Sharing our normative worlds: A theory of normative thinking

A thesis submitted for the degree of Doctor of Philosophy at the Australian National University

by

Ivan Gonzalez-Cabrera

in

August 2017

Statement of Originality

This thesis is solely the work of its author. No part of it has previously been submitted for any degree or is currently being submitted for any other degree. To the best of my knowledge, any help received in preparing this thesis, and all sources used, have been duly acknowledged.

Parts of this thesis draw on earlier published work. Chapter 3 is taken, with some changes, from Gonzalez-Cabrera (forthcoming): On social tolerance and the evolution of human normative guidance. Sections 4.4 and 4.5 of this chapter are based on worked published in Tomasello and Gonzalez-Cabrera (2017): The role of ontogeny in the evolution of human cooperation. These sections have been modified to meet university guidelines.

Signed,

Ivan Gonzalez-Cabrera

To my partner and friend, Ana.

Acknowledgments

This thesis represents the end of a long journey. The conception of the project started when I was a graduate student at the University of Tokyo. I am especially grateful to Tomoko Ishida, Ryota Morimoto, Senji Tanaka, Sohei Yajima, and all the members of the interdisciplinary group on evolutionary ethics and human ecology for their support during those early days. Thanks also to Yosaku Nishiwaki and Yoshiyuki Hirono for his encouragement to pursue a Ph.D. in philosophy. That project crystallized only years later as a graduate student at the Australian National University (ANU). My decision to move to Canberra to gain further training in philosophy of psychology and biology is one of the best that I have made. It was the stimulating environment of the School of Philosophy at ANU that allowed the core ideas of this thesis to grow and flourish. I owe special thanks to my supervisor, Kim Sterelny. My decision to start working on human cognitive evolution is largely attributable to the profound impact that his impressive and insightful body of work has had on many philosophers of my generation. He was always supportive and generous in spirit. I am honored and enriched for having been one of his students.

I learned from Kim's students almost as much I learned from him. I am grateful to Rachael Brown, Heather Browning, Carl Brusse, Brett Calcott, Adrian Currie, Ben Fraser, Jessica Isserow, David Kalkman, Stephan Kubicki, Christopher Lean, Kim Shaw-Williams, and John Zerilli for their friendship and intellectual rigor. I am also glad to have met through them much of the PhilBio community in the Sydney area including Pierrick Bourrat, Paul Griffiths, Adam Hochman, Linus Huang, Richard Menary, Dominic Murphy, and Elena Walsh. I am also grateful to Lindell Bromham, Peter Godfrey-Smith, Russell Gray, Peter Hiscock, Mateo Mameli, Claudia Passos, Kyle Stanford, and Michael Weisberg for their kindness and patience as well as for bringing clarity out of my often confused ideas.

My debts to the members of the ANU's School of Philosophy are many, but a few of them deserve special mention. Thanks to Alma Barner and Daniel Gregory for their constant companionship and insight. Thanks also to Rachael Briggs, Justin Bruner, Zoe Drayson, Edward Elliott, Anya Farennikova, Liz Irvine, Colin Klein, Leon Leontyev, Hanti Lin, Gabriel Rabin, Paolo Santorio, Wolfgang Schwarz, and Julia Staffel. Many thanks to all the faculty members of the school but, in particular, to David Chalmers, Alan Hájek, Victoria McGeer, Daniel Nolan, Philip Pettit, Nicholas

Southwood, and Daniel Stoljar, which gave me crucial advice at some point or another during my time at ANU.

I am also grateful to Cristina Bicchieri, Gilbert Harman, and Stephen Stich for allowing me to take part in the Penn-Princeton-Rutgers Graduate Seminar in Moral Psychology. I had the opportunity to discuss earlier versions of this work with them and their students who provided me with valuable feedback. Much of the material I worked out during those days did not make the cut to the final version of the thesis, but I was decisively encouraged by those discussions to do empirical research in moral psychology. Much of my recent work would not be as empirically oriented as it is now if it were not for them.

I also owe a great debt to the Department of Comparative and Developmental Psychology at the Max Planck Institute in Leipzig, which for a long period of my Ph.D. was my second home. The influence of experimental psychology on my philosophical work is obvious, and my sincerest thanks go to Michael Tomasello, Josep Call, and Malinda Carpenter for taking me in and for providing me with unprecedented opportunities. During this period, I had the opportunity to work and discuss with biologist, psychologists, and philosophers alike. All of us were interested in mutually interconnected topics ranging from the ontogeny of human social cognition to language acquisition to primate cognition. I am especially in debt with Avi Benozio, Manuel Bohn, Shona Duguid, Jan Engelmann, Sebastian Grüneisen, Susanne Hardecker, Robert Hepach, África de las Heras Martínez, Bailey House, Margherita Isella, Patricia Kanngiesser, Cathal O'Madagain, Federico Rossano, Alex Sanchez, Martin Schmelz, Barbora Siposova, Gregor Stöber, and Margarita Svetlova. Although most of the experimental work I carried out during my time in Leipzig was not included in the thesis, I have to give special thanks to Jana Jurkat, Roger Mundry, Dairon Rodriguez, Marco Schmidt, and Christoph Völter for their collaboration, guidance, and support.

I finish the writing of this dissertation in a heady six-month period in Vienna as a Writing-Up Fellow at the Konrad Lorenz Institute for Evolution and Cognition Research. I would not have been able to finish my Ph.D. without the unconditional support of Gerd Müller, Johannes Jäger, Isabella Sarto-Jackson, and Eva Lackner. Thanks also to all the fellows at the institute, Lee Altenberg, Wesley Herb Anderson, Argyris Arnellos, Lorenzo Baravalle, Riana Betzler, Daniel Brooks, James DiFrisco, Chiara Elettra Ferrario, Barbara Fischer, André Strauss, Wim Hordijk, George McGhee, Celeste Marina Pérez Ben, André Strauss, Zsóka Anna Vásárhelyi, and Berta Verd with whom I had the privilege to share the final version of this thesis.

This research was supported by an Australian Government Research Training Program (RTP) Scholarship. This project was also partially supported by a PhD Research Contract from the Max Planck Institute for Evolutionary Anthropology and a KLI Writting-Up Fellowship from the Konrad Lorenz Institute for Evolution and Cognition Research.

Abstract

This thesis focuses on the evolution of human social norm psychology. More precisely, I want to show how the emergence of our distinctive capacity to follow social norms and make social normative judgments is connected to the lineage explanation of our capacity to form shared intentions, and how such capacity is related to a diverse cluster of prototypical moral judgments. I argue that in explaining the evolution of this form of normative cognition we also require an understanding of the developmental trajectory of this capacity. For this purpose, the thesis is organized as follow. In the first chapter, I make some methodological remarks and provide the general overview and plan for the dissertation. In the second chapter, I explain what my explanatory target is and why it matters. On the view I am defending, shared intentional psychology gives rise to a special form of psychology that enables us to engage in social normative thinking. These norms are represented as shared intentional states. Moral psychology, in contrast, is more diverse. For moral judgments define a quite heterogeneous class of mental states—although some moral judgments may involve the representation and execution of norms, certainly not all of them do. I show that although much of our distinctive social norm psychology can be explained within the framework of shared intentionality, moral judgments cannot be unified in the same way. In the third chapter, I provide the baseline of social-cognitive capacities that serve as starting point for my lineage explanation. I argue that hominin social cognition was for a very long period of our evolutionary history essentially a matter of low-level cognitive and motivational processes. On this picture, bottom-up affective processes regulated the social lives of early hominins without requiring any special top-down mechanism of normative thinking such as a capacity for understanding and representing social norms. In the fourth chapter, I argue that human-like social norm psychology evolved as a result of the selective pressures that gave rise to shared intentionality, especially the demands that came from collective hunting. Yet collective hunting was not the whole story of the evolution of shared intentionality, for our capacity to form shared intentional mental states emerged from the interplay between the selective pressures that led to cooperative breeding in humans as well as organized, goal-oriented, collective hunting. Thus, I propose an evo-devo account of shared intentionality and its normative dimension since I argue that explaining the evolution of this particular form of normative thinking crucially depends on information about the developmental trajectory of this capacity. Finally, in the fifth chapter, I focus on how social norms are acquired and how the way

we learn them gives rise to some prototypical cluster of moral judgments. Thus, this chapter returns to some of themes and arguments of the first chapter by explaining how the distinction between moral judgments and nonmoral judgments can be culturally transmitted.

Table of contents

Statement of Originality i

Acknowledgments i

Abstract iv

Table of contents vi

List of figures, tables, and boxes vii

Glossary ix

Chapter 1. Introduction 1

1.1. Methodological reflections 1

1.2. Overview 6

1.3. Structure of the thesis 8

Chapter 2. Fragmenting moral judgments and unifying the psychology of norms 11

2.1. Defining moral judgments 16

2.2. What does it mean to claim that moral judgments are a unified cognitive kind? 20

2.3. Social domain theory 22

2.4. Fragmenting moral judgments 31

2.5. The psychology of norms 36

2.6. From norms to morals 42

2.7. Conclusion 46

Chapter 3. On social tolerance and the evolution of human normative thinking 47

3.1. The demonic male hypothesis 49

3.2. The puzzle of hominin evolution 53

3.3. The mosaic hypothesis 55

3.4. Evidence for the model 59

3.5. Paleoanthropological support 65

3.6. The self-domestication hypothesis 71

3.7. Normative guidance 75

3.8. Conclusion 80

Chapter 4. Big game hunting and the evolution of human social norm psychology 81

4.1. Big game hunting 84

4.2. The interdependence hypothesis 88

4.3. Did meat make us moral? 98

4.3.1. Social expectations 100

4.3.2. Social rationality 104

4.3.3. Punitive attitudes 109

4.4. Cooperative breeding 111

4.5. Evolutionary development 120
4.6. Conclusion 128
Chapter 5. Sharing our normative worlds 130
5.1. Moral nativism 131
5.2. The ontogeny of human social norm psychology 135
5.3. Norm acquisition 144
5.3.1. Acquiring social norms as shared intentional states 147
5.3.2. The mechanisms of acquisition 150
5.3.3. Norm classification 154
5.4. Moral sentimentalism 157
5.5. Language and norms 167
5.6. Conclusion 171
Chapter 6. Conclusion 173
6.1. Summary 173
6.2. Limitations and future research 176
References 180

List of figures, tables, and boxes

Figures

Figure 2.1. Diagram of the relation between joint intentional states, normative judgments, and moral judgments 13
Figure 2.2. Diagram of the relation between transgressions, signature patterns, and types of justification in Turiel’s (1983) social domain theory 24
Figure 2.3. Diagram of the relation between shared intentionality, social norms, and prototypical moral judgments 44
Figure 3.1. Partial skeleton (right) and artist’s reconstruction (left) of a relatively complete female specimen of *Ar. ramidus*, ARA-VP 1/500 67
Figure 3.2. Traditional phylogenetic arrangement of five hominid species 69
Figure 3.3. Alternative evolutionary arrangement of five hominid species 70
Figure 3.4. Model of bonobo evolution due to self-domestication 73
Figure 4.1. Sequence of stages in the lineage of shared intentionality 87
Figure 4.2. Selective feedback loop between cooperative breeding and collective hunting of big game 127
Figure 5.1. Model of norm acquisition 145

Tables

Table 5.1. List of basic cognitive and motivational infrastructure for norm acquisition 151

Technical boxes

Technical box 3.1. Comparative analysis of cooperation in great apes and humans..... 57
Technical box 4.2. Comparative analysis of cooperation in great apes and humans..... 90
Technical box 4.3. Group-mindedness and the Pleistocene-Holocene transition.......... 97

Glossary

agent-independent representation a representation is agent-independent if its content does not specify any agent—e.g., when a child realizes that a particular role in a group activity can be variably filled by different agents over time. This form of agent-independency is a matter of degree. A representation is increasingly agent-independent when it relies less on the specifics of an agent or group of agents (see Nagel, 1986).

bird's-eye view representation a bird's-eye view is a distinctive perspective to represent social interactions. This perspective is a case of functional abstraction and role/occupant distinction. The notion of a bird's-eye view representation is a closely connected but different idea from that of an agent-independent representation. For an agent could deploy, in principle, a bird's-eye view representation of a certain task only when interacting with specific partners.

bottom-up cognitive process the processing of sensory and affective information that depends more directly on features of the stimulus input (see Rauss & Pourtois, 2013). top-down control is the reverse of bottom-up processing, i.e., the processing of sensory and affective information that is driven by more cognitive processes such as goals or intentions.

cognitive process an operation that affects an agent's mental content. Examples of cognitive processes are perception, memory, language, problem-solving, and abstract thinking. Cognitive processes have hierarchical structure. Many contemporary models of processes specify a superordinate level (e.g., processes of emotional control) that controls and monitors lower level processes (e.g., emotions). Similarly, models of representation comprise higher superordinate levels that represent information in a more general form and lower subordinate levels where information is represented more specifically.

collective intentionality intentional states that we share with larger social groups, rather than specific individuals with whom we engage in simpler forms of shared intentionality. Only some shared intentional states are real collective intentional states in the sense of involving, at least to some degree, an agent-independent representation of the social interaction. The distinction between shared intentional states and full-blown collective states is not only a matter of degree but also a matter of cognitive mechanisms. Collective intentionality requires a we-mode of representation of mental states, e.g., "We believe that *p*" or "We want to do *p*” (see also ‘we-mode representation’). Since we-mode representations are representations of intentional states that are held by individuals but which make fundamental reference to a collective formed in conjunction with the other individuals (Searle, 1990), collective intentional states are also irreducible collective.

collective mental state a collective mental state is any shared intentional mental state ψ in which the subject is represented in a we-mode. A we-mode representation of the subject of a shared intentional mental state is a plural representation of the agents $A_1, A_2, ..., A_n$ who collectively are in the mental state ψ such as in “We believe that *p*” (see also ‘we-mode representation’).

corrective attitudes see punitive attitudes.

emotion the term ‘emotion’ refers here to a set of valenced behavioral and concomitant physiological responses that correlate with specific subjective experiences (McClure, Botvinick, Yeung, Greene, & Cohen, 2007). These emotional processes can be understood as a subset of automatic processes that are quick to respond and produce stereotyped effects on behavior. They can be differentiated from other automatic processes because they are valenced, *i.e.*, because they have valuative significance, carrying a level of attraction or aversion to the events that evoke them (McClure et al., 2007, p. 206). For example, Ekman and colleagues (Ekman, 1999; Ekman & Friesen, 1971) have shown that some emotions such as some forms of anger and guilt are a distinctive class of psychological phenomena marked out by their automaticity, by unique behavioral and physiological signatures, and by the existence of homologous states in other primates. In addition, LeDoux (1993, 1996) has also distinguished between ‘cognitive computations’ which yield information about stimuli and the relations between them, and ‘affective computations’ which yield information about the significance of stimuli for the organism and lead to physiological and behavioral responses appropriate to that significance. Similarly, researchers also sometimes distinguish emotions from feelings (Damasio, 1994, 1999). Emotions are automatic response repertoires, while feelings are subjective or experiential counterparts of emotions (mental representations of physiological changes that characterize and are consequent upon processing emotion-eliciting objects or states).

emotion regulation see emotional control.

emotional control the ability to exercise influence over emotion, and modulate emotion through the use of cognitive or behavioral strategies (Gross, 1998b; Lazarus & Folkman, 1984). It can be an attempt to change and regulate aspects of a situation and emotional experience prior to the generation of emotion or an attempt to alter the responses to the experience of emotion after the emotion has occurred (Gross, 1998a, 1999, 2002).

generalizability a property of social norms as defined in this thesis (see also 'social norms'). Generalizability refers to the scope of the judgment and its context-sensitivity—normative beliefs can regulate the behavior of only some individuals in specific situations or roles. Thus, normative beliefs can be characterized by a gradient of generalization or abstraction depending on how tightly they are conceived to be linked to specific individuals or situations.

I-mode representation a representation of intentional states that are fundamentally egocentric such as "I believe that *p*" or "I want *p*". When a state of affairs is represented as something that each agent individually believes, desires, or intends, this state of affairs is represented in the I-mode. The mode of a representation captures the subject's perspective or attitude on the intentional object. There are different forms of I-mode intentional states such as beliefs, desires, and intentions. Each I-mode intentional state has a we-mode counterpart. For example, an I-mode representation of an agent's intention ϕ is called an 'I-intention' of the form "I intend to ϕ", which has a we-intention counterpart of the form "We intend to ϕ" (see also 'we-mode representation').

instrumental rationality a form of rationality that focuses on the necessary and most efficient means of achieving a certain goal. Failing to do so makes an agent instrumentally irrationally. For example, if an agent is thirsty and intends to drink water, it would be rational for the agent to look for a nearby pond. All other things being equal, it would be irrational either to look for a more distant pond or not to take any means to this end.

intentional mental content the intentional content of a mental state is a propositional specification of what this mental state is about. We ascribe other agents with intentional mental content and attitudes toward them as part of our practices to explain and predict behavior (Dennett, 1987). There is a causal connection that explains the emergence and success of these practices. They are explanatory and predictively successful because the prolonged demands on cooperation and coordination in our lineage have selected not only for an increased ability to predict each other behavior through these practices but also for making our behavior more readily interpretable by them (Sterelny, 2003). Our explanatory and predictive practices are then causally connected to whatever physical implementation of those putative mental states might be through our biological and cultural evolutionary history.

intentional mental state intentional mental states are the thoughts we ascribe to an agent whose content can be roughly specified through a proposition toward which the agent takes a propositional attitude (see also ‘propositional attitude’).

intentionality a property of mental states which consists in their being directed towards objects or events. Mental states that are characterized by its intentionality are intentional mental states. Examples of intentional states are beliefs, desires, and intentions. An agent cannot have a belief, desire, or intention without these mental states being about something.

intrinsic motivation a property of social norms as defined in this thesis (see also ‘social norms’). A mental state possess intrinsic motivation when it motivates behavior as an ultimate end rather than solely as a means to other ends. An agent A is motivated to do ϕ solely as a means to an end E if and only if (i) A is motivated to do ϕ, (ii) A is motivated to achieve E, and (iii) A is motivated to do ϕ only because A believes that doing ϕ suffices (or just promotes) achieving E. On the contrary, A is intrinsically motivated to do ϕ precisely when (i') A is motivated to do ϕ and (ii') for all other ends $E_1, E_2, ..., E_n$, that A has, it is false that A is motivated to do ϕ solely because A is a means to satisfying one or more of these other ends. This means that if A is intrinsically motivated to do ϕ, A may recognize that ϕ contributes to achieve some of these other ends, but this cannot be the sole motivation for A to do ϕ. Avoiding pain is an intuitive example of intrinsic motivation. People seek to avoid pain not only because they think this will contribute to some other goals but also because avoiding pain is something we want to do for its own sake (see Sober & Wilson, 1998, pp. 200-201). Intrinsic motivation is not the same as having overriding power. For example, social norms are understood in this thesis as having intrinsic motivation, but its motivational force can be overridden by the motivational force of other mental states. What is essential for a mental state to be an ultimate end is to motivate the agent to bring about a certain state of affairs irrespective of the way it promotes other agent’s goals. In contrast, instrumental ends are those mental states that determine the means through which an agent will bring about that state of affairs.

joint intentional mental state a joint intentional mental state ψ is the mental state of an agent who joins the mental state of another. An agent *A* joins the intentional mental state ψ of another agent *B* when (i) *A* is intrinsically motivated to be in ψ because *B* is in ψ, and (ii) *A* co-represents ψ in a way that *A* and *B* are represented as a singular subject who is in ψ as opposed to each one simultaneously being in ψ. *A* can represent *A* and *B* as a singular subject who is in ψ either through recursive mindreading (see also 'recursive mindreading') or through a we-mode representation (see also 'we-mode representation'). *A* is intrinsically motivated to be in ψ because *B* is in ψ if and only if *A* is motivated to be in ψ as an end in itself because *B* is in ψ. An intentional mental state ψ is co-represented by a group of agents $A_1, A_2, ..., A_n$ when $A_1, A_2, ..., A_n$ are in ψ. An agent *A* can actively join the mental states of another agent *B* when *A* is intrinsically motivated to intentionally interpret *B*'s behavior and joins his/her mental states. Unless further qualification is made, in all that follows 'joint mental states' will denote joint, shared, and collective intentional mental states.

lineage explanation an explanation that specifies a sequence of changes which aim to show how a certain trait or mechanism could be changed into another through a set of minor modifications. The aim of these explanations is to make plausible certain phylogenetic trajectories. They can be given at the level of phenotypic change as well as at the level of developmental mechanisms (Calcott, 2009).

mental representation a theoretical construct in cognitive science, especially in computational theories of cognition, according to which mental states and processes are constituted by the occurrence, transformation, and storage of information-bearing structures (representations) of some kind. Representations are intentional mental states with semantic properties such as content, reference, and truth-conditions.

mindreading the capacity to predict, explain, or understand the behavior of other agents by attributing intentional mental states to them.

moral judgment a heterogeneous class of mental states, which although might be unified at a certain level, is not unified at the cognitive level. For this reason, I focus in this dissertation on prototypical clusters of moral judgments that are central to the philosophical tradition in moral psychology. In particular, I focus on moral judgments as defined by the domain theory of moral development

(Turiel, 1983) and judgments that are considered inescapable and authority independent by the agent (Foot, 1972; Joyce, 2006; Mackie, 1977; Williams, 1985). Moral judgments can only be partially grouped at the cognitive level in terms of the properties stipulated by these theories.

normative cognition roughly construed, normative cognition is an umbrella term that covers different kinds of symbolically mediated normative information. At a very basic level, human and animal cognition is driven, modulated, and governed by symbolically mediated information about what to do in a particular situation. When a mental state plays this role in cognition, that mental state carries normative information. Some mental states carry normative information because they govern different aspects of individual cognition and behavior in a private way, i.e., they do not govern or regulate other agents' cognitive precesses or behavior. However, humans (and more controversially other primate species) can represent normative information that is fundamentally intersubjective and social. For humans can produce, share, acquire and implement different kinds of norms and conventions that not only govern own behavior but also other agents' behavior. For example, food taboos about what pregnant women can eat are represented in a way that apply to all pregnant women regardless of whether one is pregnant or not. In this thesis, normative cognition refers more narrowly to this social form of normative cognition (see also 'social normative thinking').

normative guidance a top-down capacity to understand and respond to commands (see also 'top-down cognitive process').

normative thinking see normative cognition.

norms see social norms.

offline cognition a form of cognition that occurs when an agent is not acting but reflecting on the world and its possible actions. Agents switch to offline forms of cognition to make more careful considerations in situations that require forward planning.

online cognition online cognition is concerned with immediate input from the environment. It deals with tasks that require fast moment-by-moment processing and involves an agent's active sensorimotor engagement with the world.

practical rationality a capacity for solving the question of what one is to do through reflection. Deliberation of this kind is practical because is concerned with action and because it is assumed that reflection about action itself directly moves agents to act.

propositional attitude a propositional attitude is a mental state held by an agent such as intending, believing, desiring, and the like, which possesses a direction of fit and expresses how we regard a certain proposition. An agent can have different propositional attitudes toward the same proposition—e.g., one can believe that *p* while simultaneously desiring that *p*. A propositional attitude can have either a mind-to-fit-world direction of fit such as beliefs, which are meant to reflect the world, or world-to-fit-mind direction of fit such as desires, which are meant to influence the world.

punitive attitudes a property of social norms as defined in this thesis (see also 'social norms'). Punitive attitudes refer to the motivational force people feel to police, punish, or correct others, including themselves, when they think that they have violated a norm. These punitive attitudes are sometimes salient and explicit, while other times are less evident. For example, in some cases, the violation of a norm engenders actual physical punishment. In other cases, the perceived violation of a norm engenders punitive attitudes like anger and blame (or shame and guilt) toward the transgressor. And yet in other cases, it just drives corrective behaviors that target the transgressor without any harshness or evident signals of reprisal—e.g., when an adult correct a child for the violation a certain norm of etiquette.

rational action an action that maximizes the expected utility of the outcome according to some model o rational action. Models of rational action assume that agents' decisions are a function of their preferences among a number of available choices. In these models, rational agents use available information to assess the probabilities of events and the cost/benefit ratio to determine their choices, acting consistently with the best option available. Preferences are assumed to be both complete and transitive. Preferences are complete when agents can always say which of two alternatives they consider preferable (or whether neither is preferred to the other). Preferences are transitive if an option *A* is preferred to an option *B*, and *B* is in turn preferred to an option *C*, then *A* is also preferred to *C*.

rationality see practical rationality.

recursive mindreading recursive mindreading is the ability to embed representations about mental states inside other mental representations. For example, an agent *A* may believe that another agent *B* believes that *A* believes that *p*. This capacity allows an agent *A* to represent multiple agents $A_1, A_2, \ldots, A_n$ as a singular subject who is in a joint mental state ψ as opposed to each one simultaneously being in ψ. A is able to represent both agents A and B as a singular subject who is in a joint mental state ψ through recursive mindreading if (i) A is in a mental state ψ, (ii) A believes that B is in a mental state ψ, and (iii) A believes that B believes that A is in a mental state ψ. Unlike we-mode representations, the ability to represent multiple agents as a singular subject via recursive mindreading is constrained by our cognitive capacity to embed representations about mental states inside other mental representations.

representation see mental representation.

self-domestication a process of social selection against aggression in humans that resembles the process of adaptation of other wild animals to humans without systematic human selective breeding (Hare, Wobber, & Wrangham, 2012; Wrangham, 2011).

shared intentional mental state a shared intentional mental state ψ is the mental state of an agent who shares that mental state with other(s). A group of agents $A_1,A_2,...,A_n$ shares an intentional mental state ψ when (i') each agent A_i in the group is intrinsically motivated to be in ψ because the others are in ψ, and (ii') $A_1,A_2,...,A_n$ co-represent ψ in a way that $A_1,A_2,...,A_n$ are represented as a singular subject who is in ψ as opposed to each one simultaneously being in ψ. Each agent A_i in the group is intrinsically motivated to be in ψ because the other agents are in ψ if and only if each of the A_i agents in the group is motivated to be in ψ as an end in itself because the others are in ψ. An agent A can actively share a mental state ψ in which A is in when A is intrinsically motivated to display and signal ψ for others to join.

shared intentionality the capacity to form joint intentional states. Joint intentional states are hybrid mental states, which are characterized by the systematic co-occurrence of cognitive and motivational components. From a cognitive point of view, representing activities in a joint form requires being able to represent plural subjects of action (e.g., “We want to hunt” or “We are attending to that antelope”) and task roles or activities that are not necessarily linked to particular individuals (e.g., agent-independent representations of the different roles that our particular hunting technique requires). This creates a gradient of generalizability and abstraction depending on the scope of the plural subject of action and the degree of abstraction with which these activities are represented. From a motivational point of view, shared intentional states are intrinsically motivational, e.g., by definition, entertaining a joint intention of the form "We want to hunt" implies a motivation to hunt with others that is not purely instrumental. That is, an agent joins the activity because he/she finds its collective nature intrinsically rewarding, rather than because the agent merely thinks that the activity is instrumentally beneficial for him/her, or even instrumentally beneficial for everyone.

social normative thinking a social form of normative cognition that deals with social norms (see also 'social norms'). On the view I am defending, shared intentional psychology gives rise to a special form of psychology that enables us to engage in social normative thinking by enabling the sharing of the kind of normative mental states that govern the instrumental and practical rationality of hominid agency.

social norms normative mental states that are generalizable, intrinsically motivating, and engender punitive attitudes. A normative mental state is generalizable when multiple agents can fall within the scope of the normative mental state and that normative state is applicable to multiple counterfactual situations (see also 'generalizability'). A normative mental state is intrinsically motivating when the agent is motivated to comply with that state as an end in itself rather than as a means to fulfill other end (see also 'intrinsic motivation'). A normative mental state engenders a punitive attitude if the agent is motivated to police, punish, or correct others, including themselves, when they think that they have violated the normative state (see also 'punitive attitudes').

social rationality a form of instrumental rationality that takes place when agents share intentional states such as goals and intentions (see also 'instrumental rationality'). Social rationality gives rise to social expectations about what one should do in a given situation as well as to social expectations about what others should do in turn to achieve a common goal. For example, when hunters share a goal such as "We want meat", this goal is only achieved when all the hunting partners obtain some of the meat. In such a context, sharing the meat is socially rational.

third-person representation see bird's-eye view representation.

top-down cognitive process the processing of sensory and affective input that is driven by higher cognitive processes such as goals or intentions. These processes occur at a superordinate level in the sense that they control and monitor lower level processes that handle incoming sensory and affective information. Bottom-up processing is the reverse of top-down processing, i.e., the processing of sensory and affective information that depends more directly on features of the stimulus input (see Rauss & Pourtois, 2013).

we-mode representation intentional states held by individuals but which make fundamental reference to a collective formed in conjunction with the other individuals (Searle, 1990). They are agent-independent representations of intentional states that cannot be reduced to recursive mindreading. For example, knowledge states can be represented through recursive mindreading such as "I believe that you believe that I believe that *p*". But these mental states can also have an irreducible we-mode of representation when the plural subject of that knowledge state cannot be represented via recursive mindreading due to cognitive limitations in our capacity to embed representations about mental states inside other mental representations. The mode of a representation captures the subject's perspective or attitude on the intentional object. As in the case of I-mode representations, there are different forms of we-mode intentional states such as beliefs, desires, and intentions (see also 'I-mode representations'). A we-mode representation of an agent's intention, for instance, is called a 'we-intention'.

