S. (2002). *In gods we trust: The evolutionary landscape of religion*. Oxford: Oxford University Press.

Atran, S., & Henrich, J. (2010). The evolution of religion: How cognitive by-products, adaptive learning heuristics, ritual displays, and group competition generate deep commitments to prosocial religions. *Biological Theory*, *5*(1), 18–30. https://doi.org/10.1162/BIOT_a_00018

Avsar, R. (2020). *The evolutionary origins of markets: How evolution,psychology, and biology have shaped the economy*. New York: Routledge.

Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Science*, *14*(3), 110–118. https://doi.org/10.1016/j.tics.2009.12.006

Bar, M., Baror, S., & Ivancovsky, T. (2023). A shared novelty-seeking basis for creativity and curiosity. *Behavioral and Brain Sciences*, 1–61. https://doi.org/10.1017/S0140525X23002807

Barham, L. (2013). *From hand to handle: The first industrial revolution*. Oxford: Oxford University Press.

Barkow, J. (1992). Beneath new culture is old psychology: Gossip and social stratification. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 627–637). Oxford: Oxford University Press.

Baron-Cohen, S. (1997). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: MIT Press.

Barrett, H. C. (2015). *The shape of thought: How mental adaptations evolve*. Oxford: Oxford University Press.

Barrett, L. (2008). Out of their heads: Turning relational reinterpretation inside out. *Behavioral and Brain Sciences*, *31*(2), 130–131. https://doi.org/10.1017/S0140525X08003555

Barrett, L. (2018). What is human nature (if it is anything at all)? In R. Joyce (Ed.), *The Routledge handbook of evolution and philosophy* (pp. 194–209). London: Routledge.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577–609; discussion 610–560.

Basu, S., Dickhautb, J., Hecht, G., Towry, K., & Waymire, G. B. (2009). Recordkeeping alters economic history by promoting reciprocity. *Proceedings of the National Academy of Science*, *106*(4), 1009–1014.

Basu, S., & Waymire, G. B. (2006). Recordkeeping and human evolution. *Accounting Horizons*, *20*(3), 201–229. https://doi.org/10.2308/acch.2006.20.3.201

Baum, D. A. (2013). Developmental causation and the problem of homology. *Philosophy and Theory in Biology*, *5*(3), 1–13. http://dx.doi.org/10.3998/ptb.6959004.0005.003

Beck, J. (2018). Do non-human animals have a language of thought. In K. Andrews & J. Beck (Eds.), *The Routledge handbook of the philosophy of animal minds* (pp. 46–55). London: Routledge.

Bednarik, R. G. (2008). Beads and cognitive evolution. *Time and Mind: The Journal of Archaeology, Consciousness and Culture*, *I*(3), 285–318.

Beer, R. D. (1990). *Intelligence as adaptive behavior*. New York: Academic Press.

Benbasat, I., & Dexter, A. S. (1982). Individual differences in the use of decision support aids. *Journal of Accounting Research*, *20*(1), 1–11. https://doi.org/10.2307/2490759

Bender, A., & Beller, S. (2016). Probing the cultural constitution of causal cognition: A research program. *Frontiers in Psychology*, *7*(245), 1–6.

Bender, A., & Beller, S. (2019). The cultural fabric of human causal cognition. *Perspectives on Psychological Science*, *14*(6), 922–940. https://doi.org/10.1177/1745691619863055

Bennett, J. (1991). How to read minds in behavior: A suggestion from a philosopher. In A. Whiten (Ed.), *Natural theories of mind: Evolution, development and simulation of everyday mind reading* (pp. 97–108). Hoboken, NJ: Blackwell Publishing.

Bergemann, D., & Välimäki, J. (2008). Bandit problems. In S. N. Durlauf & L. E. Blume (Eds.), *The new Palgrave dictionary of economics: Volume 1–8* (pp. 336–340). London: Palgrave Macmillan UK. https://doi.org/10.1057/978-1-349-95121-5_2386-1

Bigelow, J. C. (1998). Functionalism in social science. In *Routledge encyclopedia of philosophy*. London: Taylor and Francis.

Billy Gibbons. (2023, February 11). *Wikipedia*. https://en.wikipedia.org/wiki/Billy_Gibbons

Binmore, K. (1998). *Game theory and the social contract, vol. II: Just playing*. Cambridge, MA: MIT Press.

Birch, J. (2021). Toolmaking and the Evolution of Normative Cognition. *Biology and Philosophy*, *36*(1), 1–26.

Blake, P. R., & Harris, P. L. (2009). Children's understanding of ownership transfers. *Cognitive Development*, *24*(2), 133–145. https://doi.org/10.1016/j.cogdev.2009.01.002

Bloom, P. (2002). Mindreading, communication and the learning of names for things. *Mind & Language*, *17*, 37–54. https://doi.org/10.1111/1468-0017.00188

Boesch, C. (2005). Joint cooperative hunting among wild chimpanzees: Taking natural observations seriously. *Behavioral and Brain Sciences*, *28*(5), 692–693. https://doi.org/doi:10.1017/S0140525X05230121

Boldrin, M., & Levine, D. (2013). The case against patents. *Journal of Economic Perspectives*, *27*, 3–22.

Boroditsky, L. (2001). Does language shape thought?: Mandarin and English speakers' conceptions of time. *Cognitive Psychology*, *43*, 1–22. https://doi.org/10.1006/cogp.2001.0748

Bowlby, J. (1958). The nature of the child's tie to his mother. *International Journal of Psychoanalysis*, *39*, 350–373.

Boyd, R. (2018). *A different kind of animal: How culture transformed our species*. Princeton, NJ: Princeton University Press.

Boyd, R., & Richerson, P. (1985). *Culture and the evolutionary process*. Chicago: University of Chicago Press.

Boyd, R., & Richerson, P. (2005). *The origin and evolution of cultures*. Oxford: Oxford University Press.

Boyd, R., Richerson, P., & Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences*, *108*(Suppl. 2), 10918–10925. https://doi.org/10.1073/pnas.1100290108

Boyer, P. (2001). *Religion explained: The evolutionary foundations of religious belief*. New York: Basic Books.

Boyer, P. (2023). Ownership psychology as a cognitive adaptation: A minimalist model. *Behavioral and Brain Sciences*, *46*, e323, Article e323. https://doi.org/10.1017/S0140525X22002527

Bramble, D. M., & Lieberman, D. E. (2004). Endurance running and the evolution of Homo. *Nature*, *432*(7015), 345. https://doi.org/10.1038/nature03052

Brandom, R. (1994). *Making it explicit*. Cambridge, MA: Harvard University Press.

Branstetter, L., & Sakakibara, M. (2001). Do stronger patents induce more innovation? Evidence from the 1988 Japanese patent law reforms. *RAND Journal of Economics*, *32*(1), 77–100.

Bringsjord, S., & Govindarajulu, N. S. (2022). Artificial intelligence. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford encyclopedia of philosophy* (Fall 2022 edition.). Stanford, CA: Stanford University Press. https://plato.stanford.edu/archives/fall2024/entries/artificial-intelligence

Brisset, N. (2019). *Economics and performativity: Exploring limits, theories and cases*. London: Routledge.

Brooks, A. S., Yellen, J. E., Potts, R., Behrensmeyer, A. K., Deino, A. L., Leslie, D. E., Ambrose, S. H., Ferguson, J. R., d'Errico, F., Zipkin, A. M., Whittaker, S., Post, J., Veatch, E. G., Foecke, K., & Clark, J. B. (2018). Long-distance stone transport and pigment use in the earliest Middle Stone Age. *Science*, *360*(6384), 90–94. https://doi.org/10.1126/science.aao2646

Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence*, *47*, 139–159.

Brosnan, S. F. (2011). Property in nonhuman primates. *New Directions for Child and Adolescent Development*, *132*, 9–22. https://doi.org/doi: 10.1002/cd.293

Brosnan, S. F., & Beran, M. J. (2009). Trading behavior between conspecifics in chimpanzees. *Journal of Comparative Psychology*, *123*(2), 181–194.

Brosnan, S. F., Grady, M. F., Lambeth, S. P., Schapiro, S. J., & Beran, M. J. (2008). Chimpanzee autarky. *PloS One*, *3*, e1518. https://doi.org/10.1371/journal.pone.0001518

Brosnan, S. F., Jones, O. D., Gardner, M., Lambeth, S. P., & Schapiro, S. J. (2012). Evolution and the expression of biases: Situational value changes the endowment effect in chimpanzees. *Evolution and Human Behavior*, *33*, 378–386. https://doi.org/10.1016/j.evolhumbehav.2011.11.009

Brosnan, S. F., Jones, O. D., Lambeth, S. P., Mareno, M. C., Richardson, A. S., & Schapiro, S. J. (2007). Endowment effects in chimpanzees. *Current Biology*, *17*, 1–4.

Brown, D. E. (1991). *Human universals*. New York: McGraw-Hill.

Brown, R. L. (2018). Animal traditions: What they are, and why they matter. In K. Andrews & J. Beck (Eds.), *The Routledge handbook of philosophy of animal minds* (pp. 362–371). Routledge.

Bulbulia, J. (2004). The cognitive and evolutionary psychology of religion. *Biology and Philosophy*, *19*, 655–686.

Buller, D. (2005). *Adapting minds*. Cambridge, MA: MIT Press.

Burgess, R. J. (2014). *The history of music production*. Oxford: Oxford University Press.

Burrowes, B., Harper, D. R., Anderson, J., McConville, M., & Enright, M. C. (2011). Bacteriophage therapy: Potential uses in the control of antibiotic-resistant pathogens. *Expert Review of Anti-infective Therapy*, *9*(9), 775–785. https://doi.org/10.1586/eri.11.90

Butterfill, S. A., & Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind & Language*, *28*(5), 606–637. https://doi.org/10.1111/mila.12036

Byrne, R. W., & Corp, N. (2004). Neocortex size predicts deception rate in primates. *Proceedings of Biological Science*, *271*(1549), 1693–1699. https://doi.org/10.1098/rspb.2004.2780

Callebaut, W., & Rasskin-Gutman, D. (2005). *Modularity: Understanding the development and evolution of natural complex systems*. Cambridge, MA: MIT Press.

Calvo, P., & Symons, J. (Eds.). (2014). *The architecture of cognition*. Cambridge, MA: MIT Press.

Camp, E. (2004). The generality constraint and categorical restrictions. *Philosophical Quarterly*, *54*(215), 209–231.

Campbell, R., & Kumar, V. (2012). Moral reasoning on the ground. *Ethics*, *122*(2), 273–312.

Carey, S. (2011). *The origin of concepts*. Oxford: Oxford University Press.

Carey, S., & Spelke, E. (1996). Science and core knowledge. *Philosophy of Science*, *63*, 515–533.

Carnesale, A., Doty, P., Hoffmann, S., Huntington, S. P., Nye, J. S., & Sagan, S. D. (1983). *Living with nuclear weapons*. Cambridge, MA: Harvard University Press.

Carruthers, P. (1998). Distinctively human thinking. In P. Carruthers & J. Boucher (Eds.), *Language and thought* (pp. 69–88). Cambridge: Cambridge University Press.

Carruthers, P. (2006). *The architecture of the mind*. Oxford: Oxford University Press.

Carruthers, P. (2013). The distinctively human mind: The many pillars of cumulative culture. In G. Hatfield & H. Pittman (Eds.), *Evolution of mind, brain, and culture* (pp. 325–346). Philadelphia: University of Pennsylvania Press. https://doi.org/doi:10.9783/9781934536605

Carter, G., & Wilkinson, G. (2013). Food sharing in vampire bats: Reciprocal help predicts donations more than relatedness or harassment. *Proceedings of the Royal Society B*, *280*, 20122573.