Chapter 1. Introduction

The goal of this dissertation is to explain and defend a naturalistic theory of normative cognition—particularly, human social norm psychology. Social norms can be roughly understood as the rules that govern group behavior. Humans seem to be a unique species in this respect since our behavior is largely governed by a widespread network of social norms. They have been extensively studied in the social sciences, although they are typically seen as exogenous variables that constrain individual behavior. However, despite their vital role in organizing human life, there has been relatively little attention in philosophy and the psychological sciences to understand this phenomenon. Much existing experimental research on the psychology of social norms is relatively recent and more often than not, focused on the understanding of moral norms. In the philosophical literature, social norms are usually seen as equilibria of coordination games or cooperative equilibria in prisoner's dilemma-type games. But there is little mention of the mental processes involved and the implications for both philosophy of mind and philosophy of psychology. This thesis aims to fill this gap by building a naturalistic account of a distinctively human form of social norm psychology and the particular kind of normative thinking it engenders. Although partial and piecemeal, this thesis is intended as an important first step towards a more comprehensive theory of human norm psychology in particular and normative thinking in general.

1.1. Methodological reflections

Naturalism is understood here as a methodological approach—one that takes the philosophical practice to be continuous with the natural sciences. On this view, philosophical theories are empirically testable conjectures. Certainly, philosophy is often understood as a matter of pure conceptual analysis aimed to find necessary truths. But I am skeptical of the prospects of understanding the mind by this method alone. Granted, conceptual clarification is a major feature of the philosophical enterprise, but it is hardly a methodological tool under the exclusive control of philosophers. Conceptual analysis is, and has been always, an important component of the scientific practice. Of course, there will be parts of this project that will require more conceptual clarification than empirical input. This will be clear at times in certain parts of this work. But the whole theoretical enterprise that this thesis attempts will be carried out by a mixture of traditional conceptual analysis, empirical data, and some formal modeling—although I do not develops these models myself, many of those on whom I rely do.

Philosophy is not a scientific enterprise on its own, but it is part of our scientific worldview, for its goal is also the pursuit of knowledge. Philosophy and science can be integrated in different ways. One is by realizing that many of the most relentless problems in science are partly philosophical in nature. Another is by noticing that philosophical ventures are often integrative. Philosophical thinking typically takes place when our subject matter does not match (or it does only in a very partial way) with any of the established sciences. Understanding human nature, for instance, is a too wide topic to be monopolized by one of the sciences alone even though individually they can make significant contributions to the overall picture. Concerning its topic, this dissertation has also a distinctive philosophical flavor since it deals with one striking aspect of human nature, namely our capacity to engage in social normative thinking. Moreover, the project aims to make a substantive contribution to our understanding of this capacity by connecting this issue with what some philosophers have called ‘shared intentionality’ (see, for instance, Bratman, 1992; Gilbert, 1989; Searle, 1990; Sellars, 1963). Thus, although this thesis should be primarily understood as a first-order project in the empirical sciences, since despite its highly conjectural nature each of its parts is independently testable, it will become clear throughout this dissertation that this first order project is directly relevant to many other standard philosophical projects. As Godfrey-Smith (2014) put it, one way to do philosophy of science is focusing on scientific disciplines as target of philosophical investigation. But another way is to use science as a tool to gain an understanding of the universe and our place within it, i.e., “[...] working out what the raw science is really telling us, and using it to put together an overall picture of the world”. (p. 4) He calls the latter ‘philosophy of nature’. This thesis can be understood as project of the latter kind.

The scope of the dissertation is broad as it is often the case in philosophy. Roughly construed, the topic of this thesis is the evolution of normative cognition and, more specifically human social norm psychology.[1] Evolutionary approaches in

---

[1] I will not provide here definitions of key terms such as ‘normative thinking’ or ‘human norm psychology’. One of the central goals of the thesis is precisely to shed light on those issues by trying to build a psychological theory of a special form of normative thinking. This is part and parcel of the naturalistic spirit of the project, since on this view, normative thinking, human norm psychology, and the like, are psychological phenomena whose very nature has to be discovered using the tools of the empirical sciences.

philosophy are now in vogue. Debunking arguments of morality, for instance, are typically grounded in evolutionary considerations and they have been shown to have wide philosophical reach. Yet historical approaches are anything but new. They have been a common strategy in the humanities and the social sciences to understand human nature. Unsurprisingly, they have been an important part of the philosopher's toolkit since an important part of understanding who we are is to understand how we came to be.

In tackling the evolution of human social norm psychology, I also focus on its development as well as its underlying cognitive and motivational mechanisms. This might indeed sound too ambitious. But as I will try to show in the following chapters, this is a necessity because I am interested in the psychological underpinnings of this form of normative thinking and how they evolved. And as I will explain in chapter 3, explaining the evolution of this form of human norm psychology requires understanding the developmental trajectory of this cognitive capacity. So, in order to explain its evolution, we need an idea about both the cognitive machinery and the developmental trajectory of human social norm psychology. I think we can make significant progress in the historical sciences through these integrative views since they allow putting together multiple independent lines of evidence, which increasingly constrain the space of possible explanations.

Reconstructing the past is certainly difficult. The causal connections between current evidence and past events are remote and the traces of these processes are usually eroded by time. Historical processes typically destroy evidence about themselves as they unfold. For example, when an organism dies, it triggers a process of cellular decay, which is only prevented by the mineralization of small portion of tissues such as bones and teeth. The chances of an organism becoming fossilized are poor; the probability of intact preservation is even lower. Fossils usually suffer damage from weathering and sedimentation before they can be discovered. These evidential problems only escalate when our explanatory targets are social-cognitive traits since they can only be indirectly inferred from the physical evidence that these processes leave behind. Unlike bones, behavior does not fossilize. Thus, evidential issues become a central problem for the historical sciences, and this problem is more acute in the case of our evolutionary history. Reconstructing the social behavior of early hominins requires serious evidential bootstrapping. No single strand of evidence can tell us everything about the social behavior of our ancestors in the distant past.

One way to do this is by comparing lineages that may serve as models of our explanatory target. Closely related species may provide valuable information about our past due to their phylogenetic connection with us—e.g., information about ancestral morphology, social organization, or baseline cognitive capacities. Also, given enough symmetries between relevant selective pressures, even unrelated taxa can offer a glimpse into our evolutionary past through parallel, convergent evolution. The comparative strategy certainly does not solve the problems of scarcity and low resolution of evidence that plague the historical sciences, but it helps us to deal with them. Admittedly, we can rarely give an accurate portrait of the past. Yet well-crafted comparative analyses within a reasonably tractable chunk of our lineage can yield meaningful conclusions. For example, comparative studies in great apes and humans have concluded that unlike children, apes do not engage in true imitation (learning focused on actions) but rather emulation (learning focused on goals and results) (Tomasello, 1996). As a result, emulation is thought to be one of the baseline capacities of hominin evolution, while true imitation is not (for a different view, see Whiten, McGuigan, Marshall-Pescini, & Hopper, 2009).[2]

Moreover, the paleoanthropological record can help us to determine the significance of the similarities and differences we find between closely related species. Fossil evidence can support comparative claims by telling us how compatible our comparative analyses are with the physical traces of the past. Phylogenetic reconstructions via common ancestry are highly sensitive to patterns of dental abrasion, skull morphology, and the like, as phylogenetic inferences generate hypotheses about our ancestors that can be later confronted with their physical imprint. For example, the number of branches in our evolutionary tree has grown significantly since ‘Lucy’ was discovered—in fact, all the three genera predating this specimen, *Sahelanthropus*

---

[2] Whiten and colleagues (2009) have argued that chimpanzees possess a minimal capacity for imitative learning since diffusion studies in chimpanzees suggest that this species have a capacity for copying local behavioral traditions and experiments involving so-called ‘ghost’ experimental conditions have shown that chimpanzees cannot learn if there is no agent demonstrating the action. However, in all these studies, the specific details of the actions are ignored. Since chimpanzees are not completely blind to goal-directed behavior, observing an action could potentially transmit information about the demonstrator’s goal, which helps the chimpanzee to learn the relevant actions to achieve the desired goal. Therefore, emulation could still be a mechanism for acquiring these behaviors since the focus of attention would be placed on the goal of the demonstrator rather than on the actions themselves. This explanation has been argued to be consistent with the primate literature on social learning (Call & Carpenter, 2002; Dindo, Thierry, & Whiten, 2008; Galef, 2009; Tennie, Call, & Tomasello, 2006, 2009; Tennie & Hedwig, 2010).

(Brunet et al., 2002, p. 150), *Orrorin* (Senut et al., 2001), and *Ardipithecus* (White et al., 2006), were discovered after this finding. The discovery of these hominins has crucially changed the way we think about the origin of human bipedalism (White, Lovejoy, Asfaw, Carlson, & Suwa, 2015). Likewise, convergent selective pressures require structurally similar ancestral environments and parallel socioecological pressures. Luckily, paleoenvironmental data frequently allows a reliable reconstruction of the geology, flora, and fauna of early hominins (see, for instance, Hart & Sussman, 2005).

All this evidence is framed in the context of the evolution of human sociality and cooperation. The type of questions I am interested in are questions like “How did the human mind change in response to our living a more prosocial, cooperative life?” and “Where does human social norm psychology fit within this change?” So, most of the data I offer in this thesis comes from well-established models of great ape and human sociality and cooperation, which are built upon comparative and paleoanthropological data. These models serve as proxies of early hominins’ and early human hunter-gatherers’ social-cognitive capacities as well as the crucial differences between, for instance, early hominin foraging practices and human foraging.

More specifically, I aim to provide in the following chapters what can be characterized as a ‘lineage explanation’ of this form of normative thinking (see Calcott, 2009), i.e., an explanation that specifies a sequence of changes that takes us from agents with an ape-like baseline capacity for social cognition to agents with human-like social norm psychology. The explanation I offer in this dissertation relies, in particular, on a model of great ape and human cooperation that has been built over the last two decades of psychological research within the theoretical framework of shared intentionality. I want to argue that human social norm psychology is closely linked to our shared intentional psychology and that this capacity is in turn connected to moral cognition—in particular, some prototypical cluster of moral judgments that is central to the philosophical tradition in moral psychology. Thus, the primary target of this thesis is social norm psychology rather than normative phenomena *per se*. Most importantly, I want to link the emergence of this capacity with the lineage explanation of our capacity to form shared intentions.

As we will see in the following chapters, the lineage explanation I want to defend in this thesis bears on some traditional metaethical debates in philosophy. Metaethics is sometimes narrowly understood as a discipline focused on moral language but it might be more broadly construed as the study of the metaphysical, epistemological, semantic, and psychological presuppositions and commitments of moral thought, talk, and practice (Sayre-McCord, 2007; Schroeder, 2012). I understand the metaethical consequences of this lineage explanation in the latter, broader sense. For example, the proposed lineage has important consequences for debates on moral nativism since it relies on the evolution, tuning, and integration of domain-general mechanisms. Likewise, it also bears on debates about the mental content of moral judgments since joint intentional states are hybrid mental states, which are partially cognitive and partially motivational.[3] As a result, the class of moral judgments that are shared normative thoughts is hybrid as well.

1.2. Overview

My goal here is to provide an explanation of how the basic psychological machinery for hominin sociality and cooperation worked at each major stage in the evolutionary trajectory leading to modern humans. Differences between adjacent stages are of great importance here. For they give us crucial information about how hominin basic social-cognitive machinery changed into another working mechanism through a sequence of relatively minor modifications. In the view I am proposing, for instance, subtle cognitive and motivational changes in the direction of shared intentionality explain not only much of the differences in cooperation between great apes and humans but also the increase in complexity of the hominin foraging practices, e.g., as it is evidenced by the gradual appearance of big game hunting in the fossil record. This sequence of changes shows the plausibility of a certain evolutionary trajectory. It would be implausible, for instance, to speculate about the origins of collaborative big game hunting in a highly aggressive and dominant hominin species (Boehm, 1999). It is true that some carnivores such as wolves and hyenas hunt in groups. But their hunting strategies were neither evolutionarily accessible to the hominin lineage nor a case of real collaborative hunting, as I will argue in chapter 4. Thus, one advantage of bringing additional information to bear on the problem is that it introduces new constraints on our evolutionary narratives—e.g., through data in comparative and developmental psychology. These hypotheses are not just-so stories. They gain their epistemic credentials because building them gets harder

---

[3] That is to say, I conceive of shared intentional states as the regular, typical, and systematic co-occurrence of these cognitive and motivational components. While I think it is possible for one to appear without the other, both in development and evolution, the regular co-occurrence and recruitment in driving human behavior is what will matter.

as the number of possible trajectories that met the imposed conditions significantly reduces.

To put it more precisely, I want to show how the machinery behind the psychological phenomenon of shared intentionality is behind human social norm psychology and how shared intentionality emerged from variation in the psychological machinery for social cognition within our lineage. I will argue that human social norm psychology relies on our shared intentional capacities because the underlying cognitive and motivational machinery behind shared intentionality accounts for a wide range of core cases of social normative thinking, which have played a central role in human evolution. In other words, my working hypothesis is that much of human social norm psychology is indeed a special form of our distinctively human shared intentional psychology.

This approach is particularly important for those interested in issues about continuity in evolutionary explanations. Lineage explanations can be given at the level of phenotypic change, but they can also be offered at the level of developmental mechanisms (Calcott, 2009). I aim to offer a lineage explanation of the latter kind. Since shared intentionality can be decomposed into a number of components, one may explain the psychological continuity of our lineage in terms of these components and how they are reassembled and rearranged in development. To be clear, when I talk about incremental evolutionary explanations, I expect the reader to see some continuity in my explanation because the components used to describe hominin social cognition can be credibly tuned and reorganized in development in such a way that they give rise to new cognitive skills such as the capacity for shared intentionality. In principle, these gradual changes in development might lead to abrupt changes in adult phenotypes. Therefore, the continuity I am arguing for is developmental, i.e., it is given in terms of gradual changes in hominin psychological development.

This approach also allows us to explain the emergence of key cognitive innovations. I argue that the ability to form shared intentional mental states was an evolutionary novelty in our lineage—very likely a unique capacity within the mammalian clade. As we will see in this thesis, an evolutionary developmental approach can use lineage explanations to explain the origins of novel cognitive capacities through subtle changes in developmental timing. This means that developmental data in psychology is central to this project since they provide crucial information to explain how novel, human unique psychological traits such as shared intentionality and its normative dimension are generated in ontogeny through relatively subtle changes in the cognitive and motivational machinery that support prosocial, cooperative behavior.

1.3. Structure of the thesis

The rest of the thesis is organized as follow. In chapter 2, I start explaining what my explanatory target is and why it matters for both philosophy and psychology. I will focus on the evolution of our general capacity to grasp social norms and to make a special class of normative judgments that I will call ‘shared intentional normative judgments’, explaining why this discussion is important for our understanding of human moral thinking. On the view I defend in this dissertation, human social norm psychology is a form of shared intentionality. These norms are represented as shared, collective intentional states that create emergent, social level facts. There might be other kinds of norms, but the relevant sense of norm I will discuss in this dissertation is social, i.e., they are norms about social interactions that are also expressed and endorsed in public contexts. Since I assume that norms have this social nature, I will not focus on normative mental states that agents keep for themselves—they will not count as norms in the relevant sense. I will try to show in this thesis that this particular form of normative thinking defines a unified cognitive kind that has played a central role in human evolution.

Moral psychology, in contrast, is more diverse. For as I will argue in the next chapter, moral judgments define a quite heterogeneous class of mental states. Sure, normative cognition and moral thinking are related to each other since moral thinking is normative thinking. But I do not believe that shared intentionality unifies moral judgments in the same way that, say, it defines a clear class of normative mental states, namely shared intentional normative thoughts. Indeed, some members of the class will be moral judgments, i.e., the ones that are the result of the kind of norm psychology I propose in this dissertation. But I am not arguing that all shared intentional normative thoughts are moral or that all moral judgments belong to the class of shared intentional normative thoughts. Shared intentionality singles out only one particular class of normative thinking, so I do not even argue that all moral judgments are a special case of shared intentional states. I left this question open to empirical scrutiny.

In chapter 3, I aim to provide the baseline of social-cognitive capacities from which my lineage explanation begins. To a large extent, this is a comparative task, i.e., I try to argue for a particular picture of early hominins based on comparative data about the social-cognitive skills of great apes. I argue that, for a very long period of our evolutionary history, social behavior was basically a matter of tuning our preexisting affective mechanisms against the backdrop of our great ape-like baseline of social-cognitive capacities. On this picture, early hominin ancestors' social lives were mainly regulated by bottom-up affective processes. It was not organized around social norms. No sensitivity to commands was required. So, at the end of this chapter, I suggest that high-level cognitive processes of normative guidance evolved as a consequence of the selective pressures that led to shared intentional states—more specifically, the demands that came from collective hunting.

I develop this view in some detail in chapter 4. I argue that shared intentionality emerged from the interplay between the selective pressures that led to cooperative breeding in humans as well as organized, goal-oriented, collective hunting. In this story, selection did not only act upon adult cognitive capacities but rather upon the entire life cycle. Very basic capacities for shared intentionality such as joint attention and pointing behavior emerged as ontogenetic adaptations for early childhood in the context of cooperative breeding. These capacities were extended in development into adulthood where they were redeployed in the context of cooperative foraging, especially collective hunting. As a result, the motivation and commitment with the shared goals and intentions of these collective activities were carefully monitored and enforced by all the group members. The coupling of these punitive or corrective attitudes with some other common features of shared intentional states was the origin of human social norm psychology. So, my claim is not only that shared intentionality was an essential precursor to normative thinking but also that some special class of shared intentional states provides a well-defined target for the psychological study of normative thinking. Human social norm psychology is a special form of shared intentionality that is linked to punitive attitudes toward norm violators.

In chapter 5, I shift the focus from the phylogeny of normative guidance to its ontogeny, particularly the process of norm acquisition. To put it poetically, I think that norm acquisition is fundamentally a process in which we share our normative worlds; hence, the title of this thesis. We acquire social norms because we share intentional states with a particular motivational profile. This motivational profile is characterized by an intrinsic motivation to comply with the norm and a variety of punitive attitudes towards norm violators. In fact, we sometimes categorize normative judgments based on this motivational profile. I argue in this chapter that the way we learn these norms not only gives rise to a prototypical cluster of moral judgments, which has been traditionally associated with the sentimentalist tradition in moral philosophy, but also that we sometimes classify these judgments as moral in virtue of their own motivational and affective profile. Take for example norms prohibiting stealing. Transgressions of this norm reliably trigger reactions of anger toward thieves or guilt for stealing. So, we learn to identify these norms and the responses they elicit as moral because a disposition to react in such a way in these situations is prototypically linked to moral responses. They form a very peculiar cluster of punitive attitudes. These affective and motivational dispositions help us to explain why we perceive certain judgments to be prototypically moral rather than just normative like the rules of conventional games. For they make us perceive normative judgments as being independent and authority independent, which are prototypical properties of moral judgments. Thus, this chapter returns to some of themes and arguments of chapter 1 by explaining how the distinction between moral judgments and nonmoral judgments can be culturally transmitted and by explaining how moral cognition can be prototype- or exemplar-based.

Chapter 2. Fragmenting moral judgments and unifying the psychology of norms

One of the central aims of this thesis is to use philosophical tools to advance our understanding of normative and moral cognition. More specifically, I want to show how and why what some philosophers have called ‘shared intentionality’ (see, for instance, Bratman, 1992; Gilbert, 1989; Searle, 1990; Sellars, 1963) can help us to understand our distinctive capacity for normative thinking and how this capacity is closely related to a particular cluster of prototypical moral judgments. Thus, the primary target of this dissertation is normative and moral cognition rather than normative or moral phenomena, i.e., the features of the world that normative and moral judgments are about.

The goal of this chapter is to explain the relationship between the type of normative thinking on which I focus in this thesis and moral judgments. On the view I propose, the distinctive shared intentional capacities of the hominin lineage help us to define an important form of normative thinking, namely shared intentional normative thoughts. These thoughts are generated by a peculiar, but important, kind of psychology of norms. Hereafter, I will understand norms as social norms. There might be other kinds of norms but my interest will focus on the psychology of this particular class of norms. Roughly speaking, social norms are rules of behavior that are considered acceptable in a group or society. These norms are represented in the form of shared intentional states, or so I will argue. When a norm is executed in cognition, it generates shared or joint intentional normative judgments. Since joint intentional states are hybrid mental states, I will argue in this chapter that shared intentional normative judgments are also hybrid mental states, i.e., states which are functionally defined by both its cognitive component and its motivational component. I will explain the essential psychological properties that this kind of normative thinking picks up in virtue of these components.

The kind of psychology of social norms I will propose accounts for a large and important spectrum of our normative cognition. Moral judgments, in contrast, are more diverse. For as I will try to argue in this chapter, they seem to define a quite heterogeneous class of mental states. Although in metaethics the term ‘moral judgment’ is sometimes used to denote a kind of speech act, the term is being treated here as denoting a mental state.[4] We can isolate different kinds of prototypical moral judgments that map onto particular psychological processes, e.g., moral judgments about harm, justice, or rights which possess a particular functional profile (see Nucci & Turiel, 1978; Nucci, Turiel, & Encarnacion-Gawrych, 1983; Turiel, 1983, 1998) or judgments about moral demands that are seen by the agent as inescapable and authority independent (see Joyce, 2001; Mackie, 1977; Mameli, 2013). But we can do this only in a fragmentary fashion because these theories define local psychological kinds that only cover particular cases (I will expand on this issue in section 2.5). To put it briefly, while we can define within the shared intentional framework a family of normative judgments as what some philosopher would call a 'natural kind' (Quine, 1969), moral judgments as a whole are not a natural kind in the psychological sciences but a family of prototypical kinds (Machery, 2012; Parkinson et al., 2011; Sinnott-Armstrong, 2008; Sinnott-Armstrong & Wheatley, 2012, 2013; Sripada & Stich, 2007; Stich, 2016). In this chapter, I will focus mainly on the negative part of this claim, i.e., that there are good reasons to doubt that moral judgments define a unified natural kind in psychology. At the end of this chapter, I will outline the general strategy that I will follow in the rest of thesis.

This way to conceive of moral judgments is important for the overall argument of the thesis. Thus, before getting into details, it may be important to explain how this chapter fits into the larger landscape of my project. For to say that moral judgments are not a unified target for the psychological sciences is not the same as saying that they cannot be a subject of study for psychology. It means that our understanding is piecemeal and that we should advance research without a preconceived view of what make some mental states to be moral judgments. Certainly, normative cognition and moral thinking are related to each other since moral thinking is always normative. This is a plausible working assumption. As a result, there are reasons to think that the kind of normative thinking that shared intentionality defines subsumes an important class of moral judgments. Since there are plausible explanations of the ontogeny and phylogeny of shared intentionality, there are grounds to connect the lineage explanation of our shared intentional capacities with the psychological machinery of prototypical kinds of moral judgments (see figure 2.1).

---

[4] If one understands metaethics more narrowly as the study of moral language, one might be inclined to treat moral judgments as some kind of speech acts rather than as a class of mental states. Of course, understood as speech acts they may perhaps express a mental state, but that is a different matter.

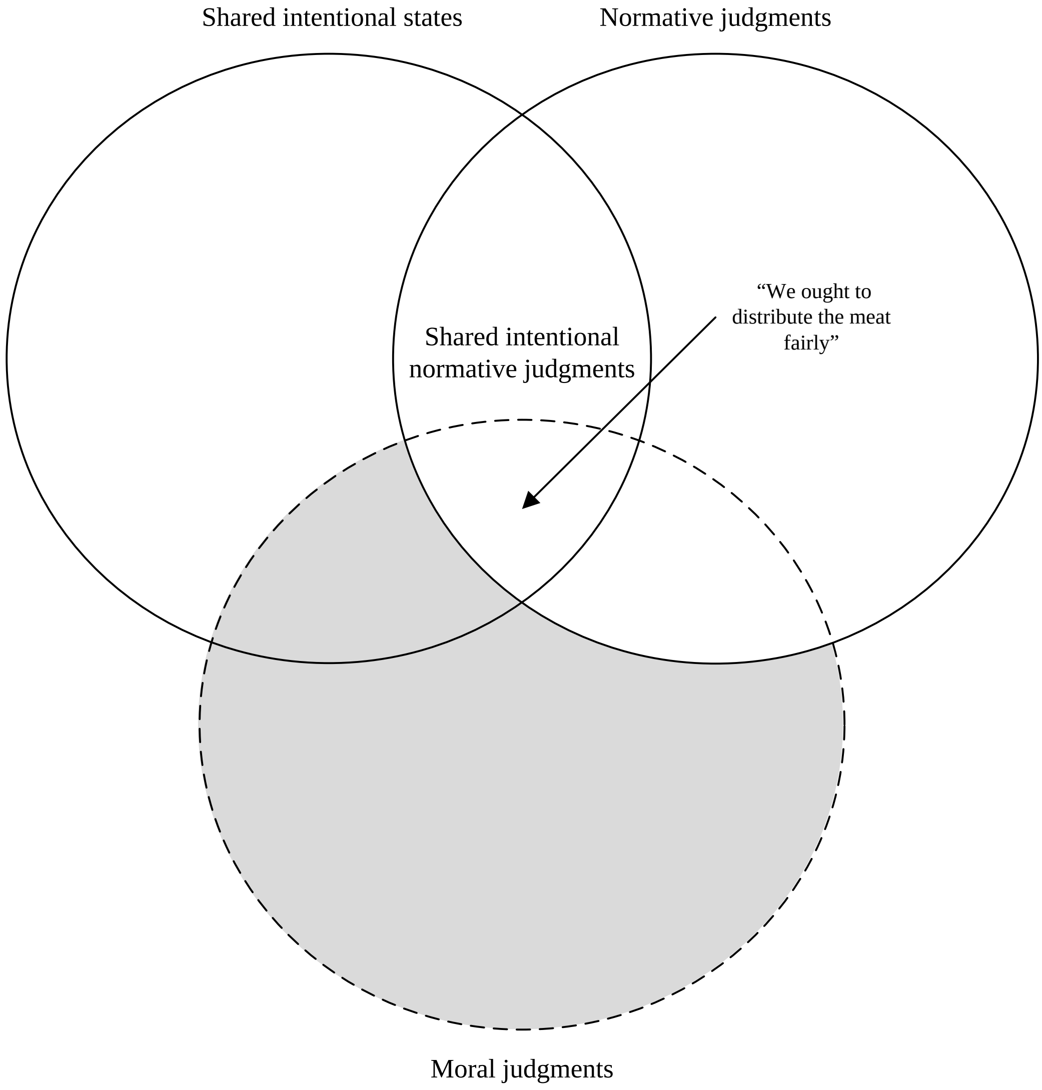

Figure 2.1. Diagram of the relation between joint intentional states, normative judgments, and moral judgments. Shadow zones represent empty zones. All moral judgments are assumed to be normative but not all moral judgments are taken to be a special class of joint intentional states, i.e., they are not shared intentional normative judgments. Normative and moral judgments include claims that are perceived by the agent as right, true, correct, valid, or justified, as well as those that are perceived by the agent as wrong, false, incorrect, invalid, or unjustified. Dashed lines indicate that the domain is not a natural kind but likely possesses a prototypical structure.

To be clear, shared intentional normative thoughts are a special and important subclass of normative thoughts. This means that there are normative thoughts that are not shared. One can imagine thoughts that regulate individual behavior in rather egocentric way such as the normative thoughts that stem from individual instrumental rational action (Camp & Shupe, 2017). Non-human animals seem to be able of some sort of instrumental reasoning. Chimpanzees, for instance, display a form of strategic reasoning to outsmart conspecifics, although they do not hold others to the same standards of rational action. Humans are a much more complex case. If I want to drink water from the drinking fountain, I think that I should go to the fountain. But unlike apes, we are inclined to hold others to our standards of rationality. When someone does not meet our expectations of what the rational course of action should be, we often (and perhaps always) judge others as stubborn, incompetent, or just simply foolish. To be clear, it is not that norms of social interaction are shared while norms of reasons are not. Instead, the point is that great apes do not share these mental states at all—these normative mental states are useful to outcompete conspecifics rather than to cooperate with them.

The evolution of the capacity for shared intentionality is foundational for our capacity to form the class of shared intentional normative judgments. Moreover, an important class of moral judgments, though not all moral judgments, are shared intentional normative thoughts—i.e., some shared intentional normative thoughts are just not moral judgments. As I will argue in chapter 4, the psychology of the norms that govern conventional games fall within the spectrum of shared intentionality. These norms apply selectively to those that we think share with us the relevant information about the rules of the game, so we enforce those norms only on them and expect others to do the same when someone violate the rules. But norms of conventional games do not lead to prototypical cases of moral judgments—at least not of the kind that we will discuss in this chapter.

Shared intentionality gives us a framework that allows singling out a particular class of normative judgments. As we will see in section 2.5, these normative judgments can be characterized by a distinctive gradient of generalizability, intrinsic motivation, and punitive attitudes. There might be other kinds of social normative thoughts that are not covered by this account. Yet as we will see throughout the thesis, they define a class of social normative judgments that is central to our social lives. They are normative because they give rise to a certain form of social rationality (e.g., about how we should share the meat among hunting partners) and because we enforce them in ourselves and others (e.g., when excluding partners who do not meet our expectations). But although judgments like “We ought to share the meat fairly” may perhaps significantly overlap with prototypical moral judgments, the framework does not define by itself any special class of moral judgments.

These judgments, however, could be said to be moral because they are linked to features of prototypical moral judgments. They could be, for instance, judgments about practical demands that are considered inescapable and authority independent by the agent, and which have traditionally played a central role in Western moral philosophy (Foot, 1972; Joyce, 2006; Mackie, 1977; Williams, 1985). These features go beyond the explanatory scope of the shared intentional framework, so explaining their emergence is explaining how the shared intentional lineage split into particular branches of moral judgments due to the ramification of our cultural practices and how we learn them. Perhaps all moral judgments are a special case of shared intentionality but I am also cautious about this generalization. For unconscious, automatic responses may also count as moral judgments, though they are not joint intentional states. For example, implicit aversive reactions against utilitarian harms are often taken to be exemplars of moral judgments, although they are not joint intentional states.

A cautionary note is in order here. In the cognitive science literature, dual process theories distinguish between two different types of cognitive process. Type 1 processes that are usually characterized as fast, automatic, and non-conscious, and type 2 processes that are typically depicted as slow, controlled, and conscious (Evans & Frankish, 2009; Frankish, 2010; Frankish & Evans, 2009; Kahneman, 2003; Stanovich, 1999; Stanovich & Toplak, 2012). The framework I will develop in the following chapters is primarily intended as an account of the second type of processes. As a result, I will leave open the question about the extent to which this account can explain implicit, automatic, normative or moral responses.

The rest of this chapter is divided as follow. In section 2.1, I will provide the philosophical background of the discussion, i.e., the problem of providing a unified account of moral judgments. In section 2.2, I will talk about what I call the ‘unification hypothesis’, i.e., the idea that there is a property (or set of properties) that unifies moral thinking as an explanatory target for psychology. In section 2.3, I will focus on some particular version of this hypothesis, which relies on the domain theory of moral development (Turiel, 1983), and explains why it fails. In section 2.4, I will argue that we are better off by fragmenting moral judgments into different classes of prototypical moral judgments. In section 2.5, I will argue that although moral judgments may not be a natural kind, an important form of norm psychology can be unified within the framework of shared intentionality. Finally, in section 2.6, I will propose that this form of norm psychology is able to subsume some prototypical classes of moral judgments.

2.1. Defining moral judgments

Some dominant views in moral psychology (Dwyer, 1999; Hauser, 2006b; Mikhail, 2011; Mikhail, Sorrentino, & Spelke, 1998) and moral development (Nucci & Turiel, 1978; Smetana & Braeges, 1990; Turiel, 1983) argue that there is a well-delimited domain of moral cognition. These views claim to reveal the essential way in which moral thinking is psychologically unified. When they do so, they make an empirical bet. They characterize moral cognition in ways that are empirically testable—e.g., they hypothesize that there is one single moral domain (Turiel, 1983) instead of several distinct ones (Haidt, 2007). This assumption of unification is in some cases implicit, for researchers not always engage with definitional issues, although they draw general conclusions from them. For example, in the developmental psychology literature (Kohlberg, 1969; Piaget, 1932), emotions play little or no role in moral cognition because the target is typically conceived as a cognitive process closely linked to certain kinds of justifications and forms of reasoning.

Sometimes these assumptions are more explicit. Lengthy discussions about what moral judgments are supposed to be are somewhat common in the scientific literature (see, for instance, Haidt, 2007; Killen & Smetana, 2015; Moll, Zahn, de Oliveira-Souza, Krueger, & Grafman, 2005; Nucci, 2001). An influential view in moral development called the ‘social domain theory’ argues, for instance, that there is a distinct and identifiable moral domain (Nucci & Turiel, 1978; Smetana & Braeges, 1990; Turiel, 1983). Social domain theorists argue that their account “[...] allows us to define what is meant by the moral domain in a manner that transcends cultural and religious boundaries” (Nucci, 2001, p. xvii). They think that moral judgments are distinctive and unified since “[...] the moral domain forms a universal core set of values around issues of human welfare and justice” (Nucci, 2001, p. xx). For they argue that moral thinking is a single organized system, or domain, of social knowledge along with others such as a societal domain (concerned with conventions, social institutions, group norms, traditions, and cultural rituals) and a psychological domain (concerned with personal goals, autonomy, identity, and individual prerogatives) (Killen & Smetana, 2015).

The question at hand here is whether we are able to provide a definition of moral judgments that is able to cover all (or at least a significant majority of) intuitive cases of moral judgments, and do so in a way that increases the explanatory power of psychological theory. What is at stake here is whether moral judgments define a single natural kind in the psychological sciences or whether moral judgment divides into several distinct subclasses. Roughly speaking, supposed natural kind terms intend to be natural ways to classify phenomena that we discover through scientific investigation. Examples of natural kinds are chemical elements, some biological taxa, some psychological categories such as the distinction between short-term and long-term memory, and different types of cells, including some complex neuronal structures.

More precisely, natural kinds can be minimally characterized as classes of objects, processes, states of affairs, and the like, that: (a) have a common property (or set of properties) that are shared by all, and only all, the members of that natural kind and (b) allow interesting (or nontrivial) inductive inferences (see, for instance, Quine, 1969). Other approaches to natural kinds may have more stringent conditions. Some accounts may require from kinds to form a hierarchy (Ellis, 2001, 2002; Kuhn, 2000) or to be the subject matter of natural laws (Lowe, 2006). But what all these definitions of natural kinds have in common is at least some version of conditions (a) and (b).

I will set aside the metaphysical issue of the naturalness of kinds and I will focus instead on the explanatory and inferential roles of natural kinds in science. For scientific theories classify objects, events, and processes, in kinds that are the target of systematic generalizations and in ways that allow us to draw nontrivial inferences about the natural world. These scientific classificatory schemes are then modified when these empirical generalizations prove to be wrong. In many empirical sciences, such as physics or chemistry, natural kinds are typically intrinsic physical kinds, but in psychology and other cognitive sciences, kinds are rather classes of states and processes that are characterized by their functional role. For arguments of multiple realizability point out that it is possible for physically diverse cognitive systems to be in the same kind of mental state. As a result, a mental kind would not be essentially realized by a certain neurophysiological kind but rather by the same functional role that particular instances of such a kind of state play in cognition. In what follows, I will focus on views in moral psychology that share this functionalist flavor.

Interpreting the moral domain as a cognitive natural kind has been the focus of recent philosophical debate (Kelly & Stich, 2008; Nado, Kelly, & Stich, 2009; Sripada & Stich, 2007; Stich, 2016). These debates gravitate around Turiel and colleagues' social domain theory of moral development (Nucci & Turiel, 1978; Smetana & Braeges, 1990; Turiel, 1983). According to social domain theorists, moral judgments are characterized through a particular functional profile which is defined by the seriousness, authority independence, generalizability, and type of justification that is used to evaluate a normative transgression. They contrast moral judgments with conventional normative judgments, such as judgments about norms of etiquette or the rules of conventional games, for moral transgressions are considered to be more serious, authority independent, generalizable, and typically justified in terms of harm, justice, and rights. Moreover, they are natural kinds in the sense above because moral judgments are defined through (a) a set of common properties that are shared by all, and only all, moral judgments, which (b) classify psychological phenomena in ways that facilitate the generalizations and predictions of the social domain theory.

Other theories that take moral judgments to be a cognitive kind are nativist views of moral cognition. These views usually explain the functional role of mental kinds in terms of the modularity of the mind (Fodor, 1983). Nativist views of moral cognition deserve, and will receive, an independent discussion later in chapter 5 when I discuss the role of norm acquisition in moral cognition. But for the purpose of this chapter, it is important to point out that even those views crucially rely on Turiel and colleagues' characterization of moral judgments and the findings of the social domain theory. For example, Mikhail (2009) considers these findings to be the most persuasive case for an innate moral faculty. Thus, the key question I would like to address in this chapter is whether is theoretically possible to unify moral judgments as a natural kind via the functional definition that the social domain theory provides.

In the view I am proposing, in contrast, there may be no single class of moral judgments in psychology but different prototypical subclasses that may pick up important features of our psychology. These subclasses may turn out to be cognitive kinds on their own. The study of memory is a well-known example that fits this pattern. For the study of memory is actually the study of different types of systems which are classified according to different information processing features—e.g., episodic and semantic memory are subclasses of declarative memory, which in turn is a subclass of long-term memory. The issue at hand is not that there is a great variety of states and processes that we can call 'memory' or 'moral judgment'. Instead, the problem is that this variety of processes does not meet the minimal conditions (a) and (b) for being a natural kind. Moreover, I will argue that if moral judgments as a whole are a natural kind in the psychological sciences, the best way to reveal this is by fragmenting our explanatory target in ways that track real differences in mental processing and then asking ourselves what these processes have in common and how they are integrated into cognition. Even if moral judgments are not a disunified class in psychology, we can discover the nature of this class bottom-up by exploring what culturally varying but prototypical subclasses of them have in common.

Cross-cultural research would be crucial for this type of project. One could examine, for instance, if moral judgments conceived in a certain way are only a Western practice, a practice only seen among large-scale societies, or perhaps a highly widespread and a central part of our social lives. There might be a partial fit, i.e., some human groups could have practices that resemble in some respects, but not in others, the ones seen in other social groups. But we can only know this if we take these antecedent concepts of moral judgments and confront them with cross-cultural data. Information about this variability, in turn, tells us something about the overall landscape of moral thinking as an explanatory target for moral psychology. For example, moral foundation theorists (Graham et al., 2013; Haidt & Graham, 2007) have argued for a more complex taxonomy of moral judgments. According to this theory, evidence in cultural psychology reveals that some innate and universally available psychological systems are the foundations of intuitive moral judgments such as care/harm, fairness/cheating, loyalty/betrayal, authority/subversion, and sanctity/degradation. A similar research agenda has also been carried out by experimental philosophers who integrate cross-cultural data to delineate the moral domain through different techniques of data exploration such as cluster analysis (Fessler et al., 2015; Machery, 2012; Nado et al., 2009; Sinnott-Armstrong & Wheatley, 2012, 2013).

To sum up, the issue at stake here is whether moral judgments, conceived as mental states, are a unified natural kind in the psychological sciences. In the next section, I will explain what specific conditions have to be met in order to unify the class of moral judgments as a natural kind in psychology. This issue has methodological implications for this thesis. If we have reasons to think that a certain hypothesis along these lines is correct, then one could address the phenomenon directly based on the set of essential properties that ones' hypothesis postulates. If correct, there would be a lineage of processes and mental capacities that lead to the emergence of this cognitive kind. One would have to worry only about explaining the emergence of those essential functional properties that define the kind and one would have explained the whole class of moral judgments. There would be no reason to worry anymore about, say, cross-cultural diversity in the same way that one does not need to bother about it to determine whether subjects with anterograde amnesia have problems with their declarative memory. Otherwise, if we cannot assume that moral judgments are natural kinds, we have to proceed in a fragmentary way. We have to focus on cases that perhaps are culturally relative, topic-specific, or functionally idiosyncratic, and then wondering how these different targets are linked to each other, if they are at all.

2.2. What does it mean to claim that moral judgments are a unified cognitive kind?

I will focus in this chapter on psychology rather than the cultural and evolutionary history of these practices because my primary focus is on the distinctive psychological features (if there are any) of moral cognition. The evolutionary and cultural history of these diverse practices might well be an organizing feature of them—e.g., species are organized in lineages despite being highly diverse. But the question at hand here is whether these organizing features can define a unified target for the psychological sciences. In fact, I will show in the next chapters that although there is a great diversity of moral judgments, parts of them at least, are historically structured in the form of a common genealogy whose branches cluster around the socioecological demands and the basic cognitive and motivational mechanisms that distinguish the history of the human lineage. That said, although I will identify a unified class of normative thinking that can be explained through mechanisms of shared intentionality, this is not the same as saying that they are specific mechanism of moral thinking, for they also explain normative judgments that are prototypically nonmoral—e.g., the representation of norms that govern conventional games.

Kitcher's (2011) evolutionary explanation of normative cognition, which I will discuss in chapter 3, is an example of this type of project. He is not interested in unifying the genealogy of human moral thinking to define a natural kind in the psychological sciences. He thinks that the distinction between moral and nonmoral normative thinking is not psychologically or cognitively fundamental—his target is our distinctive capacity for normative guidance. For him, the distinction between moral, religious, and customary practices emerges from both the evolutionary and cultural history of normative cognition. I agree on this. For the kind of complex genealogy that he proposes does not intend to tell us anything specific about the psychological mechanisms of moral judgments. He is not arguing, for instance, that the history of moral thinking reveals that the distinction between moral judgments and nonmoral judgments is psychologically robust. He is not trying to unify moral thinking as a natural kind in psychology.

One can aim to unify moral judgments in different ways such that they are potentially useful for moral psychology. As Sinnott-Armstrong and Wheatley (2012, 2013) have shown, one can try to unify these mental states by their content, phenomenology, brain mechanisms, among other dimensions. Here, however, I will focus on a prominent version of the unification hypothesis, the social domain theory of moral development (Turiel, 1983), which try to unify moral judgments by their seriousness, authority independence, generalizability, and type of justification (see also Nucci & Turiel, 1978; Nucci et al., 1983; Turiel, 1998). I will focus on this theoretical approach because many researchers in this tradition have tended to assume that there is a correct definition of moral judgments, e.g., Turiel's well-known definition of moral judgments as "[…] prescriptive judgments of justice, rights, and welfare pertaining to how people ought to relate to each other" (1983, p. 3), and on the other hand, because this is perhaps the most influential theoretical framework where a definitional approach to moral judgments has been linked to empirical research aiming to identify a unified cognitive domain.[5]

More precisely, a unification hypothesis can be understood as the conjunction of three conditions, which must be satisfied when a candidate definition aims to unify moral judgments as a natural kind:

> **Exhaustiveness:** There is a property (or set of properties) that unifies all, or almost all, the members of the class of moral judgments.
> **Specificity:** The property (or set of properties) in question should only be shared by all, or almost all, the members of that class.
> **Nontriviality:** The property (or set of properties) should allow interesting psychological generalizations and predictions.

Conditions of exhaustiveness and specificity are demarcation criteria, i.e., general conditions of satisfaction that a candidate definition should meet in order to distinguish moral judgments from nonmoral ones. Nontriviality is the condition that guarantees that there is a scientifically informative connection between the candidate definition of moral judgment and the psychological theory.

I think that all attempts to satisfy these criteria fail in one way or another. The reason is that the class of moral judgments seems to be a very heterogeneous domain across many dimensions. Thus, this diversity does not allow interesting psychological

---

[5] I understand here the domain of a cognitive mechanism as the class of representations that it can take as input. However, cognitive domains could also be defined in terms of the task they perform. For example, I will argue in chapter 5 that the mechanisms responsible for norm acquisition are not domain specific in the sense that they are not specialized learning devices for the acquisition of norms.

generalizations. One definition of moral judgments may reveal some important features of our psychology. But these definitions only work for some particular subclasses of moral judgments—they do not unify moral judgments within a single kind. Other definitions may encompass all sorts of prototypical moral judgments in a nontrivial way. However, these definitions also include judgments that would not count as intuitively moral for a large sample of the population. If this is correct, one should explain moral thinking in a fragmentary way. One should show the connection between a particular cluster of judgments (e.g., judgments about a special class of obligations) and psychological mechanisms that are not exhaustive or specific to the whole moral domain.

To put it briefly, my main goal in the following sections is to argue that the project of unifying moral judgments, in the sense of satisfying the conditions of exhaustiveness, specificity, and nontriviality, is oversimple. But this does not mean that moral cognition lack of structure or that it is methodologically intractable. For as I will develop further in the rest of this chapter, the following strategy seem to be still defensible: (i) although the class of moral judgments is broad and heterogeneous, there might be a distinctive subset of that class (e.g., judgments that are inescapable and authority independent, as some moral philosophers propose) that is both theoretically important and a central case of lay judgments; (ii) then one can offer a psychological account of that core subset; and (iii) explain the respects (which may vary from case to case) in which other relatively more peripheral cases resemble core cases.

2.3. Social domain theory

There is probably no candidate definition of moral judgement with a better empirical support that Turiel and colleagues' social domain theory.

Over the past 40 years, the most important attempt to unify moral judgments in psychology has come from the social domain theory of moral development proposed by Turiel and colleagues (for reviews, see Helwig, Tisak, & Turiel, 1990; Killen, 1991; Nucci & Lee, 1993; Tisak, 1995; Turiel, 1998). This theory began as a departure from Kohlberg's (1969) approach to moral development. According to Kohlberg, children are only preconventionally moral (Colby & Kohlberg, 1987) because they are oriented in their moral reasoning by heteronomous considerations such as fear of external sanctions and obedience to authority commands. Instead, Turiel and others (Killen, 1991; Smetana, 1995; Smetana, Killen, & Turiel, 2000; Tisak, 1995) argue that early moral thinking is legitimately moral because young children do make moral judgments that go beyond heteronomous considerations of authority and obedience (see, for instance, Killen, 1991, p. 155; Turiel, 1983, p. 148).

To make their case, developmental psychologists in this tradition focus on the capacity to distinguish moral from conventional transgressions by putting forward a set of features that aims to draw a sharp distinction between the resulting judgments (Hollos, Leis, & Turiel, 1986; Nucci, 2001; Nucci & Turiel, 1978, 1993; Nucci et al., 1983; Smetana, 1993; Smetana & Braeges, 1990; Turiel, 1978, 1983, 1998).

In particular, according to Turiel's (1983, 1998) theory, moral judgments are considered:

(i) more serious,
(ii) authority independent,
(iii) generalizable to cultures in other times and places,
(iv) and justified in terms of harm, justice, or rights.

Following Kelly and colleagues (2007), I will call this cluster of properties the 'signature moral pattern', and I will consider these features to be intended to specify a reliable nomological cluster rather than necessary conditions of moral judgments. For it is likely that unknown mechanisms might be in place such that they cluster properties (i)-(iv) together, while constraining the co-presence of other sets of properties (Boyd, 1991, 1999; Millikan, 1999).

In contrast, according to Turiel, conventional judgments are considered:

(i') less serious,
(ii') authority dependent,
(iii') applicable only to some groups in particular times and places,
(iv') and not justified in terms of harm, justice, or rights.

Similar to the case of moral judgments, I will call this cluster of properties the 'signature conventional pattern', and I will consider the co-presence of these properties a reliable law-like effect of the underlying mechanisms behind conventional normative judgments.

Briefly, Turiel and colleagues' idea is that every time we see the transgression of a social norm, our judgments about these normative transgressions possess either the signature moral pattern or the conventional one. In the former case, these judgments usually (but not necessarily) are justified in terms of harm, justice, or rights. In other words, transgressions that are considered moral produce responses with the respective signature moral pattern (i)-(iii), and then reliably elicit the pattern of justification explained in (iv) (see figure 2.2).

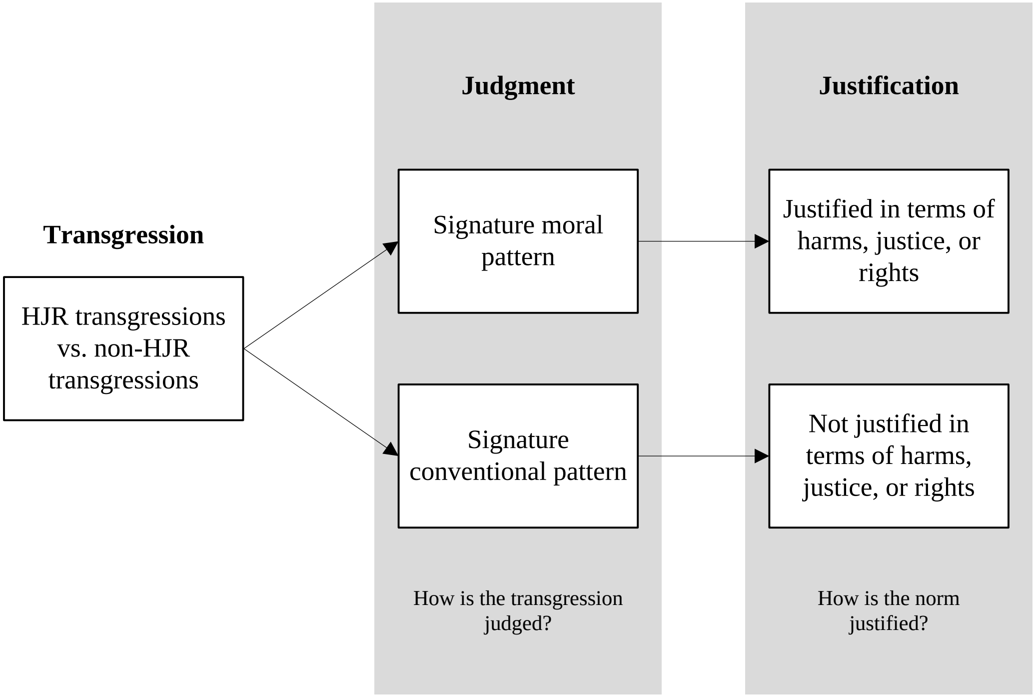

Figure 2.2. Diagram of the relation between transgressions, signature patterns, and types of justification in Turiel’s (1983) social domain theory. According to this theory, moral judgments are nomologically linked to some types of transgressions. Transgressions that involve harm, injustice, or rights violations reliably lead to judgments with the signature moral pattern. Transgressions that involve a breach of conventions without the involvement of harm, injustice, or rights violations typically lead to responses with the signature conventional pattern. Consequently, judgments of the former kind are typically justified in terms of harm, justice, or rights. In contrast, judgments of the latter class are not justified in those terms.

The most substantive claim of the theory is that we can consistently predict how the distinction will be drawn by manipulating certain features of the set of transgressions that we use as experimental stimuli. Transgressions that *prima facie* involve harm, injustice, or violation of rights are expected to be categorized as moral (Turiel, 1983). However, I will also focus on the more modest claim that the distinction between moral and conventional norms is robust with respect to the proposed patterns, i.e., that given a set of prototypically moral and conventional transgressions, they will be reliably clustered into the patterns described by the features (i)-(iv) and (i')-(iv').

Turiel and colleagues' approach is of special importance here because it relies on an impressive body of psychological data, and because it leads to nontrivial empirical generalizations and predictions about human moral thinking. This approach is supported by evidence derived from the 'moral/conventional task'. In this experimental paradigm, subjects have to differentiate different transgressions according to a questionnaire. The questionnaire is designed to determine whether the participants consider an action as a moral or as a conventional transgression following the distinctive patterns described in (i)-(iv) and (i')-(iv').

Empirical results have shown that people from different cultures (Hollos et al., 1986; Nisan, 1987; Nucci et al., 1983) and children as young as three years of age are able to distinguish moral from conventional judgments (Nucci, 1982; Nucci & Nucci, 1982; Nucci & Turiel, 1978; Smetana & Braeges, 1990; Turiel, 1983, 1998). More impressively, the pattern can be found in maltreated children (Smetana, Kelly, & Twentyman, 1984), children with autism and other cognitive disorders (Blair, 1996; Blair, Monson, & Frederickson, 2001; Nucci & Herman, 1982; Smetana et al., 1999), but not among psychopaths and children with psychopathic tendencies (Blair, 1995, 1997).

The moral/conventional task provides the most compelling evidence in favor of the idea that the distinction between moral and nonmoral domains is psychologically robust. Based on this experimental support, for instance, many evolutionary psychologists have argued for a nativist account of our moral capacities (Dwyer, 1999; Dwyer, Huebner, & Hauser, 2010; Hauser, 2006b; Mikhail, 2011; Mikhail et al., 1998) and that moral thinking relies on a distinctive class of moral computations (Cushman & Young, 2011; Hauser, 2006b; Mikhail, 2007, 2008, 2011). Though it certainly does not have to be, this approach is usually framed within the so-called 'linguistic analogy'—a view according to which moral cognition and language share several important features such as its universal development, its compulsory operation, and so on (for a more detailed discussion of this view, see chapter 5).

There are reasons, however, to think that Turiel-style moral judgments are still not able to unify the class of moral judgments in the desired way, i.e., meeting the conditions of exhaustiveness, specificity, and nontriviality. For the distinction between these forms of social evaluations and judgments is not always clear (Kelly & Stich, 2008; Kelly et al., 2007; Machery & Mallon, 2010; Nado et al., 2009; Nichols, 2004). Critics of this tradition have shown, for instance, that strong feelings of disgust trigger responses of the sort that are typical of moral transgressions (Nichols, 2004; Nichols &

Folds-Bennett, 2003). Spatial and temporal distance have also been shown to make moral transgressions be judged as conventional ones (Kelly et al., 2007). In addition, research to date on the moral-conventional distinction has relied almost exclusively on scenarios designed for young children (Kelly et al., 2007), even when the participants were adult psychopaths (Blair, 1995; Blair & Cipolotti, 2000; King, Blair, Mitchell, Dolan, & Burgess, 2006), or on large-scale societies, which share a number of key cultural features such as education and familiarity with formal legal systems (Henrich, Heine, & Norenzayan, 2010). Thus, the apparent unity of the moral domain may also be an artifact of the scenarios used or the type of societies surveyed. As we will soon see, there are sufficient reasons to think that this is indeed the case.

Turiel and colleagues' findings are empirically informative, but the problems are twofold if we interpret them as supporting the unification hypothesis. First, defining moral judgments through Turiel and colleagues' cluster of properties leads to failures of specificity. Transgressions that do not involve harm, injustice, or violation of rights can be judged (i) more serious, (ii) authority independent, and (iii) generalizable to cultures in other times and places. Empirical evidence, for example, strongly suggests that transgressions of this kind that generate strong disgust reactions have the distinctive signature moral pattern (Nichols, 2002; see also Schnall, Haidt, Clore, & Jordan, 2008; Wheatley & Haidt, 2005). These transgressions can show this pattern even when participants explicitly agree that the respective transgression is not moral—e.g., experimental evidence shows that disgust sensitivity can predict intuitive disapproval of gay people even in cases where participants do not explicitly judge homosexuality to be morally reprehensible (Inbar, Pizarro, Knobe, & Bloom, 2009b). Therefore, a transgression that may be classified as conventional, according to the experimenter's stimulus selection criteria, may have the signature moral pattern.