Chater, N. (2012). Building blocks of human decision making. In P. Hammerstein & J. R. Stevens (Eds.), *Evolution and the mechanisms of decision making* (pp. 53–68). Cambridge, MA: MIT Press.

Chater, N., & Christiansen, M. H. (2010). Language acquisition meets language evolution. *Cognitive Science*, *34*(7), 1131–1157. https://doi.org/10.1111/j.1551-6709.2009.01049.x

Cheke, L. G., Bird, C. D., & Clayton, N. S. (2011). Tool-use and instrumental learning in the Eurasian jay (Garrulus glandarius). *Animal Cognition*, *14*(441). https://doi.org/. https://doi.org/10.1007/s10071-011-0379-4

Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, MA: MIT Press.

Chomsky, N. (1968). *Language and mind*. New York: Harcourt Brace & World.

Chomsky, N. (1988). *Language and problems of knowledge (the Managua Lectures)*. Cambridge, MA: MIT Press.

Chudek, M., Zhao, W., & Henrich, J. (2013). Culture-gene coevolution, large-scale cooperation and the shaping of human social psychology. In K. Sterelny, R. Joyce, B. Calcott, & B. Fraser (Eds.), *Cooperation and its evolution* (pp. 425–458). Cambridge, MA: MIT Press.

Churchland, P. (2011). *Braintrust: What neuroscience tells us about morality*. Princeton, NJ: Princeton University Press.

Clark, A. (1997). *Being there*. Cambridge, MA: MIT Press.

Clark, A. (2008). *Supersizing the mind*. Oxford: Oxford University Press.

Clark, A. (2013). Whatever next: Predictive brains, situated agents, and the future of cognitive science. *Behavior and Brain Science*, *36*(3), 1–73.

Clatterbuck, H. (2018). The logical problem and the theoretician's dilemma. *Philosophy and Phenomenological Research*, *97*(2), 322–350.

Colagè, I., & d'Errico, F. (2018). Culture: The driving force of human cognition. *Topics in Cognitive Science*. https://doi.org/10.1111/tops.12372

Conard, N., Malina, M., & Münzel, S. (2009). New flutes document the earliest musical tradition in southwestern Germany. *Nature*, *460*, 737–740.

Conard, N., Malina, M., Münzel, S., & Seeberger, F. (2004). Eine Mammutelfenbeinflöte aus dem Aurignacien des Geißenklösterle: Neue Belege für eine Musikalische Tradition im Frühen Jungpaläolithikum auf der Schwäbischen Alb. *Archäologisches Korrespondenzblatt*, *34*, 447–462.

Cook, C., Goodman, N. D., & Schulz, L. E. (2011). Where science starts: Spontaneous experiments in preschoolers' exploratory play. *Cognition*, *120*(3), 341–349. https://doi.org/10.1016/j.cognition.2011.03.003
Coolidge, F. L., & Wynn, T. (2009). *The rise of Homo sapiens: The evolution of modern thinking*. Malden, MA: Wiley-Blackwell.
Cosmides, L., Guzmán, R. A., & Tooby, J. (2018). The evolution of moral cognition. In A. Zimmerman, K. Jones, & M. Timmons (Eds.), *The Routledge handbook of moral epistemology* (pp. 174–228). London: Routledge.
Cosmides, L., & Tooby, J. (1987). From evolution to behavior: Evolutionary psychology as the missing link. In J. Dupre (Ed.), *The latest on the best: Essays on evolution and optimality* (pp. 277–306). Cambridge, MA: MIT Press.
Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163–228). Oxford: Oxford University Press.
Cowie, F. (2003). *What's within? Nativism reconsidered*. Oxford: Oxford University Press.
Creanza, N., Fogarty, L., & Feldman, M. W. (2012). Models of cultural niche construction with selection and assortive matings. *PloS One*, *7*(8), e42744.
Cross, I. (2012). Music as an emergent exaptation. In N. Bannan (Ed.), *Music, language and human evolution* (pp. 263–276). Oxford: Oxford University Press.
Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 1149–1157. https://doi.org/doi:10.1098/rstb.2010.0319
Currie, G. (2011). The master of the Masek beds: Handaxes, art, and the minds of early humans. In E. Schellekens & P. Goldie (Eds.), *The aesthetic mind: Philosophy & psychology* (pp. 9–31). Oxford: Oxford University Press.
Curry, O. S. (2016). Morality as cooperation: A problem-centred approach. In T. K. Shackelford & R. D. Hansen (Eds.), *The evolution of morality* (pp. 27–51). New York: Springer.
Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment: Testing three principles of harm. *Psychological Science*, *17*(12), 1082–1089. https://doi.org/10.1111/j.1467-9280.2006.01834.x
Damstedt, B. G. (2002). Limiting Locke: A natural law justification for the fair use doctrine. *Yale Law Journal*, *112*(5), 1179–1221.
Darby, D. (2003). Feinberg and Martin on human rights. *Journal of Social Philosophy*, *34*(2), 199–214.
Darby, D. (2004). Rights externalism. *Philosophy and Phenomenological Research*, *68*(3), 620–634.

Davies, S. (2018). Evolution, aesthethics, and art: An overview. In R. Joyce (Ed.), *The Routledge handbook of evolution and philosophy* (pp. 359–371). London: Routledge.

de Hevia, M. D., Izard, V., Coubart, A., Spelke, E. S., & Streri, A. (2014). Representations of space, time, and number in neonates. *Proceedings of the National Academy of Sciences*, *111*, 4809–4813.

de Waal, F. B. (1996). *Good natured: The origins of right and wrong in humans and other animals*. Cambridge, MA: Harvard University Press.

de Waal, F. B. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology*, *59*, 279–300. https://doi.org/10.1146/annurev.psych.59.103006.093625

Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain*. New York: Norton.

DeAngelo, G., & Brosnan, S. F. (2013). The importance of risk tolerance and knowledge when considering the evolution of inequity responses across the primates. *Journal of Economic Behavior & Organization*, *90S*, S105–S112.

Delton, A. W., & Krasnow, M. M. (2015). Adaptationist approaches to moral psychology. In *The moral brain: A multidisciplinary perspective* (pp. 19–34). Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/9988.001.0001

Deng, H., & Zhou, J. (2016). "Juggling" behavior in wild hainan gibbons, a new finding in nonhuman primates. *Scientific Reports*, *6*(1), 23566. https://doi.org/10.1038/srep23566

Dennett, D. C. (1995). *Darwin's dangerous idea*. Cambridge, MA: MIT Press.

Dennett, D. C. (2000). Making tools for thinking. In D. Sperber (Ed.), *Metarepresentation: A multidisciplinary perspective* (pp. 17–29). Oxford: Oxford University Press.

DeScioli, P., & Kurzban, R. (2013). A solution to the mysteries of morality. *Psychological Bulletin*, *139*(2), 477–496.

Desmond, H., & Ramsey, G. (Eds.). (2023). *Human success: Evolutionary origins and ethical implications*. Oxford: Oxford University Press.

Desrochers, T. M., Burk, D. C., Badre, D., & Sheinberg, D. L. (2016). The monitoring and control of task sequences in human and non-human primates [review]. *Frontiers in Systems Neuroscience*, *9*. https://doi.org/10.3389/fnsys.2015.00185

Devaine, M., Hollard, G., & Daunizeau, J. (2014). Theory of mind: did evolution fool us? *PloS One*, *9*(e87619).

Donald, M. (2001). *A mind so rare: The evolution of human consciousness*. New York: W W Norton.

Donnelly, J. (1999). The social construction of international human rights. In T. Dunne & A. N. J. Wheeler (Eds.), *Human rights in global politics* (pp. 71–102). Cambridge: Cambridge University Press.

Doris, J. (1998). Persons, situations, and virtue ethics. *Noûs*, *32*(4), 504–530.

Dorsey, D. (2016). *The limits of moral authority*. Oxford: Oxford University Press.

Downes, S. (2018). Human nature: An overview. In R. Joyce (Ed.), *Routledge handbook of evolution and philosophy* (pp. 155–166). London: Routledge.

Downes, S., & Machery, E. (Eds.). (2013). *Arguing about human nature*. London: Routledge.

Drayton, L. A., Brosnan, S. F., Carrigan, J., & Stoinski, T. S. (2013). Endowment effect in gorillas (Gorilla gorilla). *Journal of Comparative Psychology*, *127*, 365–369. https://doi.org/10.1037/a0031902

Dretske, F. (1988). *Explaining behavior*. Cambridge, MA: MIT Press.

Dunbar, R. (2022). *How religion evolved: And why it endures*. Oxford: Oxford University Press.

Dupre, J. (2001). *Human nature and the limits of science*. Oxford: Oxford University Press.

Eden, A., Moor, J., Soraker, J., & Steinhart, E. (Eds.). (2013). *Singularity hypotheses: A scientific and philosophical assessment*. New York: Springer.

Edwards, K., & Low, J. (2017). Reaction time profiles of adults' action prediction reveal two mindreading systems. *Cognition*, *160*, 1–16. https://doi.org/10.1016/j.cognition.2016.12.004

Ekman, P. (1989). The argument and evidence about universals in facial expressions of emotion. In H. Wagner & A. Manstead (Eds.), *Handbook of social psychophysiology* (pp. 143–164). New York: Wiley.

Ekman, P., & Rosenberg, E. (Eds.). (1997). *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*. Oxford: Oxford University Press.

El Mouden, C., Andre, J. B., Morin, O., & Nettle, D. (2014). Cultural transmission and the evolution of human behaviour: A general approach based on the Price equation. *Journal of Evolutionary Biology*, *27*(2), 231–241. https://doi.org/10.1111/jeb.12296

Elster, A., & Gelfand, M. J. (2021). When guiding principles do not guide: The moderating effects of cultural tightness on value-behavior links. *Journal of Personality*, *89*(2), 325–337. https://doi.org/10.1111/jopy.12584

Elster, J. (1979). *Ulysses and the sirens*. Cambridge: Cambridge University Press.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist*, *49*(8), 709–724.

Eswaran, M., & Neary, H. M. (2014). An economic theory of the evolutionary emergence of property rights. *American Economic Journal: Microeconomics*, *6*(3), 203–226. https://doi.org/10.1257/mic.6.3.203

Everett, D. (2017). *How language began: The story of humanity's greatest invention*. New York: Norton.

Fabry, R. E. (2017). Cognitive innovation, cumulative cultural evolution, and enculturation. *Journal of Cognition and Culture*, *17*(5), 375–395.

Fabry, R. E. (2023). Narrative scaffolding. *Review of Philosophy and Psychology*, *14*(4), 1147–1167. https://doi.org/10.1007/s13164-021-00595-w

Fehr, E., & Camerer, C. (2007). Social neuroeconomics: The neural circuitry of social preferences. *Trends in Cognitive Science*, *11*(10), 419–427. https://doi.org/10.1016/j.tics.2007.09.002

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism [review article]. *Nature*, *425*, 785–791. https://doi.org/10.1038/nature02043

Fehr, E., & Gaechter, S. (2000). Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives*, *14*, 159–181.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, *114*(3), 818–868.

Felsenstein, J. (1985). Phylogenies and the comparative method. *American Naturalist*, *125*, 1–15.

Felsenstein, J. (2004). *Inferring phylogenies*. Sunderland: Sinauer Associates.

Fenici, M., & Zawidzki, T. W. (2021). The origins of mindreading: How interpretive socio-cognitive practices get off the ground. *Synthese*, *198*(9), 8365–8387.

Field, A. (2002). Review of Ofek (2001), Second Nature. *Journal of Economic History 62*(03), 922–924.

Flack, J. C., & de Waal, F. B. (2000). "Any animal whatever": Darwinian building blocks of morality in monkeys and apes. *Journal of Consciousness Studies*, *7*(1-2), 1–29.

Flemming, T. M., Jones, O. D., Mayo, L., Stoinski, T., & Brosnan, S. F. (2012). The endowment effect in orangutans. *International Journal of Comparative Psychology*, *25*, 285–298.

Fodor, J. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.

Fodor, J. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.

Fodor, J. (1990). *The theory of content*. Cambridge, MA: MIT Press.

Fodor, J., & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*, 3–71.

Fogarty, L., & Creanza, N. (2017). The niche construction of cultural complexity: Interactions between innovations, population size and the environment. *Philosophical Transactions of the Royal Society London B Biological Sciences*, *372*(1735), 20160428. https://doi.org/10.1098/rstb.2016.0428

Ford, K. M., Glymour, C., and Hayes, P. (1997). On the other hand . . . cognitive prostheses. *AI Magazine*, *18*(3), 104.

Forster, M., & Sober, E. (2011). AIC scores as evidence: A Bayesian interpretation. In M. Forster & P. S. Bandyopadhyay (Eds.), *The philosophy of statistics* (pp. 535–549). Norwell, MA: Kluwer.

Fragaszy, D. M., Biro, D., Eshchar, Y., Humle, T., Izar, P., Resende, B., & Visalberghi, E. (2013). The fourth dimension of tool use: Temporally enduring artefacts aid primates learning to use tools. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1630), 20120410. https://doi.org/doi:10.1098/rstb.2012.0410

Fraser, B. (2018). The evolution of morality and the prospects for moral realism. In R. Joyce (Ed.), *The Routledge handbook of evolution and philosophy* (pp. 333–341). London: Routledge.

Friederici, A. D. (2017). *Language in our brain: The origins of a uniquely human capacity*. Cambridge, MA: MIT Press.

Friedman, O., Neary, K. R., Defeyter, M. A., & Malcolm, S. L. (2011). Ownership and object history. *New Directions for Child and Adolescent Development*, *2011*(132), 79–89. https://doi.org/10.1002/cd.298

Friedman, O., Van de Vondervoort, J. W., Defeyter, M. A., & Neary, K. R. (2013). First possession, history, and young children's ownership judgments. *Child Development*, *84*(5), 1519–1525. https://doi.org/10.1111/cdev.12080

Gauker, C. (2003). *Words without meaning*. Cambridge, MA: MIT Press.

Gelman, S. A. (2009). Essentialist reasoning about the biological world. In A. Berthoz & Y. Christen (Eds.), *Neurobiology of "Umwelt": How living beings perceive the world* (pp. 7–16). Berlin: Springer. https://doi.org/10.1007/978-3-540-85897-3_2

Gershman, S. J., & Daw, N. D. (2017). Reinforcement learning and episodic memory in humans and animals: An integrative framework. *Annual Review of Psychology*, *68*(1), 101–128. https://doi.org/10.1146/annurev-psych-122414-033625

Geurts, B. (2019). Communication as commitment sharing: Speech acts, implicatures, common ground. *Theoretical Linguistics*, *45*(1-2), 1–30.

Gibson, K. R., & Ingold, T. (Eds.). (1993). *Tools, language and cognition in human evolution*. Cambridge: Cambridge University Press.

Gifford, D. J. (2004). How do the social benefits and costs of the patent system stack up in pharmaceuticals? *Journal of Intellectual Property Law*, *12*, 75–124.

Gigerenzer, G. (2008). *Rationality for mortals*. Oxford: Oxford University Press.

Gilbert, J. (2018). *The forgotten transition: From event to object and from tool to language*. Oud-Turnhout: Gompel & Svacina.

Gintis, H. (2007). The evolution of private property. *Journal of Economic Behavior & Organization*, *64*, 1–16.

Glickman, S. E., & Sroges, R. W. (1966). Curiosity in zoo animals. *Behaviour*, *26*(1–2), 151–187. https://doi.org/10.1163/156853966X00074

Godfrey-Smith, P. (2001). Three kinds of adaptationism. In S. H. Orzack & E. Sober (Eds.), *Adaptationism and optimality* (pp. 335–357). Cambridge: Cambridge University Press.

Godfrey-Smith, P. (2002). On the evolution of representational and interpretive capacities. *Monist*, *85*, 50–69.

Godfrey-Smith, P. (2009). *Darwinian populations and natural selection*. Oxford: Oxford University Press.

Godfrey-Smith, P. (2016). *Other minds: The octopus, the sea, and the deep origins of consciousness*. New York: Farrar, Straus and Grioux.

Godfrey-Smith, P. (2020). Gradualism and the evolution of experience. *Philosophical Topics*, *48*(1), 201–220. https://www.jstor.org/stable/48628592

Goldman, A. (2006). *Simulating minds*. Oxford: Oxford University Press.

Gopnik, A. (1996). The child as scientist. *Philosophy of Science*, *63*(4), 485–514.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychology Review*, *111*(1), 3–32. https://doi.org/10.1037/0033-295X.111.1.3

Gopnik, A., & Schulz, L. E. (2004). Mechanisms of theory-formation in young children. *Trends in Cognitive Science*, *8*(8), 371–377.

Greene, J. (2008). The secret joke of Kant's soul. In W. Sinnott-Armstrong (Ed.), *Moral Psychology* (Vol. 3, pp. 35–79). Cambridge, MA: MIT Press.

Grice, P. (1989). *Studies in the way of words*. Cambridge MA: Harvard University Press.

Griffin, A. S., West, S. A., & Buckling, A. (2004). Cooperation and competition in pathogenic bacteria. *Nature*, *430*, 1024–1027.

Griffiths, P., & Stotz, K. (2000). How the mind grows. *Synthese*, *122*, 29–51.

Guala, F. (2016a). Performativity rationalized. In I. Boldyrev & E. Svetlova (Eds.), *Enacting dismal science: New perspectives on the performativity of economics* (pp. 29–52). New York: Springer.

Guala, F. (2016b). *Understanding institutions: The science and philosophy of living together*. Princeton, NJ: Princeton University Press.

Guan, Z. H., Ma, C. Y., Fei, H. L., Huang, B., Ning, W. H., Ni, Q. Y., Jiang, X. L., & Fan, P. F. (2018). Ecology and social system of northern gibbons living in cold seasonal forests. *Zoological Research*, *39*(4), 255–265. https://doi.org/10.24272/j.issn.2095-8137.2018.045

Guo, K. (2016). Visual attention and facial identification in human and non-human animals. In L. S. Freund, S. McCune, L. Esposito, N. R. Gee, & P. McCardle (Eds.), *The social neuroscience of human-animal interaction* (pp. 33–49). Washington, DC: American Psychological Association.

Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, *20*(1), 98–116. https://doi.org/10.1007/s11211-007-0034-z

Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, *133*(4), 55–66.

Hamlin, J. K. (2015). The infantile origins of our moral brains. In J. Decety & T. Wheatley (Eds.), *The moral brain: A multidisciplinary perspective* (pp. 105–122). Cambridge, MA: MIT Press.

Hammerstein, P. (2003). Why is reciprocity so rare in social animals? A protestant appeal. In P. Hammerstein (Ed.), *Genetic and cultural evolution of cooperation* (pp. 83–94). Cambridge, MA: MIT Press.

Hammerstein, P., & Noe, R. (2016). Biological trade and markets. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1687), 20150101. https://doi.org/10.1098/rstb.2015.0101

Hansell, M., & Ruxton, G. D. (2008). Setting tool use within the context of animal construction behaviour. *Trends in Ecology and Evolution*, *23*(2), 73–78. https://doi.org/10.1016/j.tree.2007.10.006

Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, *59*, 771–786.

Hartley, T. (2019). The continuing evolution of ownership. *PloS One*, *14*(2). https://doi.org/10.1371/journal.pone.0211871

Harvey, P. (2000). *An introduction to Buddhist ethics*. Cambridge: Cambridge University Press.

Haslam, M. (2013). "Captivity bias" in animal tool use and its implications for the evolution of hominin technology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1630), 20120421. https://doi.org/doi:10.1098/rstb.2012.0421

Haugeland, J. (1985). *Artificial intelligence: The very idea*. Cambridge, MA: MIT Press.

Haugeland, J. (1999). Mind embodied and embedded. In J. Haugeland (Ed.), *Having thought* (pp. 207–237). Cambridge, MA: Harvard University Press.

Hausman, D. M. (1992). *The inexact and separate science of economics*. Cambridge: Cambridge University Press.

Hausman, D. M. (2012). *Preference, value, choice, and welfare*. Cambridge: Cambridge University Press.

Hawks, J., Hunley, K., Lee, S.-H., & Wolpoff, M. (2000). Population bottlenecks and pleistocene human evolution. *Molecular Biology and Evolution, 17*(1), 2–22. https://doi.org/10.1093/oxfordjournals.molbev.a026233

Hayashi, F. (2016). Access to electronic payments systems by unbanked consumers. *Federal Reserve Bank of Kansas City Economic Review*, *101*(3), 51–76.

Hayden, B. (1998). Practical and prestige technologies: The evolution of material systems. *Journal of Archaeological Method and Theory*, *5*(1), 1–55. https://doi.org/10.1007/BF02428415

Henrich, J. (1997). Market incorporation, agricultural change, and sustainability among the Machiguenga Indians of the Peruvian Amazon. *Human Ecology*, *25*(2), 319–351.

Henrich, J. (2015). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton, NJ: Princeton University Press.

Henrich, J. (2020). *The WEIRDest people in the world*. New York: Farrar, Straus and Giroux.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In search of Homo economicus: Behavioral experiments in 15 small-scale societies. *American Economic Review*, *91*(2), 73–78.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Henrich, N. S., Hill, K., Gil-White, F. J., Gurven, M., Marlowe, F., Patton, J. G., & Tracer, D. (2005). "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences*, *28*(6), 795–815.

Henrich, J., & McElreath, R. (2007). Dual-inheritance theory: The evolution of human cultural capacities and cultural evolution. In R. Dunbar & L. Barrett (Eds.), *The Oxford handbook of evolutionary psychology* (pp. 555–570). Oxford: Oxford University Press.

Henrich, J., & McElreath, R. (2011). The evolution of cultural evolution. *Evolutionary Anthropology*, *12*, 123–135.

Heyes, C., & Frith, C. (2014). The cultural evolution of mindreading. *Science*, *344*(6190), 1243091.

Heyes, C. M. (2012a). Grist and mills: On the cultural origins of cultural learning. *Philosophical Transactions of the Royal Society B*, *367*, 2181–2191.

Heyes, C. M. (2012b). New thinking: The evolution of human cognition. *Philosophical Transactions of the Royal Society B*, *367*, 2091–2096.

Heyes, C. M. (2013). Imitation: Associative and context-dependent. In W. Prinz, M. Beisert, & A. Herwig (Eds.), *Action science: Foundations of an emerging discipline* (pp. 309–332). Cambridge, MA: MIT Press.

Heyes, C. M. (2014). Rich interpretations of infant behaviour are popular but are they valid? *Developmental Science*, *17*, 665–666.

Heyes, C. M. (2016). Born pupils? Natural pedagogy and cultural pedagogy. *Perspectives on Psychological Science*, *11*(2), 280–295.

Heyes, C. M. (2018). *Cognitive gadgets: The cultural evolution of thinking*. Cambridge, MA: Harvard University Press.

Heyes, C. M., & Galef, B. G. (Eds.). (1996). *Social learning in animals: The roots of culture*. New York: Academic Press.

Hilbe, C., Nowak, M. A., & Sigmund, K. (2013). Evolution of extortion in iterated prisoner's dilemma games. *PNAS*, *110*(17), 6913–6918.