Of course, one could argue that the moral/conventional distinction is still useful because it tracks our implicit categorization practices. However, our implicit categorization of transgressions varies cross-culturally. One example comes from studies in the Turiel's social domain tradition. *Prima facie* moral and conventional transgressions, i.e., transgressions that involve harm, injustice, or violation of rights and transgressions that do not, can trigger the standard signature patterns among Kibbutz children and urban secular kids in Israel. Yet in traditional Arab villages, both moral and conventional judgments are taken to be (ii) authority independent and (iii) generalizable (Nisan, 1987). It is not only that they moralize what we take to be conventions, but rather that they do not moralize *à la* Turiel. The cluster that emerges from the implicit categorization practices in those communities does not match the predictions of the theory because what are taken to be prototypical conventional transgressions possess a mix of features from both the signature moral pattern and the signature conventional pattern. Therefore, moral judgments are classified into Turiel-style moral judgments and judgments that are half-way moral and half-way conventional, according to Turiel’s criteria.

Similarly, food, clothing, terms of address, sex roles, and ritual observances are usually considered moral by orthodox Hindu children (Brahmins and Dalits), and there is no normative belief viewed predominantly in conventional terms among Judeo-Christian American children under 10 (Shweder, Mahapatra, & Miller, 1987). In both of these cases, the distinction between moral and conventional domains is murky with respect to the proposed features (i)-(iv) and (i')-(iv'). For example, unlike the American adult sample, five- to seven-year-old American children tend to judge eating beef if you want to as a moral right. But for most Oriya Brahmans the taboo on eating beef is a (i) very serious, (ii) authority independent, and (iii) universal type of transgression—although the wrongness of the action is (iv') not justified in terms of harm, justice, or rights. So at least in some cultures, norms of religion and ritual are hybrids between moral and conventional norms. Children across different cultures classify the prototypical set of stimuli in different ways such that they get clustered in groups that sometimes resemble one of the patterns but frequently neither of them because what is expected to be classified as a conventional transgression, according to the social domain theory, possesses some or all the features of moral ones.

Overall, there may be a pattern of quite rigid normative judgments in early childhood that becomes more sensitive to the conventional aspects of one’s cultural environment. It is known that younger children are more likely than older children to believe that boys cannot play with dolls, that women cannot be doctors, and that people cannot eat with their fingers, even though adults typically consider all these transgressions (which do not involve harm, injustice, or violation of rights) a matter of social convention. In one child study, for instance, half of the subjects considered that social conventions of etiquette were (ii) authority independent and (iii) generalizable to other cultures (Carter & Patterson, 1982). This would suggest that the clustering pattern proposed by Turiel and colleagues would be flexible and culturally canalized, rather than a robust psychological trait emerging early in ontogeny. Indeed, one could argue that the trait could require maturation to emerge, but the more maturation it requires, the more likely it is (at least in a highly culturally embedded species like us) that the mature form will largely depend on cultural input. If this is correct, one could expect the pattern to be more labile and cross-culturally malleable, with different cultural groups carving up the normative domain in different ways that only partially overlap.

There might be identifiable socioeconomic factors that correlate with these deviations of Turiel's distinction. Turiel and colleagues' cluster of properties are usually found among WEIRD subjects (i.e., people from western, educated, industrialized, rich and democratic countries). But non-WEIRD subjects, such as low socioeconomic status participants from Brazil and the USA, normally judge conventional transgressions to be (ii) authority independent and (iii) generalizable. It has been also pointed out that participants from these populations consider a (i) serious moral transgression to eat the family dog, to clean one's toilet with the national flag or to masturbate with a dead chicken, even when these practices are kept in private (Haidt, Koller, & Dias, 1993). Hence, these studies suggest that what seems to be a robust cluster of properties can actually come apart in different ways. This is an important objection to the idea that Turiel and colleagues' cluster of properties defines a natural kind in psychology. Their characterization of moral judgments cannot be merely understood as the claim that a judgment is a moral judgment if and only if it possesses features (i)-(iv). If they define a natural kind, the signature moral pattern must form a nomological cluster, i.e., there must be a strong, lawlike tendency for the members of the cluster to co-occur. Features (i)-(iv) should not come apart. If anything, the moral/conventional syndrome is more diverse than previously assumed by social domain theorists.

Second, defining moral judgments through Turiel and colleagues' cluster of properties also leads to failures of exhaustiveness. So far I have argued that what is expected to be treated as a conventional transgression may be partially or completely treated as a moral transgression across some cultures. I have also suggested that deviations from this pattern seem to be correlated with cultural, socioeconomic, and developmental factors. Likewise, judgments about some prototypical moral transgressions may also lack some of the features of Turiel-style moral responses. People can think, for instance, that judgments prohibiting harmful actions that are not of the schoolyard variety are (ii') authority dependent and (iii') applicable only to some groups in particular times and places. Some studies, for instance, have shown that people judge to be OK to whip derelict sailors 300 years ago, but not OK to whip them now (Kelly et al., 2007). Similar results have been found when alleged confounds (Fraser, 2012) were controlled in subsequent studies (Quintelier & Fessler, 2015; Quintelier, Fessler, & De Smet, 2012).

Kelly et al. (2007) found that participants are more likely to condone harmful acts when they occur in distant times and places, or when they are not sanctioned by law or any other authority, than when it happens in the present time and in a familiar place, or when it is sanctioned by an authority or law. Fraser (2012) argues that in cases as the one depicted above, it is not the temporal differences or differences in authority dependency the ones that generate the deviation from Turiel's predictions, but rather a moral principle stating that harming others is permissible if needed as a means to a sufficiently valuable end. Subjects may have thought, for instance, that 300 years ago whipping "[…] was the only punishment fierce enough to deter sailors from drinking on duty and thus the only way to safeguard the lives that could depend on a watchful lookout's warning" (Fraser, 2012, p. 7). However, such a *post hoc* explanation is unlikely to account for this experimental results since removing temporal differences and keeping utilitarian considerations constants, yield the same result when the transgression is still sanctioned by law (Quintelier et al., 2012).

If Kelly and colleagues are right, then prototypical harmful actions would fail to trigger responses with the signature moral pattern as predicted by Turiel's theory. However, it is not just that harmful transgressions do not trigger the expected responses of the social domain theory. To give an example, if Fraser's *post hoc* explanation of the Sailor's scenario were on the right track, then there would be transgressions involving harm whose wrongness is (iv) justified by referring to the harm inflicted, but which are (iii') not generalizable and even perhaps (ii') authority dependent. For although this explanation aims to account for the variability of moral judgments, introducing a condition to control for this alternative explanation would not reveal that the subjects' judgments are authority independent. In such an explanation, it could still be the case that subjects would reason that physical harm is an appropriate punishment only in situations in which authority allows this kind of punishment.

Experimental studies show that, in Turiel's own terms, children tend to evaluate the refusal to help as morally blameworthy (Killen & Turiel, 1998; Sierksma, Thijs, & Verkuyten, 2014), but they display more negative evaluations in intragroup situations when compared to intergroup ones (Sierksma et al., 2014). This suggests that the assessment of the seriousness of a transgression is driven by factors such as ingroup loyalty, which are located outside the moral domain as delineated by Turiel and colleagues. Given this, the social cognitive domain theory would predict that children evaluate not helping others as morally blameworthy independently of the group context. Instead, these results seem to be largely compatible with Haidt's (2007) moral foundations theory—a theory that carves up moral judgments in a more pluralistic way, rather than trying to unify them across a certain set of psychological dimensions as explained in section 2.1.

Yet again, prototypical moral transgressions seem to be highly sensitive to cultural input rather than cross-culturally robust traits. One possible reason why previous studies have repeatedly failed to notice this diversity is that comparative studies have not focused on truly dissimilar social groups such as large- and small-scale societies. Another is that studies in the moral/conventional tradition, following the historical focus on children, have typically used simple binary measures, which tend to obscure what otherwise is a graded and multidimensional continuum.

In what is perhaps the most comprehensive study to date on moral judgments, Fessler et al. (2015) conducted a cross-cultural analysis of five small-scale societies: two egalitarian indigenous societies in South America with economies based on horticulture, hunting, and fishing, Tsimané (Bolivia) and Shuar (Ecuador); a semi-stratified clan-based indigenous group reliant on fishing and horticulture, Yasawa (Fiji); a clan-based rural group focused on rice agriculture, Karo Batak (Indonesia); and a clan-based horticulturalist group, Sursurunga (New Ireland, Papua New Guinea). The study also included two large-scale societies as a control for comparison: a sample from Storozhnitsa (Ukraine) and another from California (USA).

The adult subjects were presented with a range of transgressions involving harm, injustice, or violation of rights that included the following: a man stealing a stranger’s money; a man battering his wife without provocation; a man intentionally injuring a friend after been unintentionally injured by him; a man cheating a stranger in a financial transaction; a man spreading a false rumor about a rival; a man bribing a witness to blame an innocent person for an incident he initiated; and a man raping an unfamiliar woman (Fessler et al., 2015, p. 3, see also supplementary material). Then, Fessler and colleagues replaced dichotomous judgments of the acceptability of actions with judgments about the wrongness of the action on a five-point Likert scale ranging from ‘extremely bad’ to ‘extremely good’.

Similar to Kelly et al. (2007), researchers in this study asked the participants to judge those transgressions in cases in which an authority figure stated that the action was not wrong, or when the action occurred in a distant time or place. They found that participants in all these societies the transgression were systematically judged as (i') less serious when they happened a long time ago and far away, in another society. Unsurprisingly, all small-scale societies, with the exception of the Karo Batak in

Indonesia, were more likely to see a transgression involving harm, injustice, or violation of rights as (i') less serious when they were endorsed by an authority. In contrast, the two large-scale societies and the Karo Batak displayed non-significant trends in the direction of reduced severity.

In all these experiments, participants did not display all the properties of the signature response patterns (Kelly & Stich, 2008). The relevant types of transgressions also did not reliably trigger the expected responses. This is particularly important because, as it has been pointed out before, one feature that differentiates Turiel and colleagues' approach from the rest of the candidate hypotheses explored so far is that it leads to interesting empirical generalizations and predictions about the phenomenon in question. Yet transgressions that involve harm, justice, or rights can trigger the signature conventional pattern (Kelly et al., 2007; Quintelier & Fessler, 2015; Quintelier et al., 2012), and similarly, transgressions that do not involve harm, justice, or rights can trigger the signature moral pattern (Haidt et al., 1993; Nichols, 2002, 2004; Nisan, 1987).

Thus, an attempt to unify moral judgments through Turiel and colleagues' social domain theory of moral development seems to exhibit both failures of exhaustiveness and specificity. That is to say, what is expected to be treated as a conventional transgression can be partially or completely treated as a moral transgression across different cultures. Likewise, judgments about some transgressions involving harm, injustice, or violation of rights can have some of the features of Turiel's signature conventional pattern. The moral/conventional syndrome is diverse. Turiel and colleagues' cluster of properties shatters into different clusters that partially overlap with each other. If this is correct, there is no reason to understand the above account as an exhaustive and specific theory of moral judgments as opposed to a theory about a particular class of prototypical moral judgments, which seems to be correlated with particular cultural, socioeconomic, and developmental factors.

2.4. Fragmenting moral judgments

I have argued in the previous section that there are sufficient reasons to think that the social domain theory does not unify moral judgments as a psychological natural kind. This theory offers what may be the unification hypothesis with the best empirical credentials in the literature but it is certainly not the only possible candidate. One could construct other definitions and subsequently test whether they unify moral judgments as a natural kind. This is still a possible avenue of research but one that does not seem to be promising. Satisfying the conditions of exhaustiveness, specificity, and nontriviality seems too demanding. The most empirically supported psychological theory of moral judgments to date fails to meet these criteria. Of course, we could keep testing alternative unification hypotheses, but I would like to suggest a different methodology that allows us to explore different kinds of non-unificationist relationships between *prima facie* members of the class of moral judgments. More precisely, what I would suggest in the following sections is a divide-and-conquer strategy that aims to identify particular subclasses of prototypical moral judgments that can be clustered around the lineage of the psychological mechanisms for human norm psychology. In this section, I will explain what I mean by fragmenting those judgments.

The idea that moral judgments is a heterogeneous domain is a common view in contemporary philosophy (see, for instance, Flanagan, 1991; Nado et al., 2009; Parkinson et al., 2011; Scanlon, 1998; Sinnott-Armstrong, 2008; Sinnott-Armstrong & Wheatley, 2012, 2013; Stich, 2006; Taylor, 1978). For example, in the Kantian tradition, moral judgments are thought to be inescapable and authority independent if they are applicable to all the individuals regardless of their goals, desires, or preferences, and if its normative force is not derived from someone’s approval, endorsement, or capacity for coercion. In some interpretations, inescapability and authority independence are features of at least some important cluster of moral judgments, which are important for independent reasons, e.g., due to the role they play in practical deliberation. They single out a particular and central cluster of judgments, which are characterized by those properties.

I align myself with this view. Richard Joyce (2014), for instance, argues that moral judgments are inescapable and authority independent, but he does not argue that all moral judgment possess these features:

> Perhaps there is simply no fact of the matter as to whether moral rules have or lack this authoritative quality. Certainly people seem to generally imbue their moral prescriptions with this kind of strong authority, so maybe having a theory that provides this authority is a theoretical desideratum. But perhaps this authority is not an indispensable component of morality; maybe if we can make no sense of this authority and have to settle for a normative system lacking it, the system would still deserve the name ‘morality’. One way of diagnosing this situation

> would be to say that strictly speaking morality has this authoritative quality, but loosely speaking it need not. (p. 132)

Joyce seem to recognize here that not all moral judgments involve imperatives, or that not all moral imperatives are categorical imperatives, which would allow some moral judgment to be escapable or authority dependent (see also Joyce, 2006, p. 61).[6] I will follow his suggestion.

This understanding of moral judgments has been the focus of much discussion in recent evolutionary moral psychology (see, for instance, Fraser, 2010; Joyce, 2001, 2006; Mameli, 2013). One reason is methodological, for some researchers have recognized that inchoate notions of what moral judgments are hamper progress in evolutionary moral psychology. Inescapability and authority independence are then thought to help us to tackle this problem by singling out, at least, a particular class of well-defined judgments. Another reason is the philosophical implications that this view of moral judgments might have. For example, evolutionary debunkers in moral psychology have argued that the genealogy of moral judgments understood in this way undermines their epistemic status and ultimately erode their authoritative role in our practical deliberations (see Joyce, 2006). If one assumes that all and only moral judgments are inescapable and authority independent, then one could arrive at a robust form of error theory, i.e., the idea that all moral beliefs are systematically false (see Garner, 2007; Joyce, 2001; Mackie, 1977; Olson, 2010). It is not just that moral judgments in Joyce's narrow sense are false. Instead, the issue would be that moral thinking is systematically wrong across-the-board.[7]

---

[6] Joyce (2001) claims that the authority of morality is a problem that leads to error theory, while Joyce (2006) argues for a certain kind of moral nativism (according to which moral judgments did not evolve to track truths) in order to establish the epistemological conclusion that moral judgments lack justification. The passage quoted above is part of neither argument, but rather is a concession to opponents of Joyce (2001).

[7] The inescapability of moral judgments plays a central role in Joyce's (2001) debunking argument, where he thinks that no sense can be made of categorical reasons because they are entirely divorced from an agent's desires, goals, and the like. However, this feature of moral judgments seems to play a less prominent role in Joyce's (2006). For in this version of the argument, the debunking of moral judgments is thought to be a consequence of their being a product of an evolutionary process that is insensitive to moral truth. Inescapability only helps to explain why the capacity to make moral judgments promotes action and, therefore, why it is advantageous and selected for.

So, following Joyce's caution, I think that only some moral judgments are inescapable or authority independent (for a similar strategy see, Mameli, 2013). I agree that the motivational features that Joyce and others take to be diagnostic of moral demands can give us epistemic leverage to generate interesting psychological hypothesis about a special class of prototypical moral judgments. But inescapability and authority independence are neither exhaustive nor specific features of moral judgments. Giving money to charity, to give a toy example, may be seen as an act of kindness, which might also be understood as a morally good action. It seems plausible, then, to think that although one can judge these actions as morally good, they are not judgments about moral demands.[8]

It is not clear that the distinctive feature of moral judgments is their inescapability and authority independence, for not only moral judgments are inescapable and authority independent. These are features of other forms of norms, not just moral norms. A person cannot opt out of the disgustingness of certain norm violations. I might not care whether I am disgusting, but others can and will judge me to be disgusting. Similarly, disgust reactions exhibit little authority dependence (Haidt et al., 1993; Inbar et al., 2009b; Kelly, 2011; Nichols, 2004; Nichols & Folds-Bennett, 2003; Schnall, Haidt, Clore, & Jordan, 2008; Wheatley & Haidt, 2005).

There are also right and wrong ways to solve mathematical equations or to use screwdrivers. These nonmoral normative standards can also be judged as inescapable or authority independent as moral judgments. Given a certain goal, for instance, using a tool in a certain way may be the most suitable means to our ends. The use of screwdrivers can be then linked to normative standards of instrumental rationality. This means that although those normative standards are not inescapable because they are relative to a certain end, they are authority independent—their normative force is not derived from someone's approval, endorsement, or capacity for coercion.

---

[8] On this view, charity is a supererogatory act that is morally good but not strictly required. However, some views in moral philosophy argue that potentially good actions create reasons to carry out those actions. These views reject the existence of supererogatory acts on the basis that whatever is thought to be good, ought to be done (Feldman, 1986; Moore, 1903; New, 1974; Pybus, 1982). The key question here is not which philosophical view is correct, but rather whether it makes sense to think that thoughts like "Giving money to charity is morally good" are moral thoughts. I think the very philosophical debate shows that these thoughts are taken to be intuitively moral for a significant part of the population. Certainly, an important open question would be whether a normative system could count as a moral system if it does not contain normative demands anywhere. But this question goes beyond the debate at hand.

Moreover, the norms of rationality involved in mathematical proofs can be considered not only authority independent but also as inescapable as those of morality since they are quite robust standards with respect to our goals, desires, or preferences. If by an inescapable moral judgment one means a judgment that is applicable to all the individuals regardless of their goals, desires, or preferences, then one could say the same about rational judgments. Norms of rationality are considered inescapable because asking “Why should I be rational?” is asking for a reason. Thus, questioning them seems unintelligible. In fact, since practical rationality is often considered inescapable, the inescapability of moral judgments has been also traditionally linked to the inescapability of practical reasons (Kant, 1785/1998; Smith, 1994).

Inescapability and authority independence are features of some non-moral normative thoughts. Although important for philosophical reasons, not all moral judgments are judgments about moral demands in the sense of being demands that are judged by the agent as inescapable and authority independent. Inescapability and authority independence are features of some moral judgments, but not features that make judgments to be moral. These are a specific subclass of moral judgments that has attracted a lot of attention in the moral philosophy literature. They are prototypical rather than essential features of moral judgments in the sense that, if a judgment is moral, there is a good chance that such judgment is inescapable and authority independent. But these are diagnostic features of a prototypical class of moral judgments even though they do not make judgments intrinsically moral.

Finally, before going ahead, it is important to clarify why fragmentation is important for evolutionary explanations of moral thinking. Because if moral judgments form indeed a disunified psychological kind, then we should provide independent (or partially independent) genealogies explaining the different special features of each subclass. Moreover, and although beyond the scope of this dissertation, this kind of project could also help us in the long run to understand how prototypical subclasses of moral judgments are related to each other in terms of both their function and evolution. For example, if moral judgments are cognitively unified in a way that we have not discovered yet, we would expect to see an overlap in the trajectories of the different mental process that define the seemingly diverse cluster of moral judgments.

2.5. The psychology of norms

In the past few sections, I have argued that the class of moral judgments cannot be unified as natural kind in the psychological sciences. Moral judgments likely form a fragmented class of mental states. However, as I will argue in this section, an important methodological option available is to cluster prototypical moral judgments within a more encompassing class of normative thinking. In particular, I want to formulate in this section the basic proposal that I will explore in the rest of the thesis, i.e., the idea that human social norm psychology defines a cognitive kind that is able to subsume at least some important classes of prototypical moral judgments.

A theoretical possibility left open by the above discussion is that different well-defined classes of prototypical moral judgments can be clustered around a certain unified form of normative psychology (see Sinnott-Armstrong & Wheatley, 2012, 2013; Sripada & Stich, 2007). Following this line of thought, Sripada and Stich (2007) have suggested that there is a class of normative thinking that is a theoretically interesting target for psychology, even though there is no unified and cross-culturally robust moral domain (Turiel, 1983). This form of normative thinking is generated by our evolved norm psychology. This does not mean that normative cognition is entirely reducible to this psychology of norms, but rather that, unlike the moral domain, this type of psychology is a unified social-cognitive domain.

In a similar way, Tomasello and colleagues have shown that some normative judgments have a distinctive psychological profile that is very similar to the one proposed by Stich and colleagues (see Kelly & Stich, 2008; Sripada & Stich, 2007). This profile is explained through a family of well-defined psychological processes that fall under the umbrella of shared intentionality (see, for instance, Göckeritz, Schmidt, & Tomasello, 2014; Rakoczy, 2008; Rakoczy, Hamann, Warneken, & Tomasello, 2010; Rakoczy & Schmidt, 2013; Rakoczy, Wameken, & Tomasello, 2008; Schmidt, Rakoczy, & Tomasello, 2011a; Wyman, Rakoczy, & Tomasello, 2009a, 2009b). One of the central goals of the next few chapters is to show that shared intentionality and the kind normative thinking it engenders is a natural kind in the psychological sciences since they play a central explanatory role in psychological generalizations and in explaining a broad class of mental phenomena. As we will explore further in chapters 4 and 5, the framework of shared intentionality not only help us to explain the early expansion of our cooperative tendencies but also all the central aspects of the kind of norm-based cognition that I will outline in this section.

Much of the discussion in the following chapters focuses on explaining the importance of this class of normative thinking in human evolution. But the basic idea is relatively straightforward. There are reasons to assume that normative judgments are intentional states because they are judgments, i.e., they are mental states about something. Importantly, some of these judgments are private, while others are shared. That is, there are normative mental states that only play a salient role in individual cognition. Normative judgments about what is a prudential course of action are not necessarily judgments about what other agents should do because the agent may simply not care about the bad decision that other agents take—in a highly competitive social environment, instrumentally irrational agents that one has to outcompete might well be a good thing. However, in other cases, we may care about what others actually do and the social expectations they have, so we face the selective pressure to share our normative thoughts with them. It is true that in some cases the sharing of those thoughts fails. An agent may share with others his/her normative thoughts about what they should do in a given situation, but they may not join the agent's intentional state. But despite these difficulties, when the sharing of these expectations is successful, it enables more complex and efficient social coordination. It is in this sense that this kind of normative thinking is norm-based thinking. For it means that we have managed to share at some degree the social expectations and the attitudes we have toward those who violate those expectations. I think this is the reason why shared normative thinking is a central kind of normative cognition. Only the normative thoughts that we can (or we are willing to) share with others play a role in social cognition. It is only when we share these mental states that we can reliably generate the collective benefits derived from coordinating our social expectations. Thus, one of the goals of my lineage explanation is to show that this is kind of thinking has played a central role in the evolution of our lineage and convince the reader that shared normative thinking is an important part of who we are.[9]

More precisely, the view I will defend in this thesis is that an important subset of normative thoughts that are generated by our norm psychology are joint intentional

---

[9] It might be important to differentiate two different senses in which the term 'social' could be understood here. One refers to norms that are about public, social interactions. Another is refers to normative judgments that are made public, rather than private. In the sense introduced in section 1.3, social norms require both features, i.e., they are not only norms about social interactions, but also norms that are expressed and endorsed in public contexts. However, both senses of 'social' do not necessarily go hand in hand. One could make public one's self-directed norms as much as one could keep private one's normative views about how everybody should act.

states that can be characterized by a gradient of generalizability, intrinsic motivation, and punitive attitudes. Generalizability refers to the scope of the judgment and its context-sensitivity—normative beliefs can regulate the behavior of only some individuals in specific situations or roles. So, normative beliefs can be characterized by a gradient of generalization or abstraction depending on how tightly they are conceived to be linked to specific individuals or situations. Intrinsic motivation refers to the subjective pressure a person feels to comply with a certain normative belief. For sometimes normative beliefs are treated as ultimate ends, rather than as a means to other ends.[10] Similarly, punitive attitudes refer to the motivational force people feel to police, punish, or correct others, including themselves, when they think that they have violated a norm. These punitive attitudes are sometimes salient and explicit, while other times are less evident. For example, in some cases, the violation of a norm engenders actual physical punishment. In other cases, the perceived violation of a norm engenders punitive attitudes like anger and blame (or shame and guilt) toward the transgressor. And yet in other cases, it just drives corrective behaviors that target the transgressor without any harshness or evident signals of reprisal—e.g., when an adult correct a child for the violation a certain norm of etiquette.

This does not aim to account for all the range of normative beliefs, since at least some norms such as norms of instrumental rationality (e.g., about the appropriate use of corkscrews and screwdrivers) do not necessarily involve intrinsic motivation or punitive attitudes. So-called ‘prudential norms’ in philosophical theories of individual rational action are not social norms in the required sense. Normative judgments as a whole could indeed share a similar fate as moral judgments, i.e., they could be a disunified category in terms of their psychology. But normative judgments can be conceived as a supercategory that contains moral judgments as a subcategory and the class of normative judgments generated by human social norm psychology can contain as a proper subset at least some subclasses of moral judgments (see figure 2.1.). Arguably, then, a large portion of normative mental states, including some prototypical forms of moral judgments, can be understood as a class of joint intentional states.

---

[10] This does not mean that these normative beliefs have overriding power. A normative belief can be treated as an ultimate end, but its motivational force can be overridden by the motivational force of other mental states. What is essential for a mental state to be an ultimate end is to motivate the agent to bring about a certain state of affairs irrespective of the way they promote other agent’s goals. In contrast, instrumental ends are those mental states that determine the means through which an agent will bring about that state of affairs.

Joint intentional states are mental states that are partially cognitive and partially motivational. These mental states can vary across a cognitive gradient of generalizability and abstraction as much as they can across a motivational one. From a cognitive point of view, representing activities in a joint form requires being able to represent plural subjects of action (e.g., “We want to hunt” or “We are attending to that antelope”) and task roles or activities that are not necessarily linked to particular individuals (e.g., agent-independent representations of the different roles that our particular hunting technique requires). This creates a gradient of generalizability and abstraction depending on the scope of the plural subject and the degree of abstraction with which these activities are represented. This cognitive component makes at least some shared intentional states to be belief-like. For example, even entertaining a joint intention like “We want to hunt” can misrepresent the social situation, e.g., when there is no actual partner who wants to hunt with me.

From a motivational point of view, shared intentional states are intrinsically motivational, e.g., by definition, entertaining a joint intention of the form “We want to hunt” implies a motivation to hunt with others that is not purely instrumental. That is, an agent joins the activity because he/she finds its collective nature intrinsically rewarding, rather than because the agent only thinks the activity is instrumentally beneficial for him/her, or even instrumentally beneficial for everyone. The peculiar normative dimension of some of these mental states is also closely related to the motivational component of shared intentional states because this dimension can be spelled out in terms of the punitive attitudes that these states are associated with. Not all shared intentional states are shared normative mental states. But the divide between the two types of states can be made clear via the punitive attitudes that are specifically linked to the kind of shared normative thoughts that I want to single out in the thesis. To be clear, not all shared intentional states entail punitive attitudes. If we are hunting together and I let you alone in the task, this will likely trigger punitive attitudes, for collaborative hunting requires commitment. But not all shared intentional states require this degree of commitment to collective activities. For example, if we are jointly attending to the reflection of the moon on the river and I lose the interest in the event, this will likely not trigger any form of reprisal.