Hinsch, M., & Komdeur, J. (2017). What do territory owners defend against? *Philosophical Transactions of the Royal Society B*, *284*(20162356). https://doi.org/http://dx.doi.org/10.1098/rspb.2016.2356

Hitchcock, C., & Sober, E. (2004). Prediction versus accommodation and the risk of overfitting. *British Journal for the Philosophy of Science*, *55*(1), 1–34. https://doi.org/10.1093/bjps/55.1.1

Hoffman, R. R., Hayes, P. J., & Ford, K. M. (2001). Human-centered computing: Thinking in and out of the box. *IEEE Intelligent Systems*, *16*(5), 76–78.

Hollis, M. (1994). *The philosophy of social science: An introduction*. Cambridge: Cambridge University Press.

Horgan, T., & Timmons, M. (2006). *Metaethics after Moore*. Oxford: Oxford University Press.

Horowitz, A. W., & Lai, E. L. C. (1996). Patent length and the rate of innovation. *International Economic Review*, *37*(4), 785–801. https://doi.org/10.2307/2527311

Hrdy, S. B. (2011). *Mothers and others: The evolutionary origins of mutual understanding*. Cambridge, MA: Harvard University Press.

Hsu, C. W., Zhen, B., Stone, A. D., Joannopoulos, J. D., & Soljačić, M. (2016). Bound states in the continuum. *Nature Reviews Materials*, *1*(9), 16048. https://doi.org/10.1038/natrevmats.2016.48

Hu, A. G. Z., & Png, I. P. L. (2013). Patent rights and economic growth: Evidence from cross-country panels of manufacturing industries. *Oxford Economic Papers*, *65*(3), 675–698. https://doi.org/10.1093/oep/gpt011

Hume, D. (1978 [1739]). A treatise of human nature. In L. A. Selby–Bigge & P. H. Nidditch (Eds.). Oxford: Clarendon Press.

Humphrey, N. (1986). *The inner eye: Social intelligence in evolution*. Oxford: Oxford University Press.

Hutto, D. D., & Myin, E. (2012). *Radicalizing enactivism: Basic minds without content*. Cambridge, MA: MIT Press.

Ismael, J. T. (2011). Self-organization and self-governance. *Philosophy of the Social Sciences*, *41*(3), 327–351. https://doi.org/10.1177/0048393110363435

Jablonka, E., & Avital, E. (2010). *Animal traditions: Behavioural inheritance in evolution*. Cambridge: Cambridge University Press.

Jacob, P., & Scott-Phillips, T. (2021). The cultural evolution of human social cognition: Is mindreading a gadget? *Synthese 199*, 1–27. https://doi.org/10.1007/s11229-020-02620-4

Jaeger, C. B., Brosnan, S. F., Levin, D. T., & Jones, O. D. (2020). Predicting variation in endowment effect magnitudes. *Evolution and Human Behavior*, *41*(3), 253–259. https://doi.org/10.1016/j.evolhumbehav.2020.04.002

Jaffe, A. B., & Lerner, J. (2004). *Innovation and its discontents: How our broken patent system is endangering innovation and progress, and what to do about it*. Princeton, NJ: Princeton University Press.

Jagher, R. (2016). Nadaouiyeh Aïn Askar, an example of Upper Acheulean variability in the Levant. *Quaternary International*, *411*, 44–58. https://doi.org/10.1016/j.quaint.2016.02.006

Johnson, K. A., Li, Y. J., & Cohen, A. B. (2015). Fundamental social motives and the varieties of religious experience. *Religion, Brain & Behavior*, *5*(3), 197–261.

Jones, O. D. (2001). Time-shifted rationality and the law of law's leverage: Behavioral economics meets behavioral biology. *Northwestern University Law Review*, *95*, 1141–1205.

Joyce, R. (2005). *The evolution of morality*. Cambridge, MA: MIT Press.

Kabadayi, C., & Osvath, M. (2017). Ravens parallel great apes in flexible planning for tool-use and bartering. *Science*, *357*(6347), 202–204. https://doi.org/10.1126/science.aam8138

Kahane, G. (2011). Evolutionary debunking arguments. *Noûs*, *45*(1), 103–125.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, *93*, 1449–1475.

Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T.-Y., & Camerer, C. (2009). The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory. *Psychological Science*, *20*, 963–973.

Kanngiesser, P., Rossano, F., Frickel, R., Tomm, A., & Tomasello, M. (2020). Children, but not great apes, respect ownership. *Developmental Science*, *23*(1), e12842. https://doi.org/10.1111/desc.12842

Kanngiesser, P., Santos, L. R., Hood, B. M., & Call, J. (2011). The limits of endowment effects in great apes (Pan paniscus, Pan troglodytes, Gorilla, Pongo pygmaeus). *Journal of Comparative Psychology*, *125*, 436–445. https://doi.org/10.1037/a0024516

Kant, I. (1795 [1996]). Toward perpetual peace. In M. Gregor (Ed.), *The Cambridge edition of the works of Immanuel Kant: Practical philosophy* (pp. 311–352). Cambridge: Cambridge University Press.

Kaplan, H. S., Schniter, E., Smith, V. L., & Wilson, B. J. (2012). Risk and the evolution of human exchange. *Proceedings, Biological Sciences*, *279*(1740), 2930–2935. https://doi.org/10.1098/rspb.2011.2614

Karbowski, A., & Prokop, J. (2013). Controversy over the economic justifications for patent protection. *Procedia Economics and Finance* *5*, 393–402.

Kelly, D. (2018). Moral cheesecake, evolved psychology, and the debunking impulse. In R. Joyce (Ed.), *The Routledge handbook of evolution and philosophy* (pp. 342–358). London: Routledge.

Kelly, L. (2015). *Knowledge and power in prehistoric societies: Orality, memory and the transmission of culture*. Cambridge: Cambridge University Press.

Kesan, J. P. (2015). Economic rationales for the patent system in current context. *George Mason Law Review*, *22*(4), 897–924.

Kidd, C., & Hayden, B. Y. (2015). The psychology and neuroscience of curiosity. *Neuron*, *88*(3), 449–460. https://doi.org/10.1016/j.neuron.2015.09.010

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PloS One*, *7*(5). https://doi.org/e36399

Killin, A. (2018). The origins of music: Evidence, theory, and prospects. *Music & Science*, *1*, 2059204317751971. https://doi.org/10.1177/2059204317751971

Kincaid, H. (1990). Assessing functional explanations in the social sciences. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, *1990*, 341–354. http://www.jstor.org/stable/192715

King, H. C. (1955). *The history of the telescope*. High Wycombe: Charles Griffin & Co.

Kitcher, P. (1985). *Vaulting ambition: Sociobiology and the quest for human nature*. Cambridge, MA: MIT Press.

Kitcher, P. (1998). Psychological altruism, evolutionary origins, and moral rules. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *89*(2/3), 283–316. http://www.jstor.org/stable/4320823

Kitcher, P. (2011). *The ethical project*. Cambridge, MA: Harvard University Press.

Klein, R., & Edgar, B. (2002). *The dawn of human culture*. New York: Wiley.

Klein, S. B., Cosmides, L., Tooby, J., & Chance, S. (2002). Decisions and the evolution of memory: Multiple systems, multiple functions. *Psychology Review*, *109*(2), 306–329.

Klemperer, P. (1990). How broad should the scope of patent protection be? *RAND Journal of Economics*, *21*, 113–130.

Kocoń, J., Cichecki, I., Kaszyca, O., Kochanek, M., Szydło, D., Baran, J., Bielaniewicz, J., Gruza, M., Janz, A., Kanclerz, K., Kocoń, A., Koptyra, B., Mieleszczenko-Kowszewicz, W., Miłkowski, P., Oleksy, M., Piasecki, M., Radliński, Ł., Wojtasik, K., Woźniak, S., & Kazienko, P. (2023). ChatGPT: Jack of all trades, master of none. *arXiv*. https://arxiv.org/abs/2302.10724

Kohn, M., & Mithen, S. (1999). Handaxes: Products of sexual selection. *Antiquity*, *73*, 518–526.

Kumar, V., & Campbell, R. (2022). *A better ape*. Oxford: Oxford University Press.

Kummer, H. (1991). Evolutionary transformations of possessive behavior. *Journal of Social Behavior and Personality*, *6*, 75–83.

Kummer, H., & Cords, M. (1991). Cues of ownership in long-tailed macaques, Macaca fascicularis. *Animal Behavior*, *42*, 529–549.

Kurzweil, R. (2006). *The singularity is near: When humans transcend biology*. New York: Penguin.

Lakshminarayanan, V., Chen, M. K., & Santos, L. R. (2008). Endowment effect in capuchin monkeys. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *363*, 3837–3844. https://doi.org/10.1098/rstb.2008.0149

Laland, K. N., & Janik, V. M. (2006). The animal cultures debate. *Trends in Ecology and Evolution*, *21*, 542–547.

Laland, K. N., & Seed, A. (2021). Understanding human cognitive uniqueness. *Annual Review of Psychology*, *72*(1), 689–716. https://doi.org/10.1146/annurev-psych-062220-051256

Landy, D., Allen, C., & Zednik, C. (2014). A perceptual account of symbolic reasoning [10.3389/fpsyg.2014.00275]. *Frontiers in Psychology*, *5*, 275. https://www.frontiersin.org/article/10.3389/fpsyg.2014.00275

Lawson, A. (2012). *Painted caves: Palaeolithic rock art in Western Europe*. Oxford: Oxford University Press.

Lawson, G., & d'Errico, F. (2002). Microscopic, experimental and theoretical re-assessment of upper Paleolithic bird-bone pipes from Isturitz, France: Ergonomics of design, systems of notation and the origins of musical traditions. In E. Hickmann, A. D. Kilmer, & R. Eichmann (Eds.), *The archaeology of sound: Origin and organization* (pp. 119–142). Rahden/Westfalen: Marie Leidorf.

LeDoux, J. E. (2012). Evolution of human emotion: A view through fear. *Progress in Brain Research*, *195*, 431–442. https://doi.org/10.1016/B978-0-444-53860-4.00021-0

Leeson, P. T., Coyne, C. J., & Boettke, P. J. (2006). Converting social conflict: Focal points and the evolution of cooperation. *Review of Austrian Economics*, *19*, 137–147. https://doi.org/10.1007/s11138-006-7345-4

Legare, C. H. (2017). Cumulative cultural learning: Development and diversity. *Proceedings of the National Academy of Sciences*, *114*(30), 7877–7883.

Legare, C. H. (2019). The development of cumulative cultural learning. *Annual Review of Developmental Psychology*, *1*, 119–147.

Leitgeb, H. (2004). *Inference on the low level*. Dordrecht: Kluwer.

Lennie, P. (2003). The cost of cortical computation. *Current Biology*, *13*(6), 493–497.

Leslie, A. M. (1994). Pretending and believing: Issues in the theory of ToMM. *Cognition*, *50*, 211–238.

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in "theory of mind." *Trends in Cognitive Science*, *8*(12), 528–533.

Levy, W. B., & Baxter, R. A. (1996). Energy efficient neural codes. *Neural Computation* (3), 531–543.

Lewens, T. (2015). *Cultural evolution: Conceptual challenges*. Oxford: Oxford University Press.

Liberto, H. (2014). The exploitation solution to the Non-Identity Problem. *Philosophical Studies*, *167*(1), 73–88.

Lieberman, M. D. (2003). Reflexive and reflective judgment processes: A social cognitive neuroscience approach. In J. P. Forgas, K. D. Williams, & W. von Hippel (Eds.), *Social judgments: Implicit and explicit processes*. (pp. 44–67). Cambridge: Cambridge University Press.

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavior and Brain Sciences*, *43*(e1), 1–60. https://doi.org/doi:10.1017/S0140525X1900061X

Locke, J. (1689 [1983]). *A letter concerning toleration*. Indianapolis: Hackett.

Locke, J. (1987 [1866]). *The two treatises of civil government* (R. Ashcraft, Ed.). London: Routledge.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*(1), 75–98.

Lombrozo, T., Shtulman, A., & Weisberg, M. (2006). The Intelligent Design controversy: Lessons from psychology and education. *Trends in Cognitive Science, 10*(2), 56–57. https://doi.org/10.1016/j.tics.2005.12.001

Lombrozo, T., Thanukos, A., & Weisberg, M. (2008). The importance of understanding the nature of science for accepting evolution. *Evolution: Education and Outreach*, *1*(3), 290–298. https://doi.org/10.1007/s12052-008-0061-8

Luhrmann, T. M., Padmavati, R., Tharoor, H., & Osei, A. (2015). Hearing voices in different cultures: A social kindling hypothesis. *Topics in Cognitive Science*, *7*(4), 646–663. https://doi.org/10.1111/tops.12158

Luo, Y., & Baillargeon, R. (2007). Do 12.5-month-old infants consider what objects others can see when interpreting their actions? *Cognition*, *105*(3), 489–512. https://doi.org/10.1016/j.cognition.2006.10.007

Lurz, R. (2011). *Mindreading animals: The debate over what animals know about other minds*. Cambridge, MA: MIT Press.

Lurz, R. (2018). Animal mindreading: The problem and how it can be solved. In K. Andrews & J. Beck (Eds.), *The Routledge handbook of philosophy of animal minds* (pp. 229–237). London: Routledge.

Łusiak-Szelachowska, M., Międzybrodzki, R., Drulis-Kawa, Z., Cater, K., Knežević, P., Winogradow, C., Amaro, K., Jończyk-Matysiak, E., Weber-Dąbrowska, B., Rękas, J., & Górski, A. (2022). Bacteriophages and antibiotic interactions in clinical practice: What we have learned so far. *Journal of Biomedical Science*, *29*(1), 23. https://doi.org/10.1186/s12929-022-00806-1

Machery, E. (2007). Massive modularity and brain evolution. *Philosophy of Science*, *74*, 825–838.

Machery, E. (2008). A plea for human nature. *Philosophical Psychology*, *21*(3), 321–329. https://doi.org/10.1080/09515080802170119

Machery, E. (2009). *Doing without concepts*. Oxford: Oxford University Press.

Machery, E. (2017). *Philosophy within its proper bounds*. Oxford: Oxford University Press.

Machery, E. (2018). Plasticity and modularity. In R. Joyce (Ed.), *The Routledge handbook of evolution and philosophy* (pp. 268–281). London: Routledge.

Machery, E., & Barrett, H. C. (2006). Debunking adapting minds. *Philosophy of Science*, *73*, 232–246.

Machery, E., & Mallon, R. (2010). Evolution of morality. In *The moral psychology handbook.* (pp. 3–46). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199582143.003.0002

Machery, E., & Stich, S. (2022). Demographic differences in philosophical intuition: A reply to Joshua Knobe. *Review of Philosophy and Psychology*, *14*(2). DOI:10.1007/s13164-021-00609-7

MacLean, E. L. (2016). Unraveling the evolution of uniquely human cognition. *Proceedings of the National Academy of Sciences*, *113*(23), 6348–6354.

MacLean, P. D. (1985). Brain evolution relating to family, play, and the separation call. *Archives of General Psychiatry*, *42*, 405–417.

Mann, J., & Patterson, E. M. (2013). Tool use by aquatic animals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*. https://doi.org/http://doi.org/10.1098/rstb.2012.0424

Margolis, E., & Laurence, S. (Eds.). (2015). *The conceptual mind*. Cambridge, MA: MIT Press.

Martinez, S. F. (2013). Technological scaffoldings for the evolution of culture and cognition. In L. R. Caporael, J. R. Griesemer, & W. Wimsatt (Eds.), *Developing scaffolds in evolution, culture, and cognition* (pp. 249–264). Cambridge, MA: MIT Press.

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory*. Oxford: Oxford University Press.

Mayer, A., & Träuble, B. E. (2013). Synchrony in the onset of mental state understanding across cultures? A study among children in Samoa. *International Journal of Behavioral Development*, *37*, 21–28.

Maynard Smith, J., & Parker, G. A. (1976). The logic of asymmetric contests. *Animal Behavior*, *24*, 159–175.

McGeer, V. (2007). The regulative dimension of folk psychology. In D. Hutto & M. Ratcliffe (Eds.), *Folk Psychology Re-Assessed* (pp. 137–156). New York: Springer.

McGrew, W. C. (2013). Is primate tool use special? Chimpanzee and New Caledonian crow compared. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1630), 20120422. https://doi.org/doi:10.1098/rstb.2012.0422

Mellars, P. (1989). Major issues in the emergence of modern humans. *Current Anthropology*, *30*(3), 349–385.

Menary, R. (2007). *Cognitive integration*. New York: Palgrave Macmillan.

Mikhail, J. (2011). *Elements of moral cognition: Rawls' linguistic analogy and the cognitive science of moral and legal judgment*. Cambridge: Cambridge University Press.

Millikan, R. (1984). *Language, thought, and other biological categories*. Cambridge, MA: MIT Press.

Millikan, R. (1989). Biosemantics. *Journal of Philosophy*, *86*, 281–297.

Millikan, R. (2002). *Varieties of meaning*. Cambridge, MA: MIT Press.

Mirzabe, A. H., Chegini, G. R., & Khazaei, J. (2018). Sunflower petals: Some physical properties and modeling distribution of their number, dimensions, and mass. *Information Processing in Agriculture*, *5*(2), 185–198. https://doi.org/10.1016/j.inpa.2018.03.002

Mithen, S. (1990). *Thoughtful foragers: A study of prehistoric decision makers*. Cambridge: Cambridge University Press.

Mithen, S. (1999). *The prehistory of the mind: The cognitive origins of art, religion and science*. London: Thames & Hudson.

Mithen, S. (2005). *The singing neanderthals: The origins of music, language, mind and body*. London: Weidenfeld & Nicholson.

Miyagawa, S., Lesure, C., & Nóbrega, V. A. (2018). Cross-modality information transfer: A hypothesis about the relationship among prehistoric cave paintings, symbolic thinking, and the emergence of language [perspective]. *Frontiers in Psychology*, *9*(115). https://doi.org/10.3389/fpsyg.2018.00115

Moll, H., Nichols, R., & Mackey, J. L. (2021). Rethinking human development and the shared intentionality hypothesis. *Review of Philosophy and Psychology*, *12*, 453–464. https://doi.org/10.1007/s13164-020-00489-3

Monsó, S., & Andrews, K. (2022). Animal moral psychologies. In M. Vargas & J. Doris (Eds.), *The Oxford handbook of moral psychology* (pp. 388–420). Oxford: Oxford University Press.

Moore, R. (2018). Gricean communication, joint action, and the evolution of cooperation. *Topoi*, *37*(2), 329–341.

Moravec, H. (1999). *Robot: Mere machine to transcendant mind*. Oxford: Oxford University Press.

Morin, O. (2019). Did social cognition evolve by cultural group selection? *Mind & Language*, *34*(4), 530–539. https://doi.org/10.1111/mila.12252

Morley, I. (2013). *The prehistory of music: Human evolution, archaeology and the origins of musicality*. Oxford: Oxford University Press.

Moser, P. (2005). How do patent laws influence innovation? Evidence from nineteenth-century World's Fairs. *American Economic Review*, *95*(4), 1214–1236. https://doi.org/10.1257/0002828054825501

Mulcahy, N. J., & Call, J. (2006). Apes save tools for future use. *Science*, *312*(5776), 1038–1040. https://doi.org/10.1126/science.1125456

Mullins, D. A., Whitehouse, H., & Atkinson, Q. D. (2013a). The role of writing and recordkeeping in the cultural evolution of human cooperation. *Journal of Economic Behavior & Organization*, *90S*, S141–S151.

Mullins, D. A., Whitehouse, H., & Atkinson, Q. D. (2013b). The role of writing and recordkeeping in the cultural evolution of human cooperation. *Journal of Economic Behavior & Organization*, *90*, S141–S151. https://doi.org/10.1016/j.jebo.2012.12.017

Muthukrishna, M., & Henrich, J. (2016). Innovation in the collective brain. *Philosophical Transactions of the Royal Society B*, *371*(20150192). https://doi.org/10.1098/rstb.2015.0192

Narayanan, A., & Kapoor, S. (2023). GPT-4 and professional benchmarks: The wrong answer to the wrong question. *AI Snake Oil*. https://aisnakeoil.substack.com/p/gpt-4-and-professional-benchmarks

Nelson, R., Dosi, G., Helfat, C. E., Pyka, A., Saviotti, P. P., Lee, K., Dopfer, K., Malerba, F., & Winter, S. (Eds.). (2018). *Modern evolutionary economics: An overview*. Cambridge University Press.

Nichols, S. (2004). *Sentimental rules: On the natural foundations of moral judgment*. Oxford: Oxford University Press.

Nichols, S., & Stich, S. (2003). *Mindreading. An integrated account of pretence, self-awareness, and understanding other minds* (vol. 114). Oxford: Oxford University Press.

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought. *Psychology Review*, *108*(2), 291–310.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychology Review, 84*, 231–259.

Nissen, H. J. (1985). The emergence of writing in the ancient Near East. *Interdisciplinary Science Reviews*, *10*(4), 349–361. https://doi.org/10.1179/isr.1985.10.4.349

Niven, J. E., & Laughlin, S. B. (2008). Energy limitation as a selective pressure on the evolution of sensory systems. *Journal of Experimental Biology*, *211*(11), 1792–1804.

Noe, R., & Voelkl, B. (2013). Cooperation and biological markets: The power of partner choice. In K. Sterelny, R. Joyce, B. Calcott, & B. Fraser (Eds.), *Cooperation and its evolution* (pp. 131–152). Cambridge, MA: MIT Press.

Nordhaus, W. (1969). *Invention, growth and welfare*. Cambridge, MA: MIT Press.

North, D. C. (1990). *Institutions, institutional change, and economic performance*. Cambridge: Cambridge University Press.

Northcott, R., & Piccinini, G. (2018). Conceived this way: Innateness defended. *Philosophers' Imprint*, *18*, 1–16.

Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect Reciprocity *Nature*, *437*. https://doi.org/doi:10.1038/nature04131

Nowell, A. (2010). Defining behavioral modernity in the context of neandertal and anatomically modern human populations. *Annual Review of Anthropology*, *39*(1), 437–452. https://doi.org/10.1146/annurev.anthro.012809.105113

Nussbaum, M. C. (1999). Virtue ethics: A misleading category? *Journal of Ethics*, *3*(3), 163–201.

O'Neill, O. (2004). Kant: Rationality as practical reason. In A. Mele & P. Rawling (Eds.), *The Oxford handbook of rationality* (pp. 93–109). Oxford: Oxford University Press.

Oberliessen, L., & Kalenscher, T. (2019). Social and non-social mechanisms of inequity aversion in non-human animals [10.3389/fnbeh.2019.00133].

Frontiers in Behavioral Neuroscience, *13*, 133. https://www.frontiersin.org/article/10.3389/fnbeh.2019.00133

Odling-Smee, F. J., Laland, K. N., & Feldman, M. W. (2003). *Niche construction: The neglected process in evolution*. Princeton, NJ: Princeton University Press.

Ofek, H. (2001). *Second nature: Economic origins of human evolution*. Cambridge: Cambridge University Press.