Shared intentionality is a very distinctive package of cognitive and motivational mechanisms that coevolved as part of the psychological infrastructure that supported the expansion of the cooperative capacities of our lineage. It is a package built of separate elements, but, as I will argue in the rest of this dissertation, is a package of cognitive and motivational elements that share a common lineage. Functional integration of an existing capacity leads to coevolution and codevelopment—cognitive subsystems that fire together wire together. Showing that this package of cognitive mechanisms became highly integrated and entrenched in human development, as well as providing an evolutionary explanation for this, is one of the central goals of this thesis. That is to say, I agree that the component of the shared intentional infrastructure coevolved, although this is different from saying that normative thinking coevolved with shared intentionality since normative thinking is not a cognitive subsystem of the shared intentional infrastructure.

Conceiving normative beliefs as shared intentional states has also a desirable theoretical consequence. When we single out a class of normative beliefs as a subclass of shared intentional states, we draw a distinction in mental processing across a number of well-defined functional dimensions. In the above framework, normative thinking is characterized by a cognitive gradient of generalization and abstraction that captures what Tomasello and colleagues' call 'bird's-eye view' of the social interaction. This bird's-eye view is a distinctive cognitive process involved in shared intentionality—e.g., in role-reversal tasks, young children seem to integrate different roles into a single bird's-eye representational format that is not present in great apes (see, for instance, Fletcher, Warneken, & Tomasello, 2012). A social interaction is represented from a bird's-eye view when the agent representing the interaction steps away from his/her own egocentric perspective of the interaction to take into consideration the point of view of one's partners (e.g., their roles in the interaction, visual perspectives, or any other relevant perspectival information). The capacity to represent objects and events with a bird's-eye view perspective coevolved with the rest of our shared intentional infrastructure. This perspective is essentially a case of functional abstraction and role/occupant distinction. Not all shared intentional states are like this since basic forms of shared intentional states such as joint attention might be carried out through recursive mindreading capacities, which can support simple joint activities.

This cognitive process leads to some shared intentional states to be called 'collective intentional states', i.e., mental common ground that we share with larger social groups, rather than specific individuals with whom we engage in simpler forms of shared intentionality (for a more detail exposition, see Tomasello, 2015). On the view I am outlining, the distinction between joint intentional states and full-blown collective states is not only a matter of degree but also a matter of cognitive mechanisms. As previously mentioned, joint intentional states can vary along a gradient of generalizability. Over that gradient, collective intentional states are characterized as joint intentional states that we engage with a relatively large number of agents in a rather abstract way—i.e., abstracting away the details of particular agents such that it facilitates the interaction with unspecific teammates, our local tribe, and the like (for a similar distinction, see Mead, 2015). Other cases of shared intentionality are more discrete in the sense that they are joint intentional states that we engage with a relatively small number of identifiable agents—e.g., episodes of joint attention that we share with particular others. This may lead to constraints in the kind of representational machinery that we deploy in order to entertain these thoughts. When collective activities are carried out by multiple agents performing different tasks, building common ground on the basis of recursive mindreading capacities becomes cognitively implausible. I think norm-based cognition is a form of collective intentionality in that sense (see also Tomasello & Vaish, 2013; Vaish & Tomasello, 2013). They require what I will call a 'we-mode' of representation. That is, they require from the subject to conceive the norm from an increasingly agent-independent point of view that relies less on the specifics of an agent or group of agents (for a more philosophical discussion, see Nagel, 1986).

Similarly, it has also been argued that shared intentional states are intrinsically motivational (Tomasello, Carpenter, Call, Behne, & Moll, 2005). In primates, for instance, gaze following and attention-directing gestures involve individual, parallel attention to an object, rather than real joint attention. For joint attention requires knowing together that others individuals are sharing attention as well as a motivation to share attention and interest with others with no other instrumental goal in mind (see Carpenter & Call, 2013). Of course, an agent can know that other individuals are sharing attention but not being motivated to share attention and interest with others. But this type of attentional states would be parallel rather than truly joint. As they are understood here, real shared intentionality implies a motivation to share the intentional state in question. Moreover, it also implies sharing the motivational component of that mental state when the state has one, e.g., it requires from the agent not only co-represent the content of another agent's belief but also share his/her propositional attitude.

In contrast, primates do not share attention and interest with others in situations of joint attentional engagement (Bard & Vauclair, 1984; Tomasello & Carpenter, 2005; Tomonaga et al., 2004) or declarative gestures (Gómez, Sarriá, & Tamarit, 1993; Tomasello & Carpenter, 2005; Tomonaga et al., 2004). Unlike any other primate, humans have an intrinsic motivation to share psychological states without requiring any external motivation, and this intrinsic motivation predisposes them to engage in a collective activity just for the sake of it (Call, 2009; Tomasello et al., 2005). We are motivated to share these states as ultimate ends, i.e., ends that we pursue at least in part for their own sake. Even young children conceive collaborative activities as ends in themselves, rather than mere means to a personal reward (see, for instance, Gräfenhain, Behne, Carpenter, & Tomasello, 2009; Warneken, Chen, & Tomasello, 2006; Warneken & Tomasello, 2007).

All shared intentional states entail by its operational definition an intrinsic motivation. Only some shared intentional states are real collective intentional states in the sense of involving, at least to some degree, an agent-independent representation of the social interaction. Nevertheless, what differentiates this class of normative beliefs from other shared intentional states is that they distinctively engender punitive attitudes. Punitive attitudes are not essential for shared intentional states, but they are often part of the package deal of cooperation as a form of partner control. Think, for instance, about the role it would have played in early human collaborative foraging and how collaborative foraging is thought to be linked to morality (Boehm, 2012; Mameli, 2013; Tomasello & Vaish, 2013).

2.6. From norms to morals

In the previous section, I have proposed that the cognitive and motivational infrastructure of shared intentionality help us to define a psychology of norms that is cognitively unified in the sense specified in section 2.2. That is, this psychology generates shared intentional normative thoughts that define a unified cognitive kind of normative judgments. This idea follows the proposal of Sripada and Stich’s (2007) framework for the psychology of norms, which I roughly map onto the psychological mechanisms of shared intentionality. The reasons why I think shared intentional normative thoughts define a unified psychological kind will become clearer in the following chapters since they are related to the evolutionary lineage of shared intentionality and its normative dimension. In particular, I will argue in chapter 4 that this kind of normative judgments played a central role in the expansion of our cooperative lifestyle—more precisely, in the shift from a form of ape-like individualistic form of foraging to a more complex form of collective foraging. As I will try to argue, this evolutionary lineage accounts for the functional integration of the different cognitive and motivational aspects of the social norm psychology that I have outlined in the previous section.

The cautionary tale of this chapter has been that moral judgments, understood as mental states, can be a heterogeneous target for psychology. But another, more optimistic message is that there may exist a class of mental states and processes in the vicinity of that concept that can help us to manage that complexity, for we can link different prototypical types of moral judgments to a well-defined class of mental states that are normative, i.e., the class of shared intentional normative thoughts. For example, we still can ask whether in a certain small-scale society of hunter-gatherers transgressions about the distribution of resources possess the Turiel’s signature moral pattern. Or one could wonder if in those societies there is also a distinctive class of judgments that is defined by their being conceived as inescapable and authority independent (Joyce, 2006; Mackie, 1977; Mameli, 2013). These judgments could then be linked to the psychological machinery underlying shared intentionality and its distinctive normative dimension, which could help us to explain different aspects of moral judgments in virtue of their being a special class of shared intentional states.

As a consequence, prototypical forms of moral thinking could be linked to the evolutionary lineage of this ontogenetically robust psychological capacity, while the distinction between these different types of moral judgment could be explained, for instance, as a matter of our particular sociocultural history. Judgments that possess the Turiel’s signature moral pattern or which are judged to be inescapable or authority independent can be considered prototypically moral and, as I will argue in chapter 5, they can be connected to the lineage of our shared intentional capacities. Whether all types of moral judgment can be covered under the umbrella of processes of shared intentionality is an open question, although I remain skeptical. So, one hypothesis I will explore in the following chapters is that at least some prototypical cases of moral judgments, i.e., judgments with Turiel’s signature moral pattern or judgments about moral demands as defined by some philosophers (Joyce, 2006; Mackie, 1977; Mameli, 2013), stem from the evolutionary lineage of our shared intentional capacities as a result of the process of norm acquisition. I will try to show how this core capacity for normative cognition is able to ramify into these different ways to carve up normative judgments and why I think that, unlike moral psychology, the former is an ontogenetically robust psychological capacity (see figure 2.3).

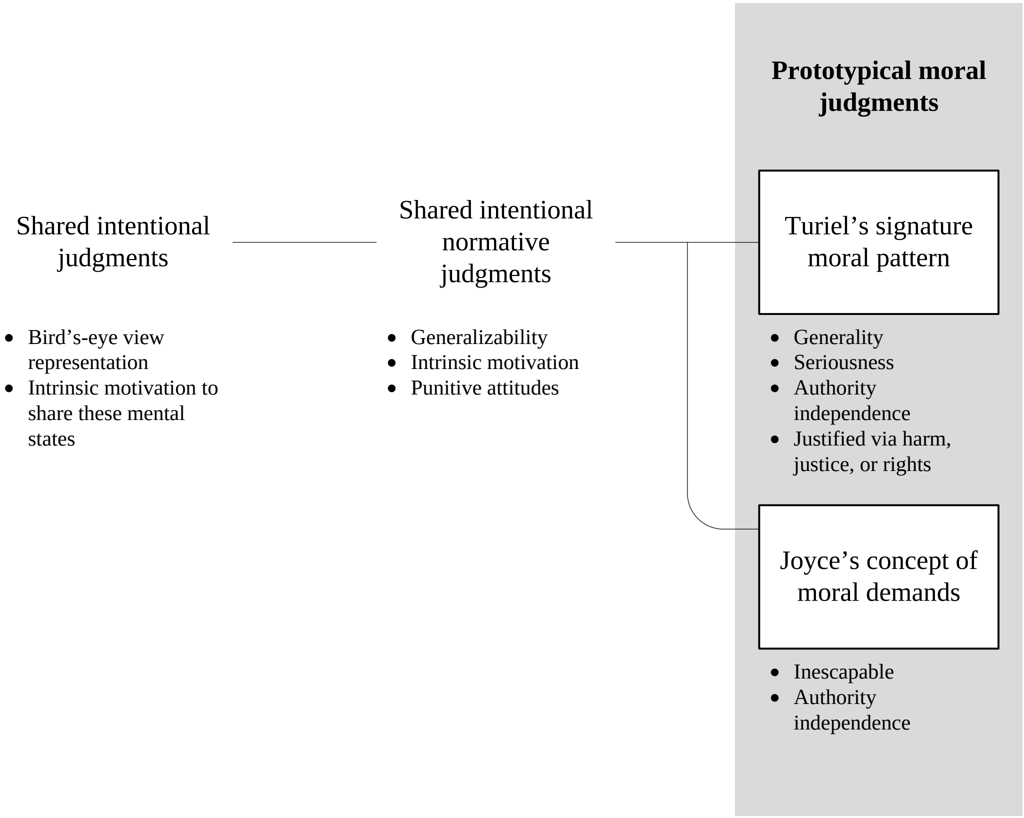

Figure 2.3. Diagram of the relation between shared intentionality, social norms, and prototypical moral judgments. The class of moral judgments fractionates into different clusters, each of which might be a culturally relative class of prototypical moral judgments. The lineage of shared intentionality leads to shared intentional normative judgments, which in turns leads to judgments with Turiel's signature moral pattern and Joyce's judgments about moral demands. Some ancestral features are retained in the lineage, while others are modified or functionally integrated downstream in the trajectory. The purpose of the lineage explanation I will provide in the following chapters is to explain the evolutionary trajectory of these features.

If it turns out to be true that there is no unified class of moral judgments in psychology, perhaps the partial overlap between different prototypical types of moral judgments is partially biologically channelized with different prototypical forms of moral judgments diverging from each other through culturally driven processes. This could help to explain, for instance, why although both Turiel's and Joyce's types of moral judgments are authority independent, these prototypes are dissimilar in other respects. For these prototypical moral judgments are partially biologically entrenched and partially culturally driven psychological phenomena. But while the distinction between different prototypical types of moral judgments, as well as the difference between moral judgments and other types of judgments, may be a matter of cultural history, the type of normative thinking that emerge from our capacities for shared intentionality are much more cross-culturally robust and entrenched in human psychological development. There is a reason to expect this since, as we will see in chapter 4, shared intentionality in general and shared intentional normative thinking in particular played a crucial role in human social evolution (see also Tomasello, 2009b, 2014; Tomasello et al., 2005; Tomasello & Vaish, 2013).

One interesting suggestion would be then that although there are different classes of prototypical moral judgments that share, at best, a certain family resemblance, these judgments can be closely linked to our shared intentional normative psychology. This type of psychology leads to a well-defined class of social normative judgments that splits into different branches of prototypical moral judgments. For example, following Mameli (2013), I will argue in chapter 5 that the tuning of the motivational gradient of punitive attitudes of shared intentional normative judgments during infancy explains how we acquire Joyce's concept of moral demands, i.e., how we develop the capacity to entertain judgments about demands that are considered inescapable and authority independent. A similar explanation via the cultural tuning of our affective responses to transgressions involving harm, injustice, or rights violations will explain the Turiel's signature moral pattern—an idea put forward by Nichols (2002) and others.

In the following chapters, I will try to explain how these different classes moral judgments are clustered around the psychological gradients that define shared intentional states—i.e., how certain shared normative mental states come to have the Turiel's signature moral pattern or how they become able to represent inescapable and authority dependent demands. A large part of my lineage explanation in chapters 3 and 4 will deal with the biological evolution of our shared intentional capacities and the genetic assimilation and biological entrenchment of shared intentional normative thinking. But the ramification of these judgments into prototypical moral judgments is culturally relative and, therefore, acquired. So, I will come back to this issue in chapter 5 when discussing the developmental trajectory of shared intentionality and its distinctive normative dimension. I will argue there that emotional and affective dispositions not only play a key constitutive role in normative thinking but also in explaining the special features that define the Turiel's signature moral pattern and Joyce's class of prototypical moral demands.

2.7. Conclusion

In this chapter, I have tried to argue that moral judgments define a quite heterogeneous class of mental states. I argued that the best theory that we have to assume that moral judgments are a unified cognitive kind, i.e., Turiel's social domain theory of moral development (Nucci & Turiel, 1978; Nucci et al., 1983; Turiel, 1983, 1998), fails in its attempt. There seems to be no property (or set of properties) that unifies moral thinking as an explanatory target for psychology. Instead, we should understand particular accounts of moral judgments as targeting particular classes of judgments that are prototypically moral. For example, judgments that are inescapable and authority independent pick up an important class of normative judgments (Joyce, 2001; Mackie, 1977; Mameli, 2013), although not all moral judgments are judgments about moral demands in the sense of being inescapable and authority independent (Joyce, 2006, 2014). Similarly, not only moral demands are inescapable and authority independent, e.g., what counts as the correct solution for a mathematical equation can also be judged as inescapable or authority independent as prototypical moral demands. Judgments that evoke Turiel's signature moral pattern or judgments about inescapability and authority-independent demands seem to be only prototypically moral. However, this does not entail the collapse of the psychological investigation of moral judgments. We can still explain in what sense a certain class of normative judgments is prototypically moral as well as why and how they share their family resemblance by linking those judgments (which may be culturally relative) to the lineage of more robust and developmentally entrenched psychological mechanisms.

Chapter 3. On social tolerance and the evolution of human normative thinking[11]

This thesis aims to provide a lineage explanation of a central form of normative thinking, i.e., an explanation that specifies a sequence of changes that takes us from agents with a certain baseline capacity for social cognition to agents with social normative thinking (see Calcott, 2009). As I have argued in the previous chapter, the class of moral judgments fractionates into different clusters, each of which might be a culturally relative class of prototypical moral judgments. However, as I have proposed at the end of that chapter, the lineage explanation of shared intentionality can help us to define an important form of norm-based cognition, which in turn leads to judgments with Turiel's signature moral pattern and Joyce's judgments about moral demands. On this view, the share intentional normative judgments that this psychology generates give rise to these subclasses of prototypical moral judgments by retaining some ancestral features of this kind of judgments, while modifying or integrating others.

I frame this lineage explanation within the hominin lineage. Providing an explanation of this kind means that I take the baseline of social-cognitive capacities to be the one of the common ancestor of chimpanzees (*P. troglodytes*), bonobos (*P. paniscus*), and humans (*H. sapiens*), for chimpanzees and bonobos are our closest living relatives. I will focus in particular on the key differences between that baseline and human normative thought. In this context, the aim of this chapter is twofold. I not only want to provide such a baseline but also to locate the emergence of the high-level cognitive processes that distinguish human normative thinking within the evolutionary trajectory that led to human social organization. The lineage here is not the lineage of organisms but of mechanisms. More specifically, since I have identified in chapter 2 a unified form of human norm psychology with our distinctive capacity for shared intentionality, I will propose at the end of this chapter that the lineage leading to the emergence of our capacity for normative guidance is the same as the one that leads to shared intentionality. The central goal of chapter 4 will be to flesh out this lineage explanation from the hominin baseline of social cognition that I defend in this chapter.

Reconstructions of the last common ancestor of chimpanzees, bonobos, and humans are in general important to understand human origins (Lovejoy, 1981; McGrew,

---

[11] This majority of this chapter is based on worked published in Gonzalez-Cabrera (forthcoming) and has been modified to meet university guidelines.

2010; Stanford & Allen, 1991; Tooby & DeVore, 1987). These discussions usually portray the *Pan*/*Homo* last common ancestor (*Pan*/*Homo* LCA hereafter) as a chimpanzee-like hominid (see, for instance, Dart, 1953; Lee & DeVore, 1969; Wrangham & Peterson, 1996; for a historical reconstruction of this debate, see also Pickering, 2013). This has long been the prevailing view in the philosophical and biological literature, and normative cognition is no exception. Recent accounts of the evolution of the human capacity for normative guidance such as Kitcher (2011) rely on this approach. I will argue that since the demonic male view and evolutionary models of normative thinking based on it no longer stand up, we need an alternative explanation of this capacity that relies on a different view of human origins.

This view has been challenged in different ways. According to some theories, the bonobo is in some important respects a suitable model of early hominins (de Waal, 1995, 2001; de Waal & Lanting, 1997; Zihlman, 1984; Zihlman & Bolter, 2015; Zihlman, Cronin, Cramer, & Sarich, 1978). Other models have focused on more distantly related great apes such as gorillas (Geary, Bailey, & Oxford, 2011; Geary & Flinn, 2001) and orangutans (Crompton, Vereecke, & Thorpe, 2008; Grehan & Schwartz, 2009; Schwartz, 1987; Schwartz, 2004), while yet another group, thinking in terms of convergent evolution rather than shared ancestry, have looked beyond the apes—e.g., through comparisons of human ancestors with savanna baboons (DeVore & Washburn, 1963; Jolly, 1970, 2001; Skybreak, 1984), capuchin monkeys (Fernandes, 1991; Perry, 1997; Perry et al., 2003), or even wolves and other social carnivores (King, 1976). Regardless of whether they are based on convergence or shared recent ancestry, what all these models have in common is the idea that reconstructions of human evolution can take advantage of much broader phylogenetic comparisons.

I will argue in this chapter that we are more justified in using an alternative model of early hominins, and perhaps even the *Pan*/*Homo* LCA, than we are in believing that early ancestors were chimpanzee-like. According to this model, early hominins were much more socially tolerant and less aggressive than usually assumed. I ground this claim in both the comparative evidence and the paleoanthropological record. As a result, I will argue that this model does not fit well with views such as the demonic male view (Wrangham & Peterson, 1996) or the killer ape hypothesis (Dart, 1953).[12]

---

[12] Something similar can be said about the man-the-hunter hypothesis (Lee & DeVore, 1969). For hunting and aggression are usually considered to be a package deal. However, the model of the *Pan*/*Homo* LCA I will propose in this paper does not rule out the idea that hunting played an important role in the evolution of normative guidance. For these reasons, I sympathize with evolutionary views such as those Pickering

More important, I will show here that such a model has important consequences for philosophical debates about the origin of our capacity for normative guidance (Kitcher, 1998, 2006, 2011). For they give us reasons to think that the emergence of top-down cognitive processes of normative guidance goes hand in hand with the emergence of shared intentional capacities (this will expand on this basic proposal in the next chapter through the connection between hominin hunting and social tolerance; see also Pickering, 2013).

The chapter is organized as follows. In section 3.1, I will explain the philosophical motivations behind genealogical approaches to human social cognition. In section 3.2, I will discuss the perils of reconstructing the social behavior of our early hominin ancestors. In section 3.3, I shall explain the specific model of these ancestors I want to propose. In sections 3.4 and 2.5, I will provide evidence in favor of this model. In section 3.6, I will discuss whether these similarities are homologies, or whether they have evolved independently. Finally, in section 3.7, I will draw out the philosophical moral of this discussion for our understanding of our capacity for normative guidance.

3.1. The demonic male hypothesis

Philosophers have argued that that the deep history of why we became moral agents is relevant to normative philosophy. For one way to understand human nature is to understand its genealogy. I take genealogical accounts to be close relatives of lineage explanations. One primary example is the role that various origin stories of morality have played in moral philosophy (Hobbes, 1668/1994; Nietzsche, 1887/1967; Rousseau, 1755/1992; see also Korsgaard, 2010). Another example is the way in which the evolutionary genealogy of our moral faculties have become a way to vindicate (Kitcher, 2006, 2011) or debunk morality (Joyce, 2006; Ruse, 1998; Street, 2006). As a result, genealogical projects in philosophy become highly sensitive to different assumptions about our hominin baseline. Depending on these assumptions, for instance, some evolutionary narratives will become more vindicatory than others (see Hobbes, 1668/1994; Hume, 1740/1978; Locke, 1689/1988; Rousseau, 1755/1992, 1762/1987; Sterelny, 2012b).

(2013) who reject the demonic male hypothesis while still defending the central role of hunting (for a complete rejection of both, see Sussman, 1999).

According to the demonic male hypothesis (Wrangham & Peterson, 1996) and the killer ape hypothesis (Dart, 1953), we evolved from a chimpanzee-like hominin whose basic social nature was characterized by hostile intergroup relations. Human and chimpanzee males share a capacity for violence because our common ancestor also possessed a genetic predisposition for such capacity (see Crofoot & Wrangham, 2010; Wrangham & Glowacki, 2012). Intergroup conflict, for instance, plays a central role in some prominent accounts of the evolution of human cooperation (Bowles, 2008, 2009; Bowles & Gintis, 2011). On this view, this predisposition not only is an important aspect of human psychology but also substantially contributed to the evolution of our lineage by constraining the path and setting the pace of human social-cognitive adaptations.

Human ancestors were distinctively aggressive, and this trait was preserved thanks to the role of war and interpersonal aggression in the evolution of our lineage. On this picture, emotional reactivity led to social groups controlled by aggressive alpha males, but the increased cognitive demands of cooperative hunting and tool-making helped us to control our aggressive tendencies. Put another way, from a cognitive point of view, human evolution can be seen as the story of the emergence of different forms of top-down control over our more disruptive and less reliable emotional nature.[13] Humans are predisposed to violence and dominance, but we overcame these limitations through the steady increase of our intellectual capabilities (see also Pinker, 2011).

This picture has implications for our view of the relative roles of rational reflection and emotion in our normative lives. For emotional reactivity would be a challenge to overcome rather than support normative thinking. But this picture radically changes, however, if a different ape species such as the bonobo, not the chimpanzee, turns out to be a comparatively better model of the social behavior of our last common ancestor. If the social world of our forebears was more cooperative and peaceful than depicted by the chimpanzee referential model, neither the killer ape hypothesis nor the demonic male view of our social nature would be completely right. I will argue in this chapter, for instance, that, to a large extent, emotional and affective processes played a central role in the evolution of peaceful and cooperative human societies, rather than being solely a matter of emerging top-down control mechanisms. Most of chapter 4 and

---

[13] Top-down control is understood here as the processing of sensory and affective information that is driven by more cognitive processes such as goals or intentions. Bottom-up processing is the reverse of top-down processing, i.e., the processing of sensory and affective information that depends more directly on features of the stimulus input (for a more detailed discussion, see Rauss & Pourtois, 2013).

5 are devoted to understand the implications of this view about the role of emotions and affective processing for human normative cognition and development.

This hypothesis has consequences for ongoing philosophical debates. For example, recently there has been a lot of interest in the connection between morality, sexual selection, and cooperation (see, for instance, Alexander, 1987; Fraser, 2010; Joyce, 2006; Kitcher, 2006, 2011; Miller, 2000, 2007; Nesse, 2007; Ruse, 1986; Ruse & Wilson, 1986; Sober & Wilson, 1998; Tomasello & Vaish, 2013). But all these theories are built on the assumption that the social organization of early hominins closely resembled the social organization of the chimpanzee. If the sexual behavior of these hominins was less characterized by high levels of intermale and intersexual aggression than in chimpanzees, then the conditions for sexual selection would be radically different. Similarly, cooperation in a more socially tolerant ancestor would be different from the type of cooperation we find in highly hierarchical and aggressive primate social groups—the cognitive challenges are different and so are the mechanisms required to face them.

I shall illustrate this point with Philip Kitcher's (2011) hypothesis about the evolution of our capacity for normative guidance. According to Kitcher, the origins of the ethical project cannot be understood neither in terms of biological altruism nor in terms of behavioral altruism. The social life of our primate ancestors required a capacity for 'psychological altruism'—roughly, a capacity to align one's desires in response to the perceived desires of others, and not in expectation of some future benefit. In other words, Kitcher understands the emergence of human altruistic capacities as the gradual evolution of the cognitive and motivational psychological mechanisms underlying them (see also Sober & Wilson, 1998). This presupposes a form of belief-desire psychology, for "[…] altruists are intentional agents whose effective desires are other-directed" (p. 20). In this view, psychological altruism fostered complex forms of cooperation, and vice versa, that ultimately led to the appearance of norms and the beginning of ethical practice.

Yet according to Kitcher, psychological altruism in chimpanzees is limited in scope, as it was also in early hominins. To overcome these limitations, Kitcher argues, ancestral hominin groups developed quickly after the split with our sister lineage a capacity for normative guidance, i.e., a capacity to understand and respond to commands. He then offers a vindicating genealogy of this capacity since "[a]n ability to apprehend and obey commands changed the preferences and intentions of some ancestral hominids, leading them to act in greater harmony with their fellows and thus creating a more smoothly cooperative society" (p. 74; see also Kitcher 2006, p. 172).

Kitcher's genealogy of our capacity for normative guidance is vindicating because it leads to ethical progress, beginning with its ancestral role in remedying failures of altruism in our chimpanzee-like hominin ancestors:

> Tens of thousands of years ago, our remote ancestors began the ethical project. They introduced socially embedded normative guidance in response to the tensions and difficulties of life together in small groups. They were equipped with dispositions to psychological altruism that enabled them to live together, but the limits of those dispositions prevented them from living together smoothly and easily. Out of their normative ventures have emerged some precepts we are not likely ever to abandon, so long, at least, as we make ethical progress, the vague generalizations that embody ethical truths. (p. 409)

On Kitcher's view, the ethical project is a form of social technology that has played a central role in the gradual improvement of our hominin social life. This role is a vindicating one. Certainly, his strategy might seem unconventional since progress is usually explained in terms of truth. Instead, he thinks that his genealogy of moral cognition can make sense of ethical truth and ethical knowledge based on this notion of progress—the second part of his book is devoted to this issue. Progress is just functional efficiency. For moral practices have an original function, namely to remedy the failures of altruism that lead to social conflict. This is what Kitcher calls 'pragmatic naturalism'. As Kitcher put it: "Pragmatic naturalism retains a notion of ethical truth for expository purposes, but it starts from the concept of ethical progress" (p. 210).