Ofek, H. (2013). MHC-mediated benefits of trade: A biomolecular approach to cooperation in the marketplace. In K. Sterelny, R. Joyce, B. Calcott, & B. Fraser (Eds.), *Cooperation and its evolution* (pp. 175–194). Cambridge, MA: MIT Press.

Olin, L. (2016). Questions for a theory of humor. *Philosophy Compass*, *11*, 338–350.

Orzack, S. H., & Sober, E. (1994). Optimality models and the test of adaptationism. *American Naturalist*, *143*(3), 361–380. https://doi.org/10.1086/285608

Osiurak, F., & Reynaud, E. (forthcoming). The elephant in the room: What matters cognitively in cumulative technological culture. *Behavioral and Brain Sciences*, 1–57. https://doi.org/10.1017/S0140525X19003236

Osmanoglu, K., & Schulz, A. (2019). It just looks the same: An evolutionary psychological account of differences in racial cognition among infants and older humans. *Review of Philosophy and Psychology*, *10*(3), 631–647.

Otsuka, M. (2003). *Libertarianism without Inequality*. Oxford: Oxford University Press. https://doi.org/10.1093/0199243956.001.0001

Oudeyer, P.-Y., & Kaplan, F. (2007). What is intrinsic motivation? A typology of computational approaches. *Frontiers in Neurorobotics*, *1*, 6. https://doi.org/10.3389/neuro.12.006.2007

Oudeyer, P.-Y., & Smith, L. B. (2016). How evolution may work through curiosity-driven developmental process. *Topics in Cognitive Science*, *8*(2), 492–502. https://doi.org/10.1111/tops.12196

Pacey, A. (1992). *The maze of ingenuity: Ideas and idealism in the development of technology* (2nd ed.). Cambridge, MA: MIT Press.

Papadopoulos, D. (2023). Shared intentionality in nonhuman great apes: A normative model. *Review of Philosophy and Psychology*, *14*(4), 1125–1145. https://doi.org/10.1007/s13164-021-00594-x

Papineau, D. (1987). *Reality and representation*. Hoboken, NJ: Blackwell.

Papineau, D. (2003). *The roots of reason*. Oxford: Oxford University Press.

Parker, S. T., Mitchell, R. W., & Boccia, M. L. (1994). *Self-awareness in Animals and Humans*. Cambridge: Cambridge University Press.

Parrish, A. E., Perdue, B. M., Evans, T. A., & Beran, M. J. (2013). Chimpanzees (Pan troglodytes) transfer tokens repeatedly with a partner to accumulate rewards in a self-control task. *Animal Cognition*, *16*(4), 627–636. https://doi.org/10.1007/s10071-013-0599-x

Parsons, T. (1951). *The social system*. London: Routledge.

Pathak-Vaidya, P., Sharma, S., & Telang, M. (2021). Bacteriophage as an antibacterial agent: A patent perspective. *Future of Microbiology*, *16*, 1327–1339. https://doi.org/10.2217/fmb-2021-0062

Pelé, M., Dufour, V., Thierry, B., & Call, J. (2009). Token transfers among great apes (Gorilla gorilla, Pongo pygmaeus, Pan paniscus, and Pan troglodytes): Species differences, gestural requests, and reciprocal exchange. *Journal of Comparative Psychology*, *123*(4), 375–384. https://doi.org/10.1037/a0017253

Penin, J. (2005). Patents versus ex post rewards: A new look. *Research Policy*, *34*(5), 641–656.

Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences, 31*(2), 109–130; discussion 130-178. https://doi.org/10.1017/s0140525x08003543

Pepperberg, I. M. (2008). Difficulties with "humaniqueness". *Behavioral and Brain Sciences, 31*(2), 143–144. https://doi.org/10.1017/S0140525X08003695

Perlovsky, L. I., Bonniot-Cabanac, M., & Cabanac, M. (2010). Curiosity and pleasure. *WebmedCentral Psychology*, *1*, WMC001275.

Peterson, C., & Wellman, H. (2009). From fancy to reason: Scaling deaf and hearing children's understanding of theory of mind and pretence. *British Journal of Developmental Psychology*, *27*, 297–310.

Pettit, P. (1996). Functional explanation and virtual selection. *British Journal for the Philosophy of Science*, *47*(2), 291–302.

Phillips, J., Buckwalter, W., Cushman, F., Friedman, O., Martin, A., Turri, J., Santos, L., & Knobe, J. (2021). Knowledge before belief. *Behavioral and Brain Sciences*, *44*, e140, Article e140. https://doi.org/10.1017/S0140525X20000618

Piantadosi, S. T. (2024). Modern language models refute Chomsky's approach to language. In E. Gibson & M. Poliak (Eds.), *From fieldwork to linguistic theory: A tribute to Dan Everett*. Berlin: Language Science Press. DOI: 10.5281/zenodo.11351540

Piccinini, G. (2010). The mind as neural software? Understanding functionalism, computationalism, and computational functionalism. *Philosophy and Phenomenological Research*, *81*(2), 269–311.

Piccinini, G., & Scarantino, A. (2010). Computation vs. information processing: Why their difference matters to cognitive science. *Studies in the History and Philosophy of Science*, *41*(3), 237–246.

Piccinini, G., & Schulz, A. (2019). The ways of altruism. *Evolutionary Psychological Science*, *5*, 58–70.

Pike, A. W. G., Hoffmann, D. L., García-Diez, M., Pettitt, P. B., Alcolea, J., De Balbín, R., González-Sainz, C., de las Heras, C., Lasheras, J. A., Montes, R.,

& Zilhão, J. (2012). U-series dating of paleolithic art in 11 caves in Spain. *Science*, *336*(6087), 1409–1413. https://doi.org/10.1126/science.1219957

Pinker, S. (1997). *How the mind works*. New York: Norton.

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, *13*(4), 707–784.

Plomin, R. (1990). *Nature and nurture: An introduction to human behavioral genetics*. Boston: Brooks/Cole.

Povinelli, D. (2003). *Folk physics for apes: The chimpanzee's theory of how the world works*. Oxford: Oxford University Press.

Premo, L. S., & Kuhn, S. L. (2010). Modeling effects of local extinctions on culture change and diversity in the Paleolithic. *PloS One*, *5*(12), e15582. https://doi.org/10.1371/journal.pone.0015582

Preston, B. (1998). Why is a wing like a spoon? A pluralist theory of function. *Journal of Philosophy*, *95*(5), 215–254.

Preston, B. (2013). *A philosophy of material culture: Action, function, and mind*. London: Routledge.

Prétôt, L., & Brosnan, S. F. (2015). The evolution of morality: A comparative approach. In J. Decety & T. Wheatley (Eds.), *The moral brain: A multidisciplinary perspective* (pp. 3–18). Cambridge, MA: MIT Press.

Prinz, J. (2002). *Furnishing the mind: Concepts and their perceptual basis*. Cambridge, MA: MIT Press.

Prinz, J. (2008). Resisting the linguistic analogy: A commentary on Hauser, Young, and Cushman. In W. Sinnott-Armstrong (Ed.), *Moral psychology* (vol. 2: The cognitive science of morality: Intuition and diversity.). Cambridge, MA: MIT Press.

Prinz, J. (2011). Is empathy necessary for morality? In A. Coplan & P. Goldie (Eds.), *Empathy: Philosophical and psychological perspectives* (pp. 211–229). Oxford: Oxford University Press.

Pryor, F. L. (1981). A survey of the economic systems of wild chimpanzees and baboons. *Journal of Economic Issues*, *15*(1), 33–59.

Pyysiainen, I., & Hauser, M. (2010). The origins of religion: Evolved adaptation or by-product? *Trends in Cognitive Science, 14*(3), 104–109. https://doi.org/10.1016/j.tics.2009.12.007

Qian, Y. (2007). Do national patent laws stimulate domestic innovation in a global patenting environment? A cross-country analysis of pharmaceutical patent protection, 1978–2002. *Review of Economics & Statistics*, *89*(3), 436–453.

Ramsey, W. (2007). *Representation reconsidered*. Cambridge: Cambridge University Press.

Ramsey, W. (2014). Systematicity and architectural pluralism. In P. Calvo & J. Symons (Eds.), *The architecture of cognition* (pp. 253–276). Cambridge, MA: MIT Press.

Rawls, J. (1999). *A theory of justice* (Revised ed.). Cambridge, MA: MIT Press.

Read, D. W. (2008). Working memory: A cognitive limit to non-human primate recursive thinking prior to hominid evolution. *Evolutionary Psychology*, *6*(4), 676–714.

Reader, S. M., & Laland, K. N. (2003). Animal innovation: An introduction. In S. M. Reader & K. N. Laland (Eds.), *Animal innovation* (pp. 3–37). Oxford: Oxford University Press.

Reed, E. S. (1996). *Encountering the world: Toward an ecological psychology*. Oxford: Oxford University Press.

Reindl, E., Bandini, E., & Tennie, C. (2018). The zone of latent solutions and its relation to the classics: Vygotsky and Köhler. In L. Di Paolo, F. Di Vincenzo, & F. De Petrillo (Eds.), *Evolution of primate social cognition: Interdisciplinary evolution research* (pp. 231–248). New York: Springer.

Renfrew, C., & Scarre, C. (Eds.). (1998). *Cognition and material culture: The archaeology of symbolic storage*. Cambridge: McDonald Institute for Archaeological Research.

Renner, E., & Zawidzki, T. W. (2018). Minimal cognitive preconditions on the ratchet. In L. D. Di Paolo, F. Di Vincenzo, & F. De Petrillo (Eds.), *Evolution of primate social cognition* (pp. 249–265). New York: Springer International. https://doi.org/10.1007/978-3-319-93776-2_16

Režek, Z. e., Dibble, H. L., McPherron, S. P., Braun, D. R., & Lin, S. C. (2018). Two million years of flaking stone and the evolutionary efficiency of stone tool technology. *Nature Ecology & Evolution 2*, 628–633. https://doi.org/10.1038/s41559-018-0488-4.

Rietveld, E. (2008). Situated normativity: The normative aspect of embodied cognition in unreflective action. *Mind*, *117*(468), 973–1001. https://doi.org/10.1093/mind/fzn050

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1996). The structure of normal human attention: The Test of Everyday Attention. *Journal of the International Neuropsychological Society*, *2*(6), 525–534. https://doi.org/10.1017/S1355617700001697

Rosati, A. G., Stevens, J. R., & Hauser, M. D. (2006). The effect of handling time on temporal discounting in two New World primates. *Animal Behavior 71*, 1379–1387.

Rosenberg, A. (2012). *Philosophy of social science* (4th ed.). Boulder, CO: Westview Press.

Ross, D. (2004). Meta-linguistic signaling for coordination amongst social agents. *Language Sciences*, *26*, 621–642.

Ross, D. (2007). H. sapiens as ecologically special: What does language contribute? *Language Sciences*, *29*, 710–731. https://doi.org/10.1016/j.langsci.2006.12.008

Rossi, M. (2014). Simulation theory and interpersonal utility comparisons reconsidered. *Synthese*, *191*(6), 1185–1210. http://www.jstor.org/stable/24021575

Rowlands, M. (2010). *The new science of the mind: From extended mind to embodied phenomenology*. Cambridge, MA: MIT Press.

Russell, S., & Norvig, P. (2009). *Artificial intelligence: A modern approach* (3rd ed.). Saddle River, NJ: Prentice Hall.

Russo, F. (2023). *Techno-scientific practices: an informational approach*. Washington, DC: Rowman & Littlefield.

Russon, A. E., & Andrews, K. (2010). Orangutan pantomime: Elaborating the message. *Biology Letters*, *7*(4). https://doi.org/10.1098/rsbl.2010.0564

Samuels, R. (2006). Is the mind massively modular? In R. Stainton (Ed.), *Contemporary debates in cognitive science* (pp. 37–56). Hoboken, NJ: Blackwell.

Sanz, C. M., & Morgan, D. B. (2013). Ecological and social correlates of chimpanzee tool use. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1630), 20120416. https://doi.org/doi:10.1098/rstb.2012.0416

Sarkissian, H., Park, J., Tien, D., Wright, J. C., & Knobe, J. (2011). Folk moral relativism. *Mind & Language*, *26*(4), 482–505. https://doi.org/10.1111/j.1468-0017.2011.01428.x

Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences*, *107*(6), 2408–2412. https://doi.org/10.1073/pnas.0908239106

Saxe, R., & Powell, L. J. (2006). It's the thought that counts: Specific brain regions for one component of theory of mind. *Psychological Science*, *17*(8), 692–699.

Saxe, R., Schulz, L. E., & Jiang, Y. V. (2006). Reading minds versus following rules: Dissociating theory of mind and executive control in the brain. *Society for Neuroscience*, *1*(3-4), 284–298. https://doi.org/10.1080/17470910601000446

Schaller, M., Park, J. H., & Kenrick, D. T. (2007). Human evolution and social cognition. In L. Barrett & R. Dunbar (Eds.), *Oxford handbook of evolutionary psychology* (pp. 491–504). Oxford: Oxford University Press.

Schliesser, E. (2011). Reading Adam Smith after Darwin: On the evolution of propensities, institutions, and sentiments. *Journal of Economic Behavior & Organization*, *77*(1), 14–22. https://doi.org/10.1016/j.jebo.2010.04.007

Schliesser, E., & Smith, G. E. (1996). Huygens' 1688 report to the Directors of the Dutch East India Company on the measurement of longitude at sea and its implications for the non-uniformity of gravity. *De zeventiende eeuw*, *12*, 198–214.

Schlingloff, L., & Moore, R. (2017). Do chimpanzees conform to cultural norms? In K. Andrews & J. Beck (Eds.), *The Routledge handbook of philosophy of animal minds* (pp. 381–389). London: Routledge.

Schouten, G. (2019). *Liberalism, neutrality, and the gendered division of labor*. Oxford: Oxford University Press.

Schroeder, M. (2010). *Noncognitivism in ethics*. London: Routledge.
Schroeder, W. (2000). Continental ethics. In H. LaFollette (Ed.), *The Blackwell guide to ethical theory* (pp. 375–399). Hoboken, NJ: Blackwell.
Schulz, A. (2011a). Simulation, simplicity, and selection: An evolutionary perspective on high-level mindreading. *Philosophical Studies*, *152*(2), 271–285. https://doi.org/10.1007/s11098-009-9476-5
Schulz, A. (2011b). Sober & Wilson's evolutionary arguments for psychological altruism: A reassessment. *Biology and Philosophy*, *26*, 251–260.
Schulz, A. (2013). Selection, drift, and independent contrasts: Defending the methodological foundations of the FIC. *Biological Theory*, *7*(1), 38–47. https://doi.org/10.1007/s13752-012-0070-2
Schulz, A. (2017). The evolution of empathy. In H. Maibom (Ed.), *The Routledge handbook of philosophy of empathy* (pp. 64–73). London: Routledge.
Schulz, A. (2018a). By genes alone: A model selectionist argument for genetical explanations of cooperation in non-human organisms [journal article]. *Biology & Philosophy*, *32*, 951–967. https://doi.org/10.1007/s10539-017-9584-0
Schulz, A. (2018b). *Efficient cognition: The evolution of representational decision making*. Cambridge, MA: MIT Press.
Schulz, A. (2021). Explaining human diversity: The need to balance fit and complexity. *Review of Philosophy and Psychology*. https://doi.org/10.1007/s13164-021-00582-1
Schulz, A. (2022). What's the point? A presentist social functionalist account of institutional purpose. *Philosophy of the Social Sciences*, *52*(1–2), 53–80. https://doi.org/10.1177/00483931211049765
Schulz, A. (2024). Core knowledge and its role in explaining uniquely human cognition: Some questions. *Behavior and Brain Sciences*, *47*: e141. doi:10.1017/S0140525X23003199
Schulz, A., & Robins, S. (2022). Episodic memory, simulated future planning, and their evolution. *Review of Philosophy and Psychology*, 1–22. https://doi.org/10.1093/bjps/axr038
Schumpeter, J. (1942). *Capitalism, socialism, and democracy*. New York: Harper & Brothers.
Schumpeter, J. (1959). *The theory of economic development*. Cambridge, MA: Harvard University Press.
Scorolli, C., Borghi, A. M., & Tummolini, L. (2018). Cues of control modulate the ascription of object ownership. *Psychological Research*, *82*, 929–954. https://doi.org/10.1007/s00426-017-0871-9
Scruton, R. (1983). Understanding music. In R. Scruton (Ed.), *The aesthetic understanding: Essays in the philosophy of art and culture* (pp. 77–100). Manchester, UK: Carcanet Press.
Seabright, P. (2010). *The company of strangers: A natural history of economic life* (revised ed.). Princeton, NJ: Princenton University Press.

Searle, J. (1980). Minds, brains and programs. *Behavioral and Brain Sciences*, *3*, 417–424.

Shahaeian, A., Peterson, C., Slaughter, V., & Wellman, H. (2011). Culture and the sequence of steps in theory of mind development. *Developmental Psychology*, *47*, 1239–1247.

Shea, J. J. (2017). *Stone tools in human evolution: Behavioral differences among technological primates*. Cambridge: Cambridge University Press.

Sherratt, T. N., & Mesterton-Gibbons, M. (2015). The evolution of respect for property. *Journal of Evolutionary Biology*, *28*, 1185–1202.

Shiffrin, S. V. (2001). Lockean arguments for private intellectual property. In S. R. Munzer (Ed.), *New essays in the legal and political theory of property* (pp. 138–167). Cambridge: Cambridge University Press.

Shrader-Frechette, K. (2006). Gene patents and Lockean constraints. *Public Affairs Quarterly*, *20*(2), 135–161. http://www.jstor.org/stable/40441433

Shultz, S., Nelson, E., & Dunbar, R. (2012). Hominin cognitive evolution: Identifying patterns and processes in the fossil and archaeological record. *Philosophical Transactions of the Royal Society of London B Biologic Science*, *367*(1599), 2130–2140. https://doi.org/10.1098/rstb.2012.0115

Shumaker, B. R. W., Walkup, K. R., & Beck, B. B. (2011). *Animal tool behavior: The use and manufacture of tools by animals* (revised and updated ed.). Baltimore: Johns Hopkins University Press.

Sichelman, T. (2010). Commercializing patents. *Stanford Law Review*, *62*(2), 341–402.

Sijilmassi, A., Safra, L., & Baumard, N. (forthcoming). “Our roots run deep”: Historical myths as culturally evolved technologies for coalitional recruitment. *Behavior and Brain Sciences*. Published online 2024:1–44. doi:10.1017/S0140525X24000013

Silberstein, M., & Chemero, A. (2012). Complexity and extended phenomenological-cognitive systems. *Topics in Cognitive Science*, *4*(1), 35–50. Https://doi.org/10.1111/j.1756-8765.2011.01168.x

Singer, P. (2011). *The expanding circle: Ethics, evolution, and moral progress*. Princeton, NJ: Princeton University Press.

Skyrms, B. (1996). *Evolution and the social contract*. Cambridge: Cambridge University Press.

Skyrms, B. (2004). *The stag hunt and the evolution of social structure*. Cambridge: Cambridge University Press.

Skyrms, B. (2010). *Signals: Evolution, learning, and information*. Oxford: Oxford University Press.

Smead, R., & Forber, P. (forthcoming). Convention and the origins of ownership. *Philosophy of Science*. https://doi.org/10.1086/710626

Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations*. London: W. Strahan and T. Cadell.

Smith, G. E. (2005). Was wrong Newton bad Newton? In J. Z. Buchwald & A. Franklin (Eds.), *Wrong for the right reasons* (vol. 127–160). New York: Springer.

Smith, L. B., & Thelen, E. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Sober, E. (1980). Evolution, population thinking, and essentialism. *Philosophy of Science*, *47*(3), 350–383.

Sober, E. (2008). *Evidence and evolution*. Cambridge: Cambridge University Press.

Sober, E., & Wilson, D. S. (1998). *Unto others: The evolution and psychology of unselfish behavior*. Cambridge, MA: Harvard University Press.

Spaulding, S. (2011). Embodied social cognition. *Philosophical Topics*, *39*(1), 141–162.

Spaulding, S. (2018). *How we understand others: Philosophy and social cognition*. London: Routledge.

Spelke, E. (2022). *What babies know*. Oxford: Oxford University Press.

Sperber, D. (1975). *Rethinking symbolism*. Cambridge: Cambridge University Press.

Sperber, D. (1996). *Explaining culture: A naturalistic approach*. New York: Wiley.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Hoboken, NJ: Blackwell.

Sripada, C. S., & Stich, S. (2006). A framework for the psychology of norms. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *Innateness and the structure of the mind* (vol. II, pp. 280–302). Oxford: Oxford University Press.

Stanford, P. K. (2018). The difference between ice cream and Nazis: Moral externalization and the evolution of human cooperation. *Behavioral and Brain Sciences, 41*(e95).

Sterelny, K. (1999). Situated agency and the descent of desire. In V. Hardcastle (Ed.), *Where biology meets psychology: Constraints, connections, conjectures* (pp. 203–220). Cambridge, MA: MIT Press.

Sterelny, K. (2003). *Thought in a hostile world: The evolution of human cognition*. Oxford: Wiley-Blackwell.

Sterelny, K. (2012a). *The evolved apprentice: How evolution made humans unique*. Cambridge, MA: MIT Press.

Sterelny, K. (2012b). From fitness to utility. In K. Binmore & S. Okasha (Eds.), *Evolution and rationality: Decision, cooperation, and strategic behaviour* (pp. 246–273). Cambridge: Cambridge University Press.

Sterelny, K. (2012c). Language, gesture, skill: The co-evolutionary foundations of language [10.1098/rstb.2012.0116]. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1599), 2141. http://rstb.royalsocietypublishing.org/content/367/1599/2141.abstract

Sterelny, K. (2016). Cumulative cultural evolution and the origins of language. *Biological Theory*, *11*(3), 173–186. https://doi.org/10.1007/s13752-016-0247-1

Sterelny, K. (2017). Artifacts, symbols, thoughts. *Biological Theory*, *12*(4), 236–247.

Sterelny, K. (2018). Culture and the extended phenotype: Cognition and material culture in deep time. In N. Albert, B. Leon De, & G. Shaun (Eds.), *The Oxford handbook of 4E cognition* (pp. 773–792). Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198735410.013.41

Sterelny, K. (2021). *The Pleistocene social contract: Culture and cooperation in human evolution*. Oxford: Oxford University Press.

Sterelny, K., Joyce, R., Calcott, B., & Fraser, B. (Eds.). (2013). *Cooperation and its evolution*. Cambridge, MA: MIT Press.

Stevens, J. R., & Stephens, D. W. (2008). Patience. *Current Biology*, *18*, R11–12.

Stich, S. (1990). *The fragmentation of reason*. Cambridge, MA: MIT Press.

Stich, S. (2007). Evolution, altruism and cognitive architecture: A critique of Sober and Wilson's argument for psychological altruism. *Biology and Philosophy*, *22*, 267–281.