As with any other genealogical argument, Kitcher's vindication of the ethical project is sensitive to issues about our hominin baseline. For his account of the role of normative guidance only makes sense in the context of a demonic male view. Male aggression is not a marginal feature of Kitcher's analysis since he takes chimpanzees, rather than bonobos, as the model for our hominid past (p. 59, footnote 40) and chimpanzee societies are male-dominated societies. This is not a marginal feature of Kitcher's analysis. In his view, the evolution of normative guidance was initially grounded in fear of punishment, and the actual beginnings of the ethical project are seen as a transition from a state of limited psychological altruism to one in which commands are followed out of fear (see, for instance, p. 87). This was so because the social life of our forebears was chimpanzee-like:

> Begin with chimpanzee societies in which a crude precursor of punishment is already present. Conflicts within these groups are often settled through the interventions of a dominant animal. Here rank or physical strength (or both as concomitants of each other) prevail, and a dispute is settled—not always, of course, through the infliction of pain or discomfort on the animal whose initial defection gave rise to the conflict. (p. 87)

In these social groups, the capacity to understand and obey commands was favored by natural selection because it helped us to avoid the cost of being punished by the dominant. Thus, Kitcher's view can be understood as a form of demonic male view.

Kitcher's account of our capacity for normative guidance is important and enlightening. But his evolutionary account relies too heavily on the so-called 'chimpanzee referential doctrine', i.e., the idea that chimpanzees are good referential models of our hominin ancestors (see Sayers, Raghanti, & Lovejoy, 2012), and a version of the demonic male view, i.e., the idea that dominance and male aggression were the cardinal challenges in the evolution of human sociality. His vindicating genealogy thus follows the typical narrative of this family of views in which top-down cognition plays the leading role in the expansion of the prosocial tendencies of our lineage, chaining the monster within. But as we will see later, if the model of our early ancestors I will propose here is right, Kitcher's account of the emergence of normative guidance would not be quite right. To the extent that his philosophical views rely on his evolutionary genealogy, they need to be reassessed in light of the plausibility of the different models of the social behavior of early hominins.

3.2. The puzzle of hominin evolution

Evolutionary explanations of cognition require a historical and a comparative context in order to determine the hominin baseline of social-cognitive capacities. This baseline can be established through research in comparative psychology (see, for instance, Kappeler & van Schaik, 2004; Tomasello & Call, 1997). Most of the supporting evidence for the proposed model I will present here comes, in particular, from the comparative literature between chimpanzees and bonobos. Chimpanzees and bonobos are our closest living relatives. According to current estimates, the human lineage diverged from the *Pan* lineage about 6 to 4.5 mya (Prüfer et al., 2012), while chimpanzees and bonobos diverged from each other more recently, about 1-2 mya. As a result, chimpanzees and bonobos are very similar in many respects, but they are also significantly different in key social and sexual behaviors (Boesch, Hohmann, & Marchant, 2002; de Waal & Lanting, 1997). This differences suggests that the *Pan/Homo* LCA could have been either chimpanzee-like or bonobo-like with respect to those key traits. More importantly, it also suggests that these social and sexual behaviors are evolutionary labile, since these differences evolved quickly from the genetic and developmental package inherited by the common ancestor of chimpanzees and bonobos.

The differences in social behavior are particularly intriguing. Chimpanzees show a clear linear dominance hierarchy among males, with male dominance over females (Goldberg & Wrangham, 1997). They also display relatively low levels of cooperation (Hirata & Fuwa, 2007). In contrast, hierarchical relationships among bonobos are not always clearly defined (Kanō, 1992). Female dominance is common, and it is based on female alliances against aggressive males (Vervaecke, de Vries, & van Elsacker, 2000). Moreover, experimental evidence also suggests that bonobos are more similar to humans in the way they solve various cooperative problems (Hare, Melis, Woods, Hastings, & Wrangham, 2007).

Sexual and play behaviors are different as well. In bonobos, sexual interaction occurs in mixed and same-sex pairings, and it is also used for conflict resolution (de Waal, 2001; de Waal & Lanting, 1997). Play behavior is common in adult bonobos, especially among females (Palagi, 2006). In contrast, chimpanzee sexual behavior is less rich and diverse. Sexual interaction does not typically occur in same-sex pairings, and (as in other primates) high-ranking males monopolize estrus females (Goodall, 1986). Unlike bonobos, play behavior is only frequent among chimpanzee infants (Goodall, 1986), and no gender bias in terms of play behavior has been found so far.

These behavioral differences are important because apes can be used as referential models, i.e., anatomical and behavioral proxies of our last common ancestor. In these models, the ethology, ecology, and cognitive skills of great apes are used to infer the traits that are most likely the ancestral condition of modern humans. These traits are either homologies (traits inherited from a common ancestor) or analogies (traits that have evolved independently due to similar selective pressures) or a combination of both. Moreover, although it is true that the recent split and stark differences between both species suggest that a wide range of social behaviors are quite plastic and evolutionarily labile, this could hardly be the whole explanation of these differences. As we will soon see, comparative studies in *Pan* show that neuroanatomical differences may be responsible for these behaviors, which indicates that these traits are not just a consequence of immediate differential responses to highly idiosyncratic socioecological factors.[14] Thus, given the behavioral differences between chimpanzees and bonobos, it is reasonable to assume that our early hominin ancestors were in part a mosaic of traits seen in both *Pan* species.[15]

This constitutes a puzzle for hominin evolution. For chimpanzees and bonobos are two very different models of our last common ancestor, especially with respect to some key social and sexual behaviors. In the next sections of this chapter, I will argue that our best model of the social behavior of early hominins is not only one that carries features of chimpanzees, bonobos, and probably other species, but also one that stresses the comparative similarities between bonobos and those early ancestors. This 'mosaic model', I claim, has important consequences for our understanding of the evolutionary trajectory of our distinctive prosocial tendencies.

3.3. The mosaic hypothesis

On the view I want to defend here, early hominins were a mosaic of different traits seen not only in chimpanzees but also in other primate species. I make this claim clear in what I call the 'mosaic hypothesis', i.e., the idea that the morphology and social

---

[14] Evolutionary lability can lead to these neuroanatomical differences. In plasticity-first hypotheses, phenotypic plasticity can produce developmental variants that might increase fitness (Levis & Pfennig, 2016). Selection can then refine the trait from an initial suboptimal version through genetic accommodation or even genetically assimilate the trait when environmental sensitivity is not favored (Moran, 1992; Waddington, 1953; West-Eberhard, 2003). However, although the robust neuroanatomical differences between chimpanzees and bonobos might be the result of some form of genetic accommodation or assimilation, they cannot be explained merely as an immediate response to environmental change or stress.

[15] Of course, this does not rule out the possibility that early hominins and the *Pan*/*Homo* LCA would have been in some respects very different from both *Pan* species. Along with comparative phylogenetic analysis (Duda & Zrzavý, 2013), the paleoanthropological record suggests that our early hominin ancestors were quite unlike chimpanzees or bonobos. Fossil evidence in *Ar. ramidus*, for instance, indicates that the *Pan*/*Homo* LCA could have possessed anatomical adaptations for bipedalism and omnivory. This evidence will be discussed in more detail in section 5.

behavior of early hominins, including the *Pan/Homo* LCA, can be reconstructed using comparative data from a wide range of extant and extinct primate species.

The key problem is then to determine which particular aspects should be included in the mosaic on the basis of the available evidence. I will focus, in particular, on a version of this hypothesis, similar to the one I have ascribed to Zihlman (Zihlman, 1984; Zihlman et al., 1978) and de Waal (de Waal, 1995, 2001; de Waal & Lanting, 1997), i.e., the idea that bonobos are to some degree a constitutive part of that mosaic. Of course, my concern here is not whether bonobos are closer to us than chimpanzees, which would not make sense given the current genetic evidence. Nor is it which species better resembles, say, the *Pan*/*Homo* LCA. I am not arguing for Zilhman's and de Waal's specific views either, but rather for what I take to be their essential insight.[16] My claim is a comparative one, namely that bonobos are in some important respects a more suitable model of the social behavior of early hominins and the *Pan/Homo* LCA than the chimpanzee, and that this undermines the demonic male hypothesis.

The overall picture of this comparative model is one in which early hominin ancestors were characterized by a level of social tolerance and prosocial skills that went beyond the usual chimpanzee referential model. This is not a minor issue. For increased social tolerance and enhanced prosocial skills diminish the role of aggression and dominance in the evolution of our lineage. Adaptations for tolerance and prosociality make the evolutionary trajectory toward seemingly distinctive human traits such as imitative learning (Galef, 1996, 2009; Tomasello, 2009a) or collaborative foraging (Tomasello, Melis, Tennie, Wyman, & Herrmann, 2012) more accessible.

The feasibility of the mosaic hypothesis and the Zilhman-de Waal conjecture is supported in the first place by genetic evidence (for a general discussion, see Pääbo, 2003). Recently, Prüfer and colleagues (2012) have completed the sequencing of the bonobo genome and have compared it to the already sequenced genome of chimpanzees and humans. According to this study, the bonobo genome is affected by incomplete lineage sorting among the three species, which occurs when an ancestral polymorphism persists, diverging only partially or not at all, within new evolutionary lineages

---

[16] For example, unlike the model I will propose in this section, Zilhman views focus on morphology (Zihlman & Bolter, 2015; Zihlman et al., 1978; Zihlman & Lowenstein, 1983). Zihlman ideas were quickly criticized on the assumption that bonobos might be an ecologically and morphologically divergent species instead of having traits homologous with early hominids (Stanford, 1998b; see also Latimer et al. 1981; Johnson 1981). As I will argue in section 5, there are reasons to think that certain aspects of the social behavior of bonobos are not as divergent as they may at first appear.

following a speciation event. This suggests that 3% of the human genome is more closely related to (i.e., more similar to homologues in) either chimpanzees or bonobos than these are to each other. They showed, in particular, that about 1.6% of the human genome is more closely related to bonobos than to chimpanzees while 1.7% of the human genome is more closely related to the chimpanzee than to the bonobo genome (Prüfer et al., 2012, pp. 2-3). Given the behavioral differences between chimpanzees and bonobos, they argue that, at least in principle, the last common ancestor of these three species could have possessed traits seen in both *Pan* species (2012, p. 527).

Technical box 3.1. Comparative analysis of cooperation in great apes and humans

Lineage sorting is the process by which, following the separation of two species, the ancestry of every gene converges to the overall phylogeny of the species. This is also called ‘coalescence’, i.e., the convergence of the genealogy of multiple gene copies backward in time into their common ancestor. Incomplete lineage sorting implies, in contrast, discordance between genealogies. For example, when one compares the genome of three related species, such as gorillas, chimpanzees, and humans, one infers that humans and chimpanzees are closer to each other than they are to gorillas through an average pattern of relatedness. This pattern suggests a more recent divergence between chimpanzees and humans than between gorillas and humans. However, this is typically only an average result. For if we compare these genomes closely, we will find specific genes and DNA regions where the gorilla sequence is more similar to the human sequence than to the chimpanzee sequence (Scally et al., 2012). If we looked at only these genes or DNA regions, we would conclude that gorillas are closer to us than chimpanzees. As a result, this phenomenon produces different evolutionary trees for humans and many other primates, depending on which DNA fragment is used for the analysis. No clear path of common ancestry between humans and various primates can be inferred. The same occurs with chimpanzees, bonobos, and humans with respect to a relatively low percentage of about 3% of our genome. What Prüfer and colleagues (2012) found was that about 1.6% of the human genome is more closely related to (i.e., more similar to homologues in) bonobos than chimpanzees while 1.7% of the human genome is more closely related to the chimpanzee than to the bonobo genome.

Certainly, the presence of incomplete lineage sorting in chimpanzees and bonobos is not an argument for the view that bonobos actually possess traits of the *Pan*/*Homo* LCA that the chimpanzee does not—although this would be possible if the chimpanzee lost an ancestral trait that bonobos and humans kept. Instead, what I am arguing here is that given the behavioral differences between chimpanzees and bonobos, incomplete lineage sorting in humans, chimpanzees, and bonobos indicates that the last common ancestor of these three species could have possessed traits seen in one but not the other *Pan* species.

This genetic evidence not only gives *prima facie* motivation for the idea that the *Pan/Homo LCA* had some bonobo-like traits. It also suggests that bonobos can be useful referential models—a methodological assumption that underlies comparative studies on ape cognition (see, for instance, Tomasello & Call, 1997). The value of bonobos as models of early hominins is likely not only limited to common ancestry, though. It is also plausible that many features we see in this extant species resemble those we see in human because both species underwent similar selective regimes, e.g., the retention of juvenile traits such as playfulness and social tolerance due to a parallel process of self-domestication. Roughly, self-domestication refers here to a process of social selection against aggression in humans that resembles the process of domestication of other wild animals to humans without systematic human selective breeding (Hare et al., 2012; Wrangham, 2011; see also section 2.6 in this chapter for discussion). Either way, I would like to argue that it is quite possible that our early hominin ancestors, and even perhaps the *Pan*/*Homo* LCA, were characterized by:

(i) group hunting behavior,
(ii) enhanced emotional control,
(iii) increased aversion against aggression (specially intermale and intergroup aggression),
(iv) enhanced brain connectivity for empathy (top-down and bottom-up control of aggressive impulses),
(v) increased mind reading skills,
(vi) increased cooperative and sharing tendencies,
(vii) non-linear/ill-defined hierarchy,
(viii) and non-exclusive male dominance.

Traits (ii)-(vi) are comparative features, i.e., they are traits of early hominins that are well above the hypothesized levels of a chimpanzee-like model of that ancestor (for a defense of a picture of early *Homo* quite similar to this, see Gamble, Gowlett, & Dunbar, 2014) . According to these features, the social life of our early ancestors was in these respects more bonobo-like than chimpanzee-like. For the evolutionary trajectory would be less constrained by our aggressive and dominant tendencies, such that overcoming them would be displaced (so to speak) from the center of gravity of our evolutionary narrative.

Moreover, from a philosophical point of view, this model would lead us to reassess naturalistic arguments based on these assumptions, such as Kitcher's evolutionary narrative of the emergence of our capacity for normative guidance. For the above model would be linked to a different picture of the trajectory of hominin social evolution and the timing of the appearance of more complex forms of social cognition. As I will argue later, the fossil record supports the view that very early in our lineage, hominins were less aggressive and more tolerant than commonly assumed by chimpanzee referential models. Moreover, current explanations of our unique human cognitive capacities assume that they emerged relatively late with the emergence of the genus *Homo*.

3.4. Evidence for the model

The features of the proposed model are closely linked to social behavior. In behavioral phylogenetics, it is possible to reconstruct an ancestor's behaviors if such behaviors are present in all of its living descendants (Boehm, 1999, 2012; Brosnan, 2006; Wrangham & Peterson, 1996). This argument relies on considerations of parsimony. To the extent that parsimony is a guide, group hunting would be characteristic of our last common ancestor. For recent evidence shows that this behavior is also present in the bonobo (Surbeck, Fowler, Deimel, & Hohmann, 2009; Surbeck & Hohmann, 2008). The same goes for some aspects of physical cognition such as tool manufacture and use (Gruber, Clay, & Zuberbühler, 2010; Ingmanson, 1996).

Since humans also possess those behavioral traits, it is possible to infer that the *Pan*/*Homo* LCA did (i) hunt in groups—although only small game, and not as a core, essential feature of their subsistence strategy. It is true that, given that traits such as tool manufacture and use are present in all great apes (Breuer, Ndoundou-Hockemba, & Fishlock, 2005; Goodall, 1964; van Schaik, Fox, & Sitompul, 1996) and also in other primate species (Chevalier-Skolnikoff, 1990; Chiang, 1967; Fernandes, 1991; Oyen, 1979; Phillips, 1998), their presence in early hominins is a more conservative phylogenetic inference than group hunting. The set of data points is significantly smaller in that case. Nonetheless, there is evidence that by 3.4 mya hominins were using stone tools to hunt large mammals (McPherron et al., 2010), which pushes the plausibility of ape-like hunting much deeper in the hominin lineage. Therefore, it is just as likely, if not more likely, that group hunting was present in the *Pan*/*Homo* LCA as assuming that it emerged very early in our lineage and then independently in *Pan*.

In addition, the neural circuitry that mediates anxiety, empathy, and the inhibition of aggression in humans is better developed in bonobos than in chimpanzees (Rilling et al., 2012). Recent comparative studies have shown that the bonobo has a more human-like circuitry for key nodes in the limbic system, including the amygdala, the hypothalamus, and the anterior insula (Rilling et al., 2012). The limbic system plays a crucial role in emotional processing, e.g., the anterior insula and the amygdala are both implicated in human empathy. Moreover, two pathways, one connecting the amygdala and the anterior cingulate cortex, and another connecting the amygdala and the ventromedial prefrontal cortex, are larger in bonobos than chimpanzees. The former is implicated in emotion regulation in humans while the latter enables the restraint of aggression via top-down suppression of aggressive impulses from the amygdala (Davidson, Putnam, & Larson, 2000; Meyer-Lindenberg et al., 2006; Pezawas et al., 2005). The same pathway may also be involved in controlling aggressive impulses through a bottom-up relay of perceived distress in others to the ventromedial prefrontal cortex that inhibits antisocial behavior (Blair, 2007, 2008). This would mean that, for example, even if cooperating were a good decision for instrumental reasons, those instrumental reasons would have problems influencing chimpanzee behavior because they often would require controlling aggressive impulses via a relatively underdeveloped pathway. Similarly, there would also be little influence from bottom-up biases against instrumental forms of aggression because perceived cues of distress from others would not bias behavior. This would be different in bonobos and humans. When our amygdala senses that our actions are causing someone else distress, we may use that pathway to adjust our behavior in a prosocial direction.

Insofar as the above neurobiological traits are examples of fine-grained similarities, then parsimony suggests that the early hominins possessed (ii) enhanced emotional control, (iii) increased aversion against aggression, and (iv) enhanced brain connectivity for empathy with respect to a hypothetical chimpanzee-like model of these ancestors. A broader look at the neurobiology of other empathic and tolerant primate species gives some additional support to this view. For callitrichid monkeys, for instance, are quite socially tolerant but their social behavior relies on somewhat different neural circuitry. They possess small brains and their empathic behavior is mediated by physiological responses that are especially geared to cooperative breeding (Fernandez-Duque, Valeggia, & Mendoza, 2009). This indicates that empathy and emotion regulation are not necessarily related to an increase in gray and white matter connectivity as in bonobos and humans, which makes a hypothesis about convergent evolution less likely.

Bonobos are also more socially tolerant than chimpanzees, especially when co-feeding (Hare et al., 2007). They show a stronger stress hormone response to feeding competition (Wobber et al., 2010). They have also been described as more nervous and shy than chimpanzees (de Waal & Lanting, 1997; Herrmann, Hare, Cissewski, & Tomasello, 2011). As in humans, these differences in temperament are associated with enhanced social-cognitive skills. Studies with young children, for instance, show a strong connection between shyness and mindreading skills (Wellman, Lane, LaBounty, & Olson, 2011). Similarly, bonobos outperform chimpanzees in tasks related to mindreading, while chimpanzees are more skilled at tasks requiring the use of tools and an understanding of physical causality (Herrmann, Hare, Call, & Tomasello, 2010; see also Rosati & Hare, 2012; Wobber, Herrmann, Hare, Wrangham, & Tomasello, 2014)).

Differences in mindreading skills, however, cannot be explained solely on the basis of social tolerance. These differences are products of a particular neural system for understanding the intentional states of others. The medial prefrontal cortex and the temporoparietal junction are known to be implicated in mindreading capabilities in humans (Gallagher & Frith, 2003; Saxe & Kanwisher, 2003; Saxe, Xiao, Kovacs, Perrett, & Kanwisher, 2004). Thus, the fact that bonobos also have increased gray matter in the dorsomedial prefrontal cortex compared with chimpanzees seems to be telling. Mindreading skills in apes are typically linked to competitive contexts (Call & Tomasello, 2008), but there is no reason to think that food and mating competition is stronger in bonobos than chimpanzees. Thus, explaining this increased capacity in bonobos through a convergent selective gradient seems problematic.

Levels of tolerance also affect sharing behavior in *Pan*. Chimpanzees share food with conspecifics only under some circumstances—e.g., food transfer from mother to offspring (Ueno & Matsuzawa, 2004) or when the food is not valuable and not monopolizable (Blurton-Jones, 1987; de Waal, 1989; Gilby, 2006). However, peaceful food sharing in wild bonobos seems to contradict the usual sharing-under-pressure hypothesis (Yamamoto, 2015). As the possessor of a food resource becomes satiated over the time, the relative value of the resource decreases for the possessor with respect to the non-possessors. Thus, the sharing under pressure hypothesis predicts that the utility costs caused by the pressure exerted by the non-possessors (e.g., the risk of aggression or the time invested in monopolizing a resource whose value diminish with time) is the explanation of food transfer. However, under experimental conditions, active and voluntary food sharing also seems to be present in bonobos (Hare & Kwetuenda, 2010), even among strangers and when food is easily monopolizable (Tan & Hare, 2013). In a slightly different way, Bullinger and colleagues (2013) have argued that bonobos do not have a preference to feed together, but rather a preference to be together. Either way, inhibition of aggression and social tolerance are linked to sharing.

Moreover, experimental evidence supports what has been called the ‘emotional reactivity hypothesis’. Recent studies suggest that selection on emotional reactivity critically shapes a species’ ability to solve social problems (Hare et al., 2005; Hare & Tomasello, 2005). This hypothesis, for instance, predicts that bonobos will cooperate more successfully in food-retrieval tasks than chimpanzees because tolerance levels are higher in bonobos. So, although experimentally both species have been shown to be equally successful at cooperating when food is difficult to monopolize, tests with monopolizable food have shown that bonobos are much more able to cooperate than chimpanzees. For example, in Hare and colleagues’ (2007) food retrieval paradigm, a food resource was placed on a platform such that the resource could only be retrieved if two subjects pulled both ends of a rope at the same time. When the food was difficult to monopolize because there were two piles of food placed at either end of a platform, chimpanzees and bonobos performed the task equally well. But when food was placed in a single pile in the center of the platform and, therefore, was easily monopolizable, the bonobos outperformed the chimpanzees, *i.e.*, bonobos cooperated more often, and after successful cooperation, they co-fed more. Same results were obtained with different cooperative partners, even though the bonobos were naïve before testing while the chimpanzees had been previously tested using similar experimental setups (Melis, Hare, & Tomasello, 2006a, 2009).

Given the differences in temperament between chimpanzees and bonobos, it is at least as plausible that early hominins possessed (v) increased mindreading skills and (vi) increased cooperative and sharing tendencies with respect to a hypothetical chimpanzee-like model of the *Pan*/*Homo* LCA as it is to adopt the standard chimpanzee referential model. This is a non-negligible difference in social-cognitive abilities. The fact that these differences are correlated with particular neurobiological similarities between bonobos and humans also deserves attention. The chimpanzee's mindreading and cooperative capacities cannot simply be taken to represent the ones of early hominins.

What is more, even if the *Pan*/*Homo* LCA was very chimpanzee-like, the social-cognitive differences seen in bonobos suggest that a fairly tolerant and cooperative primate was relatively readily evolvable from the genetic and development package inherited by the *Pan*/*Homo* LCA. Tolerance and cooperation did not depend on the subsequent evolution of high-end cognitive capacities beyond those possessed by great apes. Evolutionary plasticity could have been enough to adapt the three lineages to different socioecologies, which suggests that social tolerance and other social-cognitive traits could have evolved very early as a result the evolutionary plasticity of our common ancestor. There are good reasons to think that aggression and dominance are not traits deeply entrenched in our lineage. So, one could argue that chimpanzee-like sociality and cognition did not impose sharp constraints on evolutionary options, somewhat independently of the actual character of our last common ancestor.

Another aspect of the social behavior that one could expect to be influenced by structural similarities in the hypothalamus and the amygdala is sex. These brain regions play a central role in controlling sexual behavior in non-human animals (Breedlove, Watson, & Rosenzweig, 2010; Newman, 1999) as well as in processing visual sexual stimuli in humans (Hamann, Herman, Nolan, & Wallen, 2004; Karama et al., 2002). This suggests that bonobos could use sex to reduce tension and anxiety produced by everyday social interactions (de Waal, 2001; de Waal & Lanting, 1997; Hare et al., 2007; Hohmann, Mundry, & Deschner, 2009; Parish, 1996). For example, the enlarged bonobo amygdala could elicit anxiety while motivating sexual behavior that helps reduce it (Rilling et al., 2012).

Naturally, sexual behavior in all the three species has important differences. But a crucial similarity between bonobos and humans is that both species use sexual behavior in a social context. Unlike chimpanzees, female bonobos are continuously sexually active and attractive. So, in bonobos and humans, sexual intercourse can be initiated at any point, which in turn increases bonding between individuals. Bonobos with lower testosterone levels and attenuated testosterone responses engage more often in amicable relationships with unrelated females and have greater reproductive success (Surbeck, Deschner, Schubert, Weltring, & Hohmann, 2012). Therefore, bonobo males benefit from affiliative long-term association with females (Hohmann & Fruth, 2003;

Hohmann, Gerloff, Tautz, & Fruth, 1999; Surbeck et al., 2012), which facilitates more egalitarian and more peaceful social lives.

Similarly, hypothalamus size and amygdala size have been shown to predict social play frequency in non-human primates but not nonsocial play (Lewis & Barton, 2006). Bonobos—females more than males—seem to use play to assess physical skills, the willingness of other individuals to invest in a relationship, and to strengthen already existing social bonds. Adult bonobos play much more frequently than chimpanzees. This asymmetry is important because it has been experimentally shown that both species use grooming and play as social currency (Schroepfer-Walker, Wobber, & Hare, 2015). Play is a valuable social interaction and can be used to establish social preferences depending on the amount of playful interactions between individuals. Thus, play behavior could also have a crucial role in the bonobo social organization and its typically weak dominance hierarchy.

To the extent that the above neurobiological similarities are correlated with the more egalitarian social structure of bonobos, they would suggest that early hominins lived in (vii) less hierarchical and arguably (viii) less male-dominated social groups with respect to a hypothetical chimpanzee-like model of the *Pan*/*Homo* LCA. Explanations of the evolution of the bonobo usually argue that reduced male aggression toward females was sexually selected (Hare et al., 2012; Wrangham & Peterson, 1996). But it is at least equally likely that this trait was inherited from the common ancestor, especially in light of the fact that the traditional evolutionary scenario for the split between chimpanzees and bonobos is not supported by our current knowledge about the formation of the Congo River (Takemoto, Kawamoto, & Furuichi, 2015).[17]

Granted, this is not conclusive evidence for the mosaic hypothesis or the particular model I have offered in the previous section. However, even if the case for the model is not compelling enough, we have good reasons to think that the social behavior of early hominins, including the *Pan*/*Homo* LCA, was in many respects not

---

[17] According to this hypothesis, the formation of the Congo River isolated an ancestral population of the common ancestor of chimpanzees and bonobos around 2 mya (Wrangham, 1993; Wrangham & Peterson, 1996). This population did not have to compete with gorillas for resources, which allowed females to form coalitions and resist the advances of males. Since coercion was not an efficient mating strategy, sexual selection favored less aggressive males. This led to the evolution of bonobos and their distinctively low levels of aggression. However, the current geological evidence contradicts this scenario because it indicates that the present Congo River was formed much earlier, around 34 mya.

chimpanzee-like. At the very least, the chimpanzee referential model should not be anymore the default assumption.

3.5. Paleoanthropological support

Although certainly thin, the above evidence suggests that the *Pan*/*Homo* LCA was in some respects more bonobo-like than chimpanzee-like. In this section, I will argue that even if the *Pan*/*Homo* LCA was not characterized by the features ascribed in the model, we still have reasons to think that they evolved very early in our lineage. For the paleoanthropological evidence suggests that early hominins were much more socially tolerant than the chimpanzee referential doctrine actually tells us.

Fossil evidence is central to whatever model of our hominin ancestry we choose. Referential models are constrained by phylogenetic inferences. After all, phylogenetic analysis can be understood as a form of referential modeling—one that does not rely on individual species as models but species traits' distributions as indicators of their conserved or derived nature (Duda & Zrzavý, 2013; Moore, 1996; Whiten et al., 2010; Wrangham, 1987). On top of this, fossil evidence particularly restricts the scope and shape of these models. Generally speaking, for instance, referential models are either based on homology through shared descent (McGrew, 1981) or analogy through convergent evolution (DeVore & Washburn, 1963; Fernandes, 1991; Jolly, 1970, 2001; Perry, 1997; Perry et al., 2003; Skybreak, 1984). The above model can be considered rather neutral regarding this issue.