Stich, S. (2016). Why there might not be an evolutionary explanation for psychological altruism. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *56*, 3–6. https://doi.org/10.1016/j.shpsc.2015.10.005

Stich, S. (2018). Do we really externalize or objectivize moral demands? *Behavioral and Brain Sciences*, *41*, e113, Article e113. https://doi.org/10.1017/S0140525X18000183

Stich, S., Doris, J., & Roedder, E. (2010). Altruism. In J. D. a. t. M. P. R. Group (Ed.), *The moral psychology handbook* (pp. 147–205). Oxford: Oxford University Press.

Stout, D., & Chaminade, T. (2012). Stone tools, language and the brain in human evolution. *Philosophical Transactions of the Royal Society of London B Biological Science*, *367*(1585), 75–87.

Strawson, P. F. (1962). Freedom and resentment. *Proceedings of the British Academy*, *48*, 1–25.

Street, S. (2006). A Darwinian dilemma for realist theories of value. *Philosophical Studies*, *127*, 109.

Strevens, M. (2003). The role of the priority rule in science. *Journal of Philosophy*, *100*, 55–79.

Suddendorf, T., & Butler, D. L. (2013). The nature of visual self-recognition. *Trends in Cognitive Science*, *17*, 121–127.

Sugden, K. F. (1981). A history of the abacus. *Accounting Historians Journal*, *8*(2), 1–22. https://doi.org/10.2308/0148-4184.8.2.1

Sugiyama, L. S., & Sugiyama, M. S. (2003). Social roles, prestige, and health risk. *Human Nature*, *14*(2), 165–190.

Suhler, C. L., & Churchland, P. (2011). Can innate, modular "foundations" explain morality? Challenges for Haidt's moral foundations theory. *Journal of Cognitive Neuroscience*, *23*(9), 2103–2116. https://doi.org/10.1162/jocn.2011.21637

Teece, D. J. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy*, *15*, 285–290.

Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: On the evolution of cumulative culture. *Philosophical Transactions of the Royal Society of London B Biological Science*, *364*(1528), 2405–2415. https://doi.org/10.1098/rstb.2009.0052

Tennie, C., & Over, H. (2012). Cultural intelligence is key to explaining human tool use. *Behavioral and Brain Sciences*, *35*(4), 242–243. https://doi.org/10.1017/S0140525X11001968

Tetzlaff, M., & Carruthers, P. (2008). Languages of thought need to be distinguished from learning mechanisms, and nothing yet rules out multiple distinctively human learning systems. *Behavioral and Brain Sciences, 31*(2), 148–149. https://doi.org/doi: 10.1017/S0140525X08003749

Tibble, L., & Carvalho, S. (2018). Rethinking the evolution of property and possession: A review and methodological proposition. *Evolutionary Anthropology, 27*(6), 285–296. https://doi.org/10.1002/evan.21748

Toda, K., & Platt, M. L. (2015). Animal cognition: Monkeys pass the mirror test. *Current Biology*, *25*(2), R64–R66. https://doi.org/10.1016/j.cub.2014.12.005

Todd, P., & Benbasat, I. (1992). The use of information in decision making: An experimental investigation of the impact of computer-based decision aids. *MIS Quarterly*, *16*(3), 373–393. https://doi.org/10.2307/249534

Toh, Y. L. (2021). When paying bills, low-income consumers incur higher costs. Federal Reserve Bank of Kansas City; Payments System Research Briefing; Issue: November 23, 2021.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Tomasello, M. (2021). *Becoming human: A theory of ontogeny*. Cambridge, MA: Harvard University Press.

Tomasello, M. (2022). *The evolution of agency*. Cambridge, MA: MIT Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral & Brain Sciences*, *28*, 675–735.

Tomasello, M., & Herrmann, E. (2010). Ape and human cognition: What's the difference? *Current Directions in Psychological Science*, *19*(1), 3–8.

Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind* (pp. 19–136). Oxford: Oxford University Press.

Tooby, J., & Cosmides, L. (2010). Groups in mind: The coalitional roots of war and morality. In H. Høgh-Olesen (Ed.), *Human morality and sociality: Evolutionary and comparative perspectives* (pp. 191–234). London: Palgrave-Macmillan.

Tooby, J., Cosmides, L., & Barrett, H. C. (2005). Resolving the debate on innate ideas: Learnability constraints and the evolved interpenetration of motivational and conceptual functions. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind: Structure and contents* (pp. 305–337). Oxford: Oxford University Press.

Torii, M. (1974). Possession by non-human primates. *Contemporary Primatology*, 310–314.

Torrance, A. W., & Tomlinson, W. M. (2009). Patents and the regress of useful arts. *Columbia Science and Technology Law Review*, *10*, 130–168.

Triki, Z., Wismer, S., Rey, O., Ann Binning, S., Levorato, E., & Bshary, R. (2019). Biological market effects predict cleaner fish strategic sophistication. *Behavioral Ecology*, *30*(6), 1548–1557. https://doi.org/10.1093/beheco/arz111

Trinkhaus, E., & Shipman, P. (1993). *The Neandertals: Changing the image of mankind*. New York: Knopf.

Tse, P. U. (2008). Symbolic thought and the evolution of human morality. In W. Sinnott-Armstrong (Ed.), *Moral psychology* (vol. 1: The evolution of morality: Adaptations and innateness, pp. 269–297). Cambridge, MA: MIT Press.

Urton, G. (2018). Quipus and Yupanas as imperial registers: Reckoning and recording in Tahuantinsuyu In S. Alconini & A. Covey (Eds.), *The Oxford handbook of the Incas* (pp. 601–622). Oxford: Oxford University Press.

Vaesen, K. (2012). The cognitive bases of human tool use. *Behavioral and Brain Sciences*, *35*(4), 203–218.

van Basshuysen, P. (2023). Austinian model evaluation. *Philosophy of Science*, *90*(5), 1459–1468.

Van Gelder, T. (1995). What might cognition be, if not computation? *Journal of Philosophy*, *91*, 345–381.

van Schaik, C. P., & Burkart, J. M. (2011). Social learning and evolution: The cultural intelligence hypothesis. *Philosophical Transactions of the Royal Society of London B Biological Science*, *366*, 1008–1016.

van Schaik, C. P., & Pradhan, G. R. (2003). A model for tool-use traditions in primates: Implications for the coevolution of culture and cognition. *Journal of Human Evolution*, *44*(6), 645–664.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.

Ventola, C. L. (2015). The antibiotic resistance crisis: Part 2: Management strategies and new agents. *P T*, *40*(5), 344–352.

Wadley, L., Hodgskiss, T., & Grant, M. (2009). Implications for complex cognition from the hafting of tools with compound adhesives in the Middle Stone Age, South Africa. *Proceedings of the National Academy of Sciences*, *106*(24), 9590–9594. https://doi.org/10.1073/pnas.0900957106

Waldron, J. (1990). *The right to private property* (vol. 40). Oxford: Clarendon Press.

Wallach, W., & Allen, C. (2010). *Moral machines: Teaching robots right from wrong*. Oxford: Oxford University Press.

Wang, Z., Wei, X.-X., Stocker, A. A., & Lee, D. D. (2016). Efficient neural codes under metabolic constraints. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, & R. Garnett (Eds.), *Advances in neural information processing systems* (vol. 29, pp. 4619–4627). Red Hook, NY: Curran Associates.

Ward, Z. B. (2022). Cognitive variation: The philosophical landscape. *Philosophy Compass*, *17*(10), e12882. https://doi.org/10.1111/phc3.12882

Warneken, F., Lohse, K., Melis, A. P., & Tomasello, M. (2011). Young children share the spoils after collaboration. *Psychological Science*, *22*(2), 267–273. https://doi.org/10.1177/0956797610395392

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *311*(5765), 1301–1303. https://doi.org/10.1126/science.1121448

Wasserman, E. A. (2008). On possible discontinuities between human and nonhuman minds. *Behavioral and Brain Sciences, 31*(2), 151–152. https://doi.org/10.1017/S0140525X08003774

Weber, B. H., & Depew, D. J. (Eds.). (2003). *Evolution and learning: The Baldwin Effect reconsidered*. Cambridge, MA: MIT Press.

Weber, M. (1958). *The Protestant ethic and the spirit of capitalism*. New York: Scribner.

West, S. A., El Mouden, C., & Gardner, A. (2011). Sixteen common misconceptions about the evolution of cooperation in humans. *Evolution and Human Behavior*, *32*(4), 231–262. https://doi.org/10.1016/j.evolhumbehav.2010.08.001

Westra, E., & Andrews, K. (2022). A pluralistic framework for the psychology of norms. *Biology and Philosophy*, *37*(5), 1–30.

Westra, E., & Nagel, J. (2021). Mindreading in conversation. *Cognition*, *210*, 104618.

Whiten, A. (1995). When does smart behavior-reading become mind-reading? In P. Carruthers & P. Smith (Eds.), *Theories of theories of mind* (pp. 277–292). Cambridge: Cambridge University Press.

Whiten, A., & Byrne, R. W. (Eds.). (1997). *Machiavellian intelligence II: Extensions and evaluations*. Cambridge: Cambridge University Press.

Wilson, B. J. (2020). *The property species: Mine, yours, and the human mind*. Oxford: Oxford University Press.

Wilson, D. S. (2002). *Darwin's cathedral: Evolution, religion, and the nature of society*. Chicago: University of Chicago Press.

Wilson, D. S., Gowdy, J. M., & Rosser Jr., J. B. (2013). Rethinking economics from an evolutionary perspective. *Journal of Economic Behavior & Organization*, *90S*, S1–S2.

Wimsatt, W. (2007). *Re-engineering philosophy for limited beings*. Cambridge, MA: Harvard University Press.

Winner, L. (1980). Do artifacts have politics? *Daedalus*, *109*(1), 121–136.

World Bank. Trade (% of GDP). https://data.worldbank.org/indicator/NE.TRD.GNFS.ZS

Wylie, A. (2002). *Thinking from things: Essays in the philosophy of archaeology*. Berkeley CA: University of California Press.

Wynn, T. (1993). Two developments in the mind of early Homo. *Journal of Anthropological Archaeology*, *12*, 299–322.

Wynn, T., & Coolidge, F. L. (2011). The implications of the working memory model for the evolution of modern cognition. *International Journal of Evolutionary Biology*, *2011*, 741357–741357. https://onlinelibrary.wiley.com/doi/10.4061/2011/741357

Yates, R. R. (1999). New intervention strategies for reducing antibiotic resistance. *Chest*, *115*(3, Suppl. 1), 24S–27S. https://doi.org/10.1378/chest.115.suppl_1.24S

Zawidzki, T. W. (2011). How to interpret infant socio-cognitive competence. *Review of Philosophy and Psychology*, *2*(3), 483–497.

Zawidzki, T. W. (2013). *Mindshaping: A new framework for understanding human social cognition*. Cambridge, MA: MIT Press.

Zawidzki, T. W. (2018). Mindshaping. In S. Gallagher, A. Newen, & L. D. Bruin (Eds.), *The Oxford handbook of 4E cognition* (pp. 735–754). Oxford: Oxford University Press.

Zeberg, H., & Pääbo, S. (2021). A genomic region associated with protection against severe COVID-19 is inherited from Neandertals. *Proceedings of the National Academy of Sciences*, *118*(9), e2026309118. https://doi.org/doi:10.1073/pnas.2026309118

Zollman, K. (2018). The credit economy and the economic rationality of science. *Journal of Philosophy*, *115*(1), 5–33.

Zou, Y., Turvey, S. T., Cui, J., Zhang, H., & Gong, W. (2022). Recent recovery of the world's rarest primate is not directly linked to increasing habitat quality [original research]. *Frontiers in Ecology and Evolution*, *10*. https://doi.org/10.3389/fevo.2022.953637

Index

For the benefit of digital users, indexed terms that span two pages (e.g., 52–53) may, on occasion, appear on only one of those pages.