Paleoanthropological evidence, however, suggests that even if some aspects of the proposed model are not homologies, i.e., ancestral traits of the *Pan*/*Homo* LCA that have been retained by bonobos and humans, they might have evolved fairly early in our lineage. This view is supported by fossil evidence from *Sahelanthropus*, *Orrorin*, and *Ardipithecus*, which indicates that our lineage was less aggressive and less male-dominated than assumed by the traditional chimpanzee referential model because these extinct homininae species lack the adaptations for agonistic male-male competition present in other living primates—i.e., large canine and body-size sexual dimorphism (Brunet et al., 2002; Haile-Selassie, 2001; Haile-Selassie, Suwa, & White, 2004; Haile-Selassie & WoldeGabriel, 2009; Lovejoy, 2009; Pickford & Senut, 2001; Suwa et al., 2009; White, Asfaw, et al., 2009; see also Pickering, 2013). Haile-Selassie and colleagues (Haile-Selassie et al., 2004; Haile-Selassie, Suwa, & White, 2009), for instance, see enough anatomical continuity between all three to suggest that they might belong to the same genus—pushing the earliest known appearance of *Ardipithecus* one million years deeper into prehistory, based on *Sahelanthropus*, at 7 mya (Pickering, 2013, pp. 28-29).

Early hominins and the *Pan*/*Homo* LCA could also have been very different from both *Pan* species. Fossil evidence from *Ar. ramidus*, for instance, indicates that this early hominin was well-adapted to bipedalism, although it retained arboreal capabilities (Lovejoy, Latimer, Suwa, Asfaw, & White, 2009; White, Asfaw, et al., 2009). This means a more human-like locomotion system, quite different from that seen in any extant ape. Another important difference is that *Ar. ramidus* appears to be neither a ripe fruit specialist like *Pan*, nor a folivorous browser like *Gorilla*, but rather a more generalized omnivore (Suwa et al., 2009; White, Asfaw, et al., 2009). However, the same fossil evidence also suggests that the social behavior of the *Pan*/*Homo* LCA was in many important respects more bonobo-like than chimpanzee-like and that this social behavior is likely to be an ancestral condition.

Evidence from *Ardipithecus ramidus* is particularly telling. The fossil record of this ancestor is rich and the completeness of some remains makes sex assessment relatively reliable (White et al., 2015; White, Suwa, & Asfaw, 1994). Dating estimates place the earliest remains of this hominin at circa 4.4 mya, close enough to the split between these two lineages as to make this ancestor highly relevant for reconstructing the morphology and behavior of the *Pan*/*Homo* LCA (see figure 3.1).

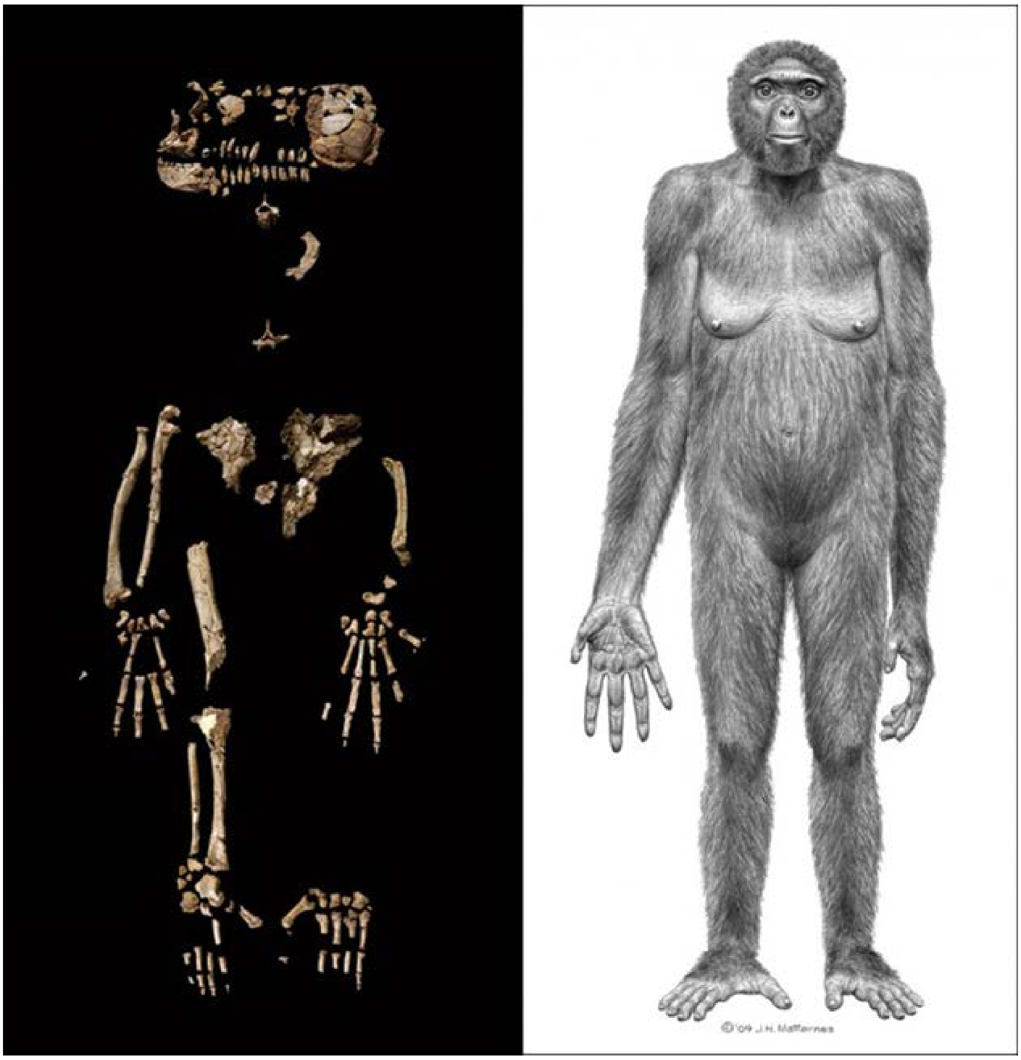

Figure 3.1. Partial skeleton (right) and artist’s reconstruction (left) of a relatively complete female specimen of *Ar. ramidus*, ARA-VP 1/500. The specimen, nicknamed ‘Ardi’, is estimated to be around 120 centimeters tall and to weigh around 50 kilograms. The remains were dated as about 4.4 mya based on their stratigraphic position between the volcanic strata of the Gaala Tuff Complex and the Daam Aatu Basaltic Tuff. Source: http://www.scien§cemag.org/

Large canine and body-size sexual dimorphism are important because they are traits associated with agonistic male-male competition. But *Ar. ramidus* remains reveal that this hominin was characterized by reduced canine teeth and low body size dimorphism. In basal dimensions, the canines of *Ar. ramidus* are approximately as large as those of female chimpanzees and male bonobos, although their crown heights are shorter; they are comparable to those of *Australopithecus anamensis* and *Australopithecus afarensis* (Suwa et al., 2009; White, Asfaw, et al., 2009; White et al., 1994; White et al., 2006; see also supplementary material of Suwa, Kono, et al., 2009).

They are also ‘feminized’ in shape. The size of the upper canine tooth is not only similar to that of females but also less sharp than those of chimpanzees.

Reduced canine teeth dimorphism is a common feature of the hominin clade (Brace, 1972; Greenfield, 1992; White, Asfaw, et al., 2009; Wolpoff, 1976). Along with *Ar. ramidus*, this trait is seen in *Sahelanthropus* (Brunet et al., 2002, p. 150), *Orrorin* (Senut et al., 2001), and *Ar. kadabba* (Haile-Selassie, 2001; Haile-Selassie et al., 2004; Haile-Selassie & WoldeGabriel, 2009). Since the canine tooth is usually used as a weapon in intermale and intergroup conflicts, the less pronounced upper canine teeth suggests that early hominins, including *Ar. ramidus*, were characterized by relatively little intermale and intergroup aggression compared to chimpanzees.

Similarly, *Ar. ramidus* is also expected to have shown little sexual dimorphism in body size—comparable to that of chimpanzees or humans, as opposed to orangutans or gorillas (White, Asfaw, et al., 2009, p. 80; White et al., 2015, p. 4881). In higher primates, body size dimorphism is usually coupled with strong canine dimorphism (Plavcan & van Schaik, 1997). Using dimorphism to infer behavior in early hominids is usually problematic because their unique combination of minimal canine size dimorphism and intense body mass dimorphism (Plavcan & van Schaik, 1997). But this is not the case in *Ar. ramidus*, which is characterized by a combination of reduced canine and body size dimorphism, unlike *Pan* and other early hominids. As a consequence, lack of sexual dimorphism seems to indicate that males did not compete against each other for dominance.

While intermale and intergroup aggression is frequent among chimpanzees, *Ar. ramidus* possessed low levels of agonistic male-male competition (Clark & Henneberg, 2015; Suwa et al., 2009)—and even, perhaps, male-female codominance as in bonobos (Suwa et al., 2009, p. 57). We cannot be sure about these aspects of the social behavior of our early ancestors, but we can infer them indirectly. For early hominins do not seem to have any of the adaptations for agonistic male-male competition present in other living primates. This would suggest that the frequency of intermale and intergroup aggression seen in chimpanzees is likely a derived condition, which in turn compromises its status as the behavioral model of the *Pan*/*Homo* LCA.

Parallel evolution does not always seem to give us the most parsimonious reconstruction of these traits. Chimpanzees are more sexually dimorphic than bonobos and humans, and australopithecines were more sexually dimorphic than both extant *Pan* species. Therefore, to the extent that australopithecines are direct ancestors of modern humans (and not a paraphyletic sister lineage, which they may be), this loss of sexual dimorphism must have not only occurred twice independently, in *Pan* and in *Homo*, but also in *Ar. ramidus* (see figure 3.2).

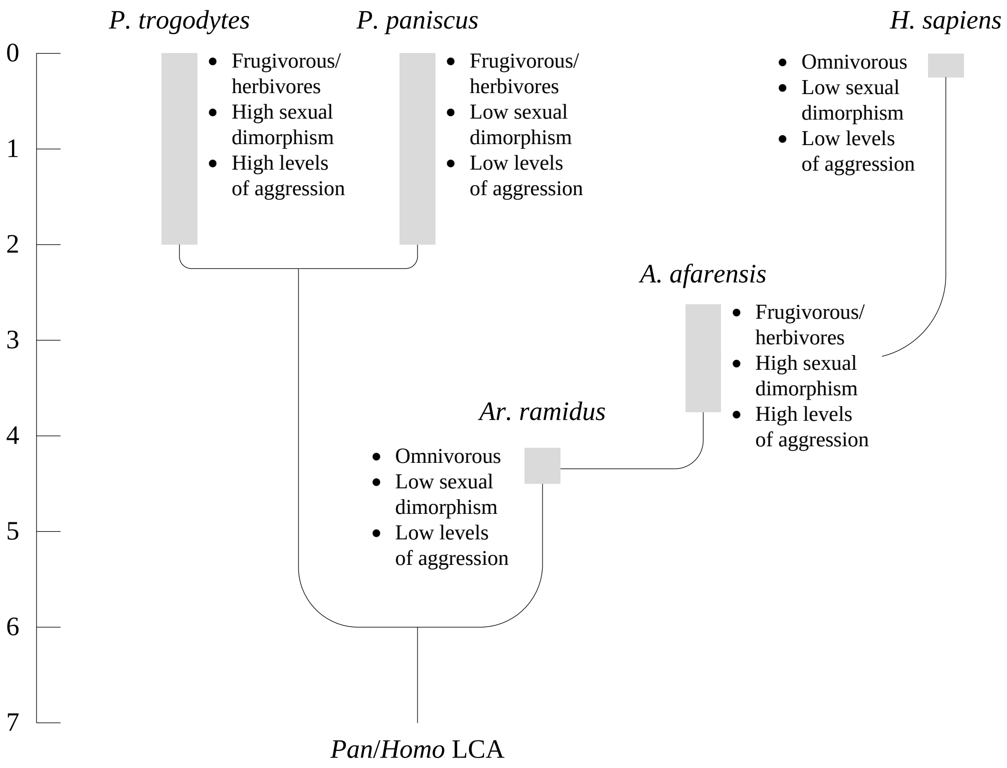

Figure 3.2. Traditional phylogenetic arrangement of five hominid species. According to this view, *Ar. ramidus* and *A. afarensis* are direct ancestors of humans (*H. sapiens*). But *Ar. ramidus*, bonobos (*P. paniscus*), and humans are characterized by low sexual dimorphism and low levels of intermale and intergroup aggression while *A. afarensis* and chimpanzees (*P. troglodytes*) show increased levels of sexual dimorphism and aggression. On top of that, *Ar. ramidus* and humans are characterized by an omnivorous diet and a similar dentition, although *A. afarensis* and *Pan* have specialized masticatory apparatus.

Another option would be to suggest that low sexual dimorphism and reduced intermale and intergroup conflict is, in fact, the ancestral condition, with a pattern of increasing dimorphism in australopithecines and chimpanzees. Australopithecines would be a paraphyletic sister lineage (an alternative pointed out to me by Kim Shaw-Williams, personal communication), or not as sexually dimorphic as it has often been claimed (McHenry, 1991; Plavcan, Lockwood, Kimbel, Lague, & Harmon, 2005; Reno, McCollum, Meindl, & Lovejoy, 2010; Reno, Meindl, McCollum, & Lovejoy, 2003). In this way, the evolutionary trajectory of the human lineage could be explained without postulating so many evolutionary reversals, or by invoking less drastic shifts. However, this would challenge the current picture of *Australopithecus* as a very aggressive, highly sexually dimorphic genus, or even its place as direct human ancestor—a hypothesis that, nonetheless, would be worth exploring (see figure 3.3).

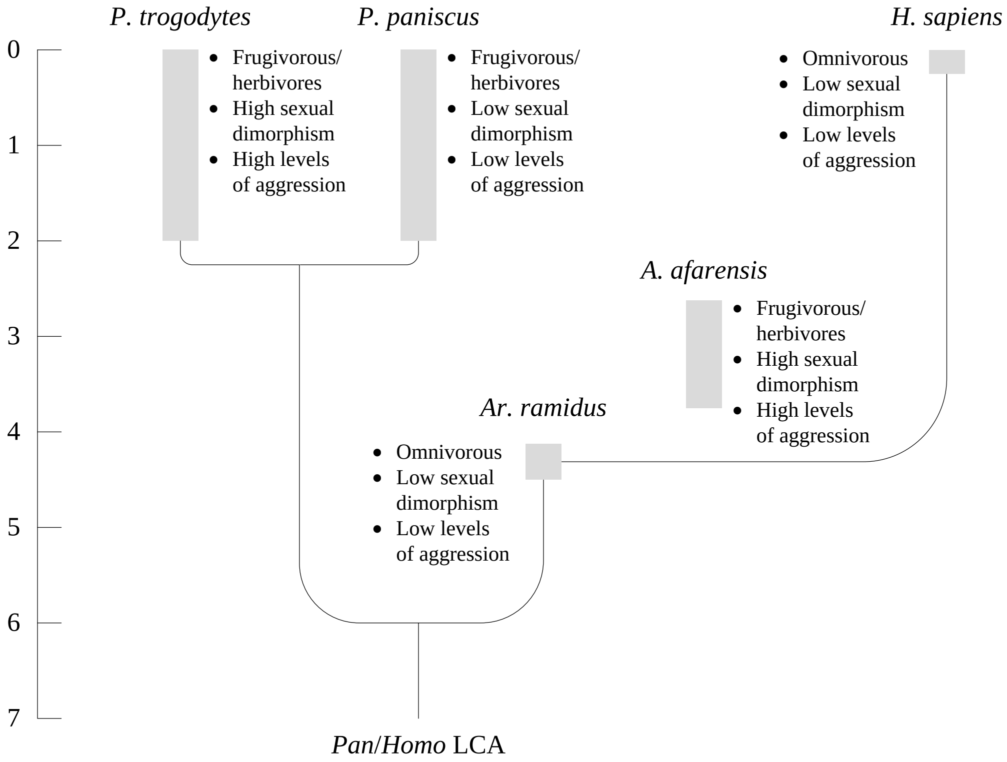

Figure 3.3. Alternative evolutionary arrangement of five hominid species. According to this view, *Ar. ramidus* is a direct ancestor of humans (*H. sapiens*), while *A. afarensis* is part of a paraphyletic sister lineage. A dashed line bypassing *A. afarensis*, but connecting *Ar. ramidus* with humans represents this alternative phylogenetic relation. This provides a more parsimonious reconstruction by making low sexual dimorphism and reduced intermale and intergroup conflict the ancestral condition.

Placing *A. afarensis* as our direct ancestors is an important issue if we assume that australopithecines were indeed a very aggressive, highly sexually dimorphic genus. If they are our direct ancestors, then the social behavior of early hominins was likely chimpanzee-like before the split between *Pan* and *Homo*. If they are a side branch, and early hominins were ardipithecines all the way through to encephalization, then early hominins were probably very unchimpanzee-like, more similar to bonobos and humans at least in some important respects of their social and sexual behavior. Except that on either reconstruction, it seems that social and sexual behavior come out as evolutionarily plastic, so we cannot put very much weight on phylogenetic parsimony inferences.

It is not clear whether body size dimorphism in australopithecines is a consequence of male-male competition (Plavcan & van Schaik, 1997) since their canines have a variety of features inconsistent with their use as a weapon (Greenfield, 1992). There are multiple reasons that could potentially explain the increase in body size dimorphism in australopithecines, e.g., reduction in female body size (Leigh & Shea, 1995), predator defense (Clutton-Brock, Harvey, & Rudder, 1977), the Rench's rule, i.e., the idea that size dimorphism will increase with increasing body size when the male is the larger sex (see Fairbairn, 1997), or even niche differentiation, as in raptors where the female is typically larger than the male.

Although there is no necessary link between the specific features ascribed by the model and the paleoanthropological evidence, the common theme of reduced (or controlled) aggression in early hominins stands, which would make these ancestors behaviorally more similar to bonobos and humans than chimpanzees in this respect. Even if the *Pan*/*Homo* LCA was very different from the proposed model, a decrease in these aggressive tendencies seems to have occurred very early in our lineage. If this were the case, many of the features ascribed by the proposed model would be cases of parallel evolution rather than common ancestry. They could be explained, for instance, by parallel heterochronic shifts of ancestral developmental patterns that have cascading effects on social behavior. For as we will see in the next section, there is an explanation for the emergence of similar social behavioral traits in bonobos and humans due to a process of self-domestication, which lead to the retention of juvenile traits.

3.6. The self-domestication hypothesis

In the last two sections, I have argued that the demonic male view of human evolution and the chimpanzee referential doctrine very likely do not give us an accurate picture of early hominins. In section 3.4, I have argued that comparative evidence suggests that the *Pan*/*Homo* LCA was in some respects more bonobo-like than chimpanzee-like. In section 3.5, I have argued that even if the *Pan*/*Homo* LCA is not correctly described by the model, these features could have evolved very early in our lineage since the paleoanthropological evidence suggests that early hominins were much more socially tolerant than the chimpanzee referential doctrine actually tells us. This is a plausible evolutionary scenario given the evidence that many of the social-cognitive traits of the model are evolutionary plastic. As a result, many of the features ascribed by the proposed model could be cases of parallel evolution rather than common ancestry—i.e., humans and bonobos would share these traits via an ancestral character that independently evolved in similar ways.[18]

The increased social tolerance and the enhanced prosocial skills that the model presupposes, occurred either before the split between the ape and hominin lineages or very early in the evolutionary history of the latter. They can be explained either by parallel evolution, or by shared recent ancestry, or (quite likely) by a combination of both—i.e., with some features of the model being cases of parallel evolution in bonobos and humans while others being inherited from the common ancestor. This would help explain why the social life of our early ancestors was in these respects more bonobo-like than chimpanzee-like and, consequently, why these two species are so different. But although I think convergent, parallel evolution is possible, I am skeptical that all the comparative claims of the model could be explained by this process alone. As a result, there are reasons to think that at least some aspects of the proposed model are real homologies inherited from the *Pan*/*Homo* LCA. To be clear, this debate should not affect the overall picture of the thesis. For both sides in the debate only differ in terms of the timing and the underlying evolutionary processes that account for the baseline of my lineage explanation, rather than in terms of the baseline itself. But the same kind of processes behind scenarios of parallel evolution could have also played an important role much later in the evolution of the genus *Homo*—in particular, when our ancestors became cooperative breeders (Hawkes, 2014; Hrdy, 2016; Tomasello & Gonzalez-Cabrera, 2017).

The best case for parallel evolution comes from the so-called ‘self-domestication hypothesis’ (Hare et al., 2012; Henrich, 2016; Wrangham, 2011; for a precursor of this theory, see Lorenz, 1940). On this view, the differences seen between chimpanzees and bonobos are a consequence of selection against aggression that led to a syndrome of changes observed in domestic animals, including the retention of juvenile traits. Domesticated animals show a number of morphological, physiological, behavioral, and cognitive modifications that seem to be correlated, e.g., variations in body coloration, cranial shape, dentition, brain size, activation of the hypothalamic-pituitary-adrenal axis, and problem-solving abilities. These changes are thought to resemble the ones undergone by bonobos since their split from the chimpanzee lineage. As a result, it has been argued that the alleged similarities between bonobos and humans might be

---

[18] I understand parallel evolution as a special case of convergent evolution, where similar traits have appeared independently in different closely related taxa such that the traits in question evolved from the same ancestral character (McGhee, 2011, p. 3).

analogous traits due to a parallel process of selection against aggression, rather than common ancestry.

The key evidence in favor of this scenario relies on the developmental trajectory of the skull in chimpanzees and bonobos. It is sometimes argued that the cranial development of the chimpanzee closely resembles the ontogenetic pattern of the gorilla (Shea, 1983, 1989), while the bonobo skull remains small and juvenilized when compared to chimpanzees, other great apes, and australopithecines (Kappelman, 1996; Lieberman, Carlo, de Leon, & Zollikofer, 2007; Schultz, 1941). For this reason, a chimpanzee-like trajectory of cranial ontogeny is sometimes considered the most parsimonious reconstruction of the ancestral pattern, while the bonobo developmental trajectory is taken to be derived (Wrangham & Pilbeam, 2001). In this line of argument, Hare and colleagues (2012) have suggested that the non-aggressive behavior seen in bonobos would be also derived because brain development is known to be correlated to cranial ontogeny in particular ways (see figure 3.4).

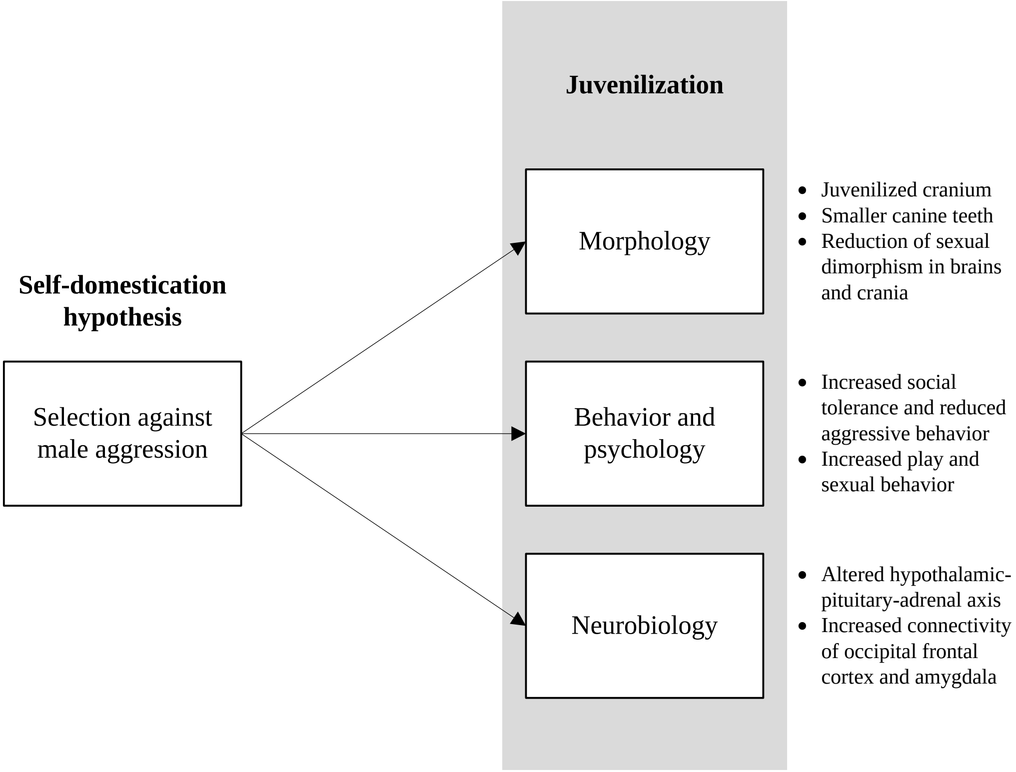

Figure 3.4. Model of bonobo evolution due to self-domestication. Selection against male aggression leads to a process of juvenilization that is correlated with morphological, behavioral, and neurobiological changes. If the developmental trajectory of the bonobo skull proof to be derived in this way, then this would support the idea that the non-aggressive behavior of bonobos is also derived.

Juvenilization (paedomorphosis) may account only for few aspects of the bonobo cranium. If the differences at hand are just a matter of delayed development, we should expect overlapping ontogenetic trajectories that only differ in timing. But most of the differences in neurocranial shape between chimpanzees and bonobos are a result of divergent developmental trajectories, rather than just maturational delay (Lieberman, 2011; Lieberman et al., 2007; Mitteroecker, Gunz, Bernhard, Schaefer, & Bookstein, 2004; Mitteroecker, Gunz, & Bookstein, 2005; Williams, Godfrey, & Sutherland, 2002). This suggests that the bonobo skull follows a different developmental path, which partially contradicts scenarios such as the self-domestication hypothesis (Berge & Penin, 2004; Mitteroecker et al., 2004; Penin, Berge, & Baylac, 2002; Ponce de Leon & Zollikofer, 2001; Williams, Godfrey, & Sutherland, 2003; Zelditch, Sheets, & Fink, 2000). In other words, if differences in neurocranial shape between chimpanzees and bonobos are not just a matter of maturational delay, then at least some of the features of the proposed model could be due to common ancestry.

Similarly, not all shape differences between chimpanzees and bonobos can be attributed to paedomorphosis. For example, in chimpanzees and bonobos, significant expansions of the areas corresponding to the frontal lobes of the cerebrum occur at two different ontogenetic stages, but they are not explained through maturational delay. These expansions seem to be associated with other changes in skull morphology, e.g., a rotation of the orbits toward the midline, which in turn is closely related to alterations in other endocranial regions (Durrleman, Pennec, Trouve, Ayache, & Braga, 2012).

Pure developmental change in the timing of events (heterochrony) can explain neither human craniofacial morphology nor the differences among the African apes. The human skull is already markedly different at birth when compared to other great apes in shape space and size-shape space (Mitteroecker et al., 2004). This suggests that differences in craniofacial morphology between *Homo* and *Pan* are caused by small genetic differences affecting early ontogeny.[19] This makes problematic to assume that human skull morphology is the result of global heterochrony.

Subsequent multivariate analyses have rejected the hypothesis of global heterochrony in the cranium of *Pan* as well as regional heterochrony for the lower face, the upper face, and the neurocranium (Mitteroecker et al., 2005). Paedomorphosis can explain only a modest proportion of the variation between chimpanzees and bonobos

---

[19] In fact, Mitteroecker and colleagues' (2004) study gives some support to the old idea that bonobos come closer in their overall ontogenetic pattern to the great ape ancestor (Ciochon, 1983; Zihlman, 1984; Zihlman et al., 1978; Zihlman & Lowenstein, 1983).

(Lieberman, 2011; Lieberman et al., 2007). One region of the bonobo skull may have evolved relative to its ancestral form through one pattern of heterochrony, while another component of the skull might have evolved via a different heterochronic pattern, or by a completely different mechanism (e.g., by the repositioning of regions, by evolutionary novelty, or by the addition of bone and other tissues). In humans, for instance, the braincase seem to be an example of peramorphosis (a heterochronic process in which individuals of a species mature past adulthood), whereas the face appears to be an example of paedomorphosis (Lieberman, 2011; Shea, 1989). Important modifications in the cranial base in the hominin lineage are also connected to the evolution of bipedalism rather than paedomorphosis (Kimbel, Suwa, Asfaw, Rak, & White, 2014). As a consequence, skull morphology in bonobos and humans is not explained in its entirety by paedomorphosis and, therefore, neither is the behavioral and neurobiological changes associated with domestication.

To a large extent, this hypothesis depends on the idea that self-domestication leads to the retention of juvenile features of the skull. If paedomorphosis explains only some of the similarities between bonobos and humans, then this would bear on the hypothesis itself. It is still possible, though, that some important aspects of the behavior of bonobos and humans such as the retention of play behavior into adulthood might be explained through some version of the self-domestication hypothesis since this feature might be fundamentally mediated by hormonal mechanisms. However, domestication typically leads to a syndrome of morphological, behavioral, and neurobiological changes that crucially includes changes in cranial morphology (Wilkins, Wrangham, & Fitch, 2014).

3.7. Normative guidance

The model I have defended in this chapter has important philosophical consequences for descriptive theories of ethics. For it gives us a different picture of the evolution and nature of our capacity for normative guidance (Kitcher, 1998, 2006, 2011), i.e., our capacity to grasp norms and to make normative judgments. Kitcher is not the only one implicitly relying on some form of the demonic male view (see, for instance, Boehm, 2012; Bowles, 2008, 2009; Bowles & Gintis, 2011; Dubreuil, 2010a; Prinz, 2007). However, he is a primary example of this view in the context of an evolutionary approach to normative cognition that is close to mine—his account, for instance, does not aim to unify human moral thinking as a natural kind in psychology, but rather to provide an account of the cognitive capacities that enables our distinctive human norm psychology.

In section 3.1, I suggested that Kitcher's account of the emergence of the capacity for normative guidance is a particular form of the demonic male view. Kitcher's evolutionary scenario relies on a chimpanzee-like social environment where dominance and aggression are the key driving forces behind human evolution. On Kitcher's account, dominant alpha males punish anyone who disrupts the established social order, and this makes normative guidance, at least initially, psychologically grounded in fear. In addition, as in the demonic male view, the evolution of our capacity for normative guidance is in part the story of the gradual expansion of top-down mechanisms of control (in the form of some sensitivity to commands) over our less reliable emotional nature. If an agent is able to understand the normative structure of its chimpanzee-like social environment, that agent will be able to avoid the costs imposed by aggressive alpha males. The motivational force to obey these commands comes for free in this case since they help the agent to avoid situations in which the anticipated consequences are feared or disliked.

The above model, then, bears important consequences for Kitcher's view of normative guidance and its function. For, according to him, normative guidance has to be more explicit, more a matter of offline cognition. But the model of early hominins I presented in section 3.3 strongly suggests that neither the demonic male view nor Kitcher's (2011) account of our capacity for normative guidance are plausible. On this model, the social world of our last common ancestor is not male-dominated (viii), their social organization is less hierarchical (vii) and social cohesion is less regulated by aggression and fear of punishment (vi). On the contrary, this ancestor is characterized by its enhanced emotional control (ii), increased aversion against aggression (iii), empathy and positive emotions (iv), and enhanced perspective taking capacities (v). If this is correct, normative guidance would not have been the result of selection for avoiding punishment by very aggressive and authoritative alpha males. Kitcher's vindicating genealogy becomes murky. The tendencies of some individuals to monopolize resources and to impose social order through aggression would have been largely regulated in our lineage through more bottom-up affective processes—e.g., because perceived cues of distress would inhibit violent behavior. No sensitivity to commands is required. No norms are invoked. Another explanation would be necessary.

Kitcher's evolutionary account of normative guidance is not the only available explanation. It is also not the best. I think a better explanation of the shift toward normative guidance could be framed in terms of shared intentionality (Tomasello, 2015; Tomasello & Carpenter, 2007; Tomasello et al., 2005). Shared intentionality seems to account for much of the distinctive features of human psychology. It has been argued, for instance, that such capacity is responsible for the appearance of joint attention, cooperative communication, imitative learning, and teaching, which are at the basis of cultural learning and the social norms and traditions we see in human culture (Call, 2009; Tomasello, 2014). Although joint activities and behavioral traditions are common among great apes, humans substantially differ from other apes in their underlying psychological mechanisms. Chimpanzees and bonobos can attribute some psychological states such as perceptions and goals to others (Tomasello, Call, & Hare, 2003), but they are neither intrinsically motivated to share those psychological states nor are able to represent these mental states in a joint, collective fashion (Call, 2009).

Shared intentionality helps us to explain much of our distinctive cooperative tendencies within the ape lineage (Tomasello, 2009b). In this view, ecological triggers forced a shift from a very individualistic lifestyle to more collaborative one, characterized by cooperative breeding and increasingly complex forms of collaborative foraging (Hawkes, 2014; Tomasello et al., 2012). But the expansion of our collaborative capacities was only possible by the emergence of the cognitive and motivational infrastructure for sharing intentional states such as attention, goals, beliefs, and desires. For example, by the time of the emergence of *Homo*, females have to forage, leaving their offspring to the care of other group members. Under the custody of strangers, children needed to engage the caregiver's attention for its own sake (Hawkes, 2012; Tomasello & Gonzalez-Cabrera, 2017). This engagement required not only simultaneously attending to the same objects and events but also some visual referencing such as a quick look to the caregiver's face or eyes to track the caregiver's focus of attention (Carpenter & Call, 2013; see also Bakeman & Adamson, 1984). They would have needed to see whether the caregiver was also looking to them, and then to express affective contact with the adult (Hobson & Hobson, 2007).

The phenomenon of shared intentionality is also known to be closely related to the emergence of normative thinking (see, for instance, Göckeritz et al., 2014; Rakoczy, 2008; Rakoczy, Brosche, Warneken, & Tomasello, 2009; Rakoczy et al., 2008; Schmidt et al., 2011a; Schmidt & Tomasello, 2012; Tomasello, 2015). As we shall see in chapter 5, developmental studies show that at a very young age children are able to understand social norms such as the ones that govern conventional games (Rakoczy, 2008; Rakoczy, Brosche, et al., 2009; Schmidt & Tomasello, 2012). Children strongly object to transgressions of norms of conventional games by using normative language, e.g., “No! It does not go like this!”, or by intervening directly on the transgressor in order to correct him, e.g., “Take that one” or “Not this way!” (Rakoczy et al., 2008).

As explained in the previous chapter, shared intentionality gives us a framework that allows singling out a particular class of normative judgments, which is characterized by a distinctive gradient of generalizability, intrinsic motivation, and punitive attitudes. The cognitive component of joint intentional states allows to represent judgments such as “We ought to do *x*” in ways that are increasingly general and abstract regarding their plural subject. The motivational component in turn allows us to entertain normative thoughts whose motivational component is shared, e.g., the motivation to comply with the normative judgment or to admonish those who do not. As we will see in chapter 4, this form of normative thinking required the interaction of a distinctively human line of development for sharing psychological states with an ancient developmental trajectory of increased social tolerance and enhanced mindreading skills.

Other primates do form social expectations but they lack the capacity to form normative expectations that are socially shared (von Rohr, Burkart, & van Schaik, 2011). Normative expectations depend for their emergence and maintenance on shared acceptance and commitment. Joint goals of the form “We intend to do *x*” have some normative weight. If we intend to drink water from the pond, we should go to the pond. These shared mental states require at least minimal commitment. For the mutual interlock of motivational states is one of the features that make these shared intentional states different from a mere accidental overlap of intentions—e.g., “I want to go drink water from the pond” and “You want to drink water from the pond”. Certainly, this normative weight becomes stronger and more evident the more is at stake for the parties involved. The reason why these states underwent this increase in normative force is evolutionary, for as I will argue in chapter 4 the increasing demands on cooperative activities such as foraging required commitment and social enforcement to guarantee their success. In modern humans, if someone unexpectedly abandons the joint activities that these states bring about, others may demand an explanation and censure their partner (see, for instance, Warneken et al., 2006; Warneken, Gräfenhain, & Tomasello, 2012; Warneken, Hare, Melis, Hanus, & Tomasello, 2007). Even mostly prelinguistic children exhibit glimpses of this normative dimension when they try to reengage an unresponsive partner in a collaborative activity (Warneken et al., 2006; Warneken et al., 2012; Warneken et al., 2007). The attempts of the child to reengage collaborative partners seem to be attempts to restore a shared activity, and in this way, they could well be an incipient form enforcement.

Given that much of the empirical work on this psychological phenomenon comes from the comparative literature, the theory of shared intentionality offers a helpful framework to put normative guidance within an evolutionary context. This capacity, for instance, is thought to be closely linked to the selective pressures resulting from cooperative activities such as cooperative breeding and collaborative foraging. The former is often considered a previous step for the full emergence of shared intentionality (Hawkes, 2012, 2014; Tomasello et al., 2012) because, although cooperative breeding leads to greater prosocial skills, it does not entail in itself higher cognition (Burkart et al., 2014; Burkart, Hrdy, & van Schaik, 2009; Burkart & van Schaik, 2010). For this reason, it has been argued that the selective pressures of collaborative foraging, which are more cognitively demanding in terms of coordination, would explain the emergence of the type of complex cognition underlying shared intentionality, starting with *Homo erectus* and continuing with *Homo heidelbergensis* (Tomasello et al., 2012). Since it is only with the emergence of collaborative foraging that we can fully explain the emergence of shared intentionality, it would be only until then that we could expect social norms to emerge—i.e., understanding them as mutually known expectations bearing social force and enforced by third parties.

In sum, following the suggestion put forward in chapter 2, I think that social normative thinking can be understood as a special case of shared intentionality. In the next chapter, I will explain in more detail the connection between these two aspects of human cognition by arguing that our capacity for normative guidance was selected to avoid disappointing a relationship partner's expectations in a more tolerant social environment when hominins became more interdependent foragers (Tomasello et al., 2012). Norms would then be represented by the agent as shared expectations about how individuals ought to behave in a given situation, i.e., they would be represented as joint intentional states. These expectations were necessary to carry out tasks that required complex coordination such as collaborative foraging and more so to build the kind of collective cultural institutions that are the distinctive feature of behaviorally modern humans.

3.8. Conclusion

I have begun this chapter explaining the philosophical motivations behind debates about the reconstruction of early hominins, and the perils of reconstructing the social behavior of these ancestors. Then, I have argued that we are more justified in using an alternative model of the *Pan*/*Homo* LCA based on a mosaic hypothesis of human origins as we are in adopting the chimpanzee referential model. According to the model, our last common ancestor would have been a large mosaic of traits seen not only in both *Pan* species but also in other primate species. In order to support this model, I have gathered evidence from a wide range of disciplines, especially paleoanthropology and comparative psychology. I have particularly focused on comparative studies between humans and both *Pan* species, as well as discussed whether these similarities are homologies or analogies. I have argued that this alternative model does not support popular views of human evolution such as demonic male view (Wrangham & Peterson, 1996) or the killer ape hypothesis (Dart, 1953). Finally, I have argued that this alternative model suggests a different evolutionary account of our distinctive capacity for normative guidance (Kitcher, 1998, 2006, 2011).

Chapter 4. Big game hunting and the evolution of human social norm psychology[20]

In chapter 3, I have argued that our capacity for normative guidance may be linked to the lineage of our shared intentional capacities. Now, in this chapter, I will argue that normative guidance evolved as a consequence of the selective pressures of collaborative hunting. This is a hypothesis I have not ruled out in chapter 3 since the model I proposed there was still compatible with the idea that early hominins used tools and hunt in groups—a model in line with views that decouple hominin hunting from aggression (see Pickering, 2013). The tamer picture of our early ancestors I offered in that chapter will play a crucial role here in two different ways. One is to provide a cognitive baseline for my lineage explanation of normative guidance. Another is facilitating the gradual expansion of early hominin cooperative capacities since collaborative foraging would be implausible in a very aggressive and dominant lineage.

Wolves, hyenas, and lions are carnivores with dominance hierarchies, but their hunting strategies were not readily accessible to the hominin lineage since we lack the speed, the canines, the claws, and the physical advantage of these predatory mammals. They hunt in groups as some great apes do but their hunting behavior is not collaborative. Collaborative hunting as the one we see among hunter-gatherers requires shared intentionality, for it relies on joint goals and plans, task and role division, and shared commitment, which are psychological features much more readily accessible to the tamer model of early hominins I have put forward in chapter 3. It is sometimes informally argued that hunting behavior such as the one seen in social carnivores involves some of these capacities. However, these claims are based on anecdotic evidence rather than behavioral data collected under controlled conditions. No evidence today seems to indicate that non-human animals possess real shared intentional capacities. More importantly, great apes with whom we share a common lineage systematically fail tests of shared intentionality (Call, 2009; Carpenter & Call, 2013; Tomasello & Carpenter, 2007; Tomasello et al., 2005). There are phylogenetic reasons that make hunting in great apes much more relevant for comparative purposes. I will explain how a deficit in shared intentionality captures the difference between the hunting behavior of chimpanzees and that of hominin hunters. More generally, my aim

---

[20] Sections 4.4 and 4.5 of this chapter are based on worked published in Tomasello and Gonzalez-Cabrera (forthcoming) and has been modified to meet university guidelines.

in this chapter is to propose a lineage explanation whose starting point is an early hominin with high levels of social tolerance and basic mindreading skills, and as an endpoint, agents with human-like capacities for shared intentionality, including its subsidiary normative dimension, i.e., the distinctive punitive attitudes they engender and their consequences for partner choice.

However, I think that collaborative hunting is only part of the story of the evolution of shared intentionality. For this capacity emerges early in ontogeny while collaborative hunting is typically an adult activity. This poses the nontrivial question of how shared intentionality was pushed down into early developmental stages. In some cases, once these collaborative capacities become central to the lives of adults, it is not too difficult to provide a story about how these skills could have been co-opted for secondary functions early in ontogeny. For example, Sterelny's (2012a) apprentice learning model suggests that the demands of social learning integrate juveniles into the economic lives of adults as a key aspect of their education, i.e., they learn by supervised doing. But it is less obvious, for instance, how basic shared intentional capacities such as joint attention and pointing are learned skills or how they became so entrenched in early development. For this reason, I will argue in this chapter that the initial driving forces in the evolution of these basic forms of shared intentionality were the selective pressures derived from a cooperative breeding niche while its distinctive normative dimension emerged as a result of the selective pressures on collaborative hunting.

The concept of shared intentionality has roots that are both philosophical (Bratman, 1992; Gilbert, 1989; Searle, 1990; Sellars, 1963) and psychological (Tomasello & Carpenter, 2007; Tomasello et al., 2005). Roughly speaking, shared intentional states are hybrid mental states. I understand these states as the systematic co-occurrence of a cognitive and a motivational component. They have a cognitive component, e.g., the representation of shared goals, the awareness of other peoples' intentional states, and the representation of others' actions, as well as an intrinsic motivational component—i.e., these states possess motivational force which is partially explained by the shared nature of those goals and intentions.[21] It is true that one of these

---

[21] Often, we are motivated to do something because other people are also motivated to do it. However, it is psychologically possible (and very likely indeed) to be motivated to do something for different reasons. If we intend to go to the pub, we genuinely share this intention only if you are intrinsically motivated to do it because I am also joining, and vice versa. Yet this does not mean that one may not have other reasons to go to the pub. Perhaps I also want to go to the pub to drink a beer. Thus, if you cancel our appointment for some reason, I may still go to the pub because I am still individually motivated to do so.

components may appear without the other, both in development and in evolution. But it is the regular co-occurrence and the role this co-occurrence plays in driving human behavior that is relevant for a mental state to be a shared intentional state. Since this is an essential feature of shared intentional states, this hybrid nature is also an essential feature of the normative judgments that our shared intentional capacities are able to engender as a special subclass. Depending on the specifics of this mixture of components, some normative states would appear more belief-like while others would seem more desire-like.

On the view I propose, basic forms of shared intentionality such as joint attention and pointing behavior emerged as ontogenetic adaptations for early childhood in a cooperative breeding context. These capacities were subsequently extended in ontogeny into adulthood because they were co-opted for foraging, especially in the form of collective hunting. Collaborative foraging in early *Homo* required increased motivation and commitment to shared goals compared to cooperative breeding. Foraging is riskier. It requires more time and energy, is dangerous, and has a higher cost of opportunity. As a consequence, these shared intentional states were carefully monitored and enforced by all the group members. The coupling of these punitive attitudes with some other common features of shared intentional states was the origin of social normative thinking. Thus, shared intentionality helps us to identify a well-defined target for the psychological study of normative thinking. For at least some of these shared intentional states regulate collective activities in a way that is distinctively normative—they are mental states that are not only generalizable in scope and intrinsically motivational but also able to engender punitive attitudes, for they are enforced in oneself and others (for a similar proposal, see Sripada & Stich, 2007).

Moreover, as I have suggested in chapter 2, within this broad category of normative thinking, we can isolate different kinds of prototypical moral judgments that latch onto particular psychological processes. For example, in virtue of the cognitive and motivational profile of these mental states, we can single out prototypical moral judgments about harm, justice, or rights, which possess the signature moral pattern described by social domain theorists (Nucci & Turiel, 1978; Nucci et al., 1983; Turiel, 1983, 1998). Similarly, we can account for a particular class of prototypical moral demands, which are seen by the agent as inescapable and authority independent (Joyce, 2001; Mackie, 1977; Mameli, 2013), depending on how robust the motivations to comply and enforce compliance are across different counterfactual scenarios. I will develop this view in more detail in chapter 5.

The rest of the chapter is organized as follow. In section 4.1, I will briefly explain the relation between big game hunting and my lineage explanation of normative thinking. In section 4.2, I will introduce Tomasello and colleagues' (2012) account of the origins of shared intentionality, which identifies collaborative foraging as the main driver in the evolution of our shared intentional capacities. I will argue that this explanation leaves open the question of why these capacities emerge so early in ontogeny. In section 4.3, I will defend a view according to which changes in our rearing environments rather than challenges in adult social interaction were the responsible for the early emergence of this capacity. In section 4.4, I will then explain what I think was the decisive contribution of collaborative hunting of big game. Finally, in section 4.5, I will put forward a model of the role of ontogeny in the evolution of human normative cognition that takes into account the tradeoffs that emerge from the selective pressures associated to collaborative foraging and those stemming from cooperative breeding environments.

4.1. Big game hunting

Part of the legacy of the killer ape hypothesis and the demonic male view I have discussed in the previous chapter is that hunting (in particular, big game hunting) played a crucial role in the evolution of human social cognition, including moral cognition (see, for instance, Boehm, 2008, 2012; Mameli, 2013; see also Gibbard 1990, pp. 66-67, 139-140). Admittedly, hunting also plays a central role in Tomasello and colleagues' (2012) 'interdependence hypothesis' of the evolution of shared intentionality—a capacity that is central to my lineage explanation of normative cognition. But recent accounts of the evolution of hunting have begun to emphasize the role of impulse control in ambush hunting while others, focused on pursuit and endurance running, highlight the importance of collaboration in such forms of hunting (see, for instance, Bramble & Lieberman, 2004; Henrich, 2016; Lieberman, Bramble, Raichlen, & Shea, 2009).

I also think that some version of this idea is correct. For I believe that the transition toward cooperative foraging contributed to the evolution of shared intentionality and, as a consequence, to the emergence of human normative thinking. But the evolutionary account I will propose here differs from the usual evolutionary account of shared intentionality in that it emphasizes the role of cooperative breeding in shaping the basic infrastructure of normative thinking. Moreover, the evolutionary account that I will offer focuses more on the early expansion of human cooperative foraging in general rather than on big game hunting in particular. Although there are reasons to think that big game hunting indeed played a decisive role in shaping our human-unique capacities of normative thinking, I will argue in this chapter that its specific role can only be fully understood within a much broader framework of human cooperation and sociality. For the evolution of big game hunting needs to be framed in the context of an already quite tolerant lineage as the one offered in the previous chapter. Big game hunting was important. But in order to understand its emergence we need to decouple aggression and hostile propensities from human hunting prowess (Pickering, 2013). Moreover, although I will argue that cooperative breeding facilitated the early emergence of our shared intentionality capacities, their evolution was certainly much more gradual and complex than any model exclusively focused on cooperative breeding or collective foraging would predict. For the basic social capacities that emerged in the context of cooperative breeding were also co-opted to facilitate social and cooperative in low-risk contexts such as play. In this context of peer interaction, children could learn many of the social, cognitive, and coordination skills that facilitate obligate cooperative foraging. The stronger the demands of these skills among adult foragers, the more maturation and preparation they require, which could have led to the early developmental assimilation of more complex capacities for shared intentionality.

In other accounts of human cognition, normative thinking emerged significantly later, e.g., after 100 kya as a consequence of a shift to a mixed economy (see, for instance, Sterelny, 2012b). But these accounts typically focus on the emergence of explicit social norms rather than the motivational and cognitive machinery that support our capacity of normative guidance. Although the first signs of behaviorally modern humans around 100 kya (Bar-Yosef Mayer, Vandermeersch, & Bar-Yosef, 2009; Henshilwood & d'Errico, 2011; Henshilwood, d'Errico, & Watts, 2009; Schwarcz et al., 1988) suggest an expansion of norm-guided behavior, I will argue that the basic cognitive capacities for normative guidance appeared much earlier in our lineage—resembling the relatively early emergence of cooperative communication with respect to later arrival and diversification of natural languages (for a discussion about the parallel between language and normative cognition, see chapter 5).

The above discussion bears important consequences for the overall argument of this thesis. Since I shall argue that shared intentionality leads to a robust form of normative psychology, I will be in a position to explain the connection between distinctive, easily identifiable clusters of prototypical moral judgments and the psychological literature on shared intentionality. But the connection between moral judgments and shared intentionality is likely just prototypical and fragmentary (see chapter 2). They are not explained by robust inheritance of biological mechanisms, as is so in the case of shared intentionality, but rather by ontogeny through channels of cultural transmission.

The unifying features of moral psychology as I see them are local, historical, and culturally contingent rather than the result of robust and domain-specific internal mental processes. While there are reasons to think that shared intentionality and normative guidance are generatively entrenched since other aspects of human cognition depend on the reliable emergence of these capacities (see Wimsatt, 1999), there is no similar rationale in the case of moral judgments. For as we have seen in chapter 2, even though normative cognition might be a unified domain, people seem to make the moral/conventional distinction in different ways (see Nucci & Turiel, 1978; Nucci et al., 1983; Turiel, 1983, 1998), depending on particular cultural, socioeconomic, and developmental factors. As a result, selection would favor plasticity and increased sensitivity to the relevant sociocultural variables instead of developmental entrenchment or genetic assimilation, i.e., when selection causes plastic phenotypes to lose their environmental sensitivity over evolutionary time (Levis & Pfennig, 2016).

In this view, shared intentionality lead to a particular form of norm psychology that later ramifies in various types of prototypical moral judgments which depend on specific cultural contexts. Very likely, the kind of norm psychology I am defending in this dissertation has branches that extend into other subclasses of social norms such as religious norms, but I will specifically focus on in this thesis in the ramification of this account for moral cognition. In particular, I will argue that this form of normative psychology shares a common genealogy that stems from our capacity for shared intentionality. A lineage explanation of this shared intentional psychology provides a unified target for norm psychology due to the prominent role that shared intentionality played in the hominin expansion of our cooperative capacities. We not only can identify the distinctive motivational and cognitive mechanisms of shared intentionality but also link them to the particular selective pressures that shaped human cognition (see figure 4.1).

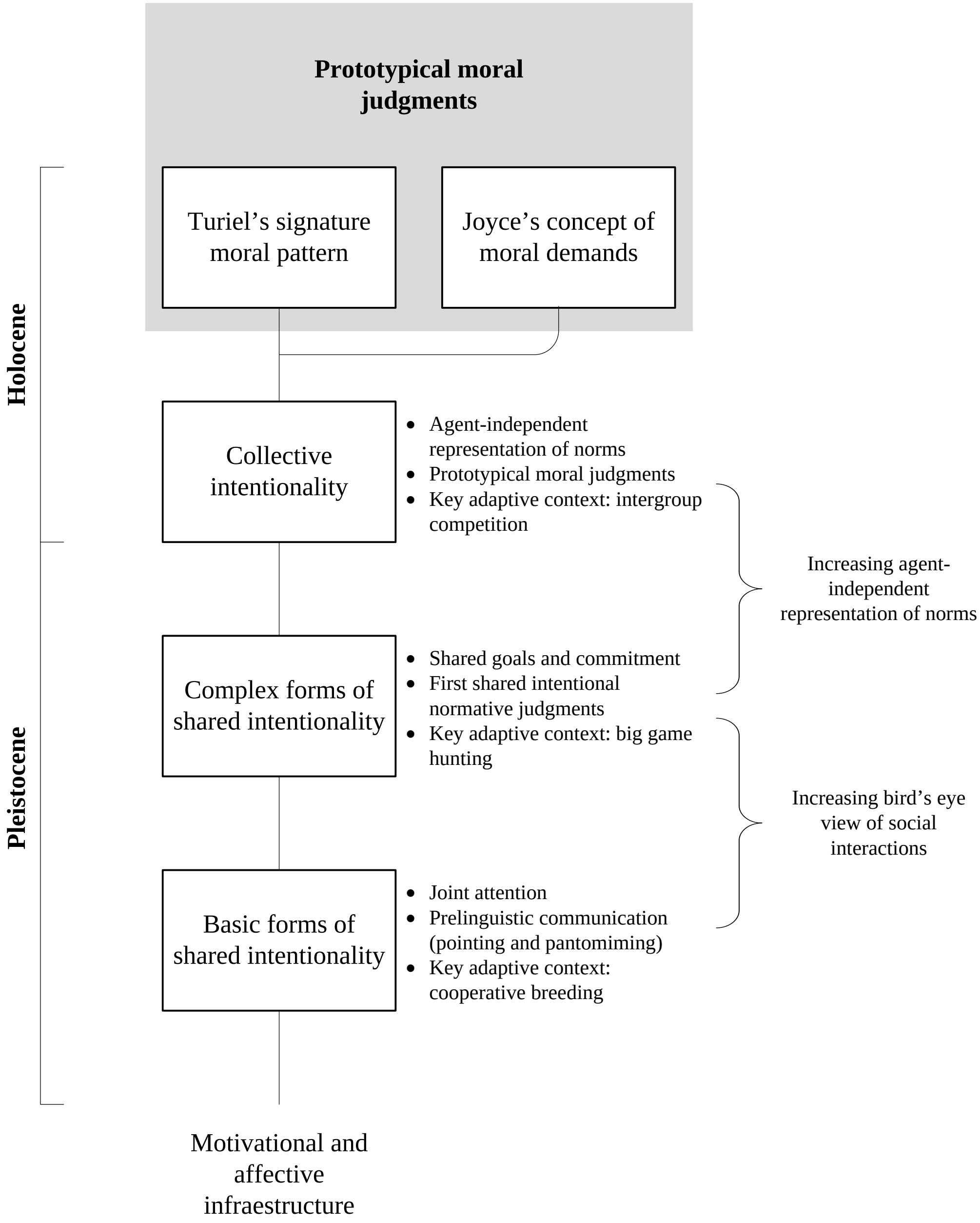

Figure 4.1. Sequence of stages in the lineage of shared intentionality. Cooperative breeding lead to basic forms of shared intentionality, which allows an incipient bird's-eye view of social interactions. These capacities are enhanced due to the selective pressures of collaborative hunting, generating ever more complex forms of shared intentionality—including our capacity to make shared intentional normative judgments. Finally, growing intergroup competition selects for increasingly agent-independent representation of norms, which ultimately lead to the emergence of prototypical moral judgments.

4.2. The interdependence hypothesis

Evolutionary accounts of shared intentionality are important for my lineage explanation of normative thinking because they can incorporate both the motivational and cognitive components of this form of cognition. They not only provide an evolutionary explanation of the motivational mechanisms of norm compliance, but it also gives us an account of the distinctive mode of representation of these social standards. The evolutionary origins of shared intentionality, however, have also been traditionally linked to hominin hunting. Tomasello and colleagues (Tomasello, 2014; Tomasello et al., 2012; Tomasello & Vaish, 2013), for instance, have argued that the key step in the evolution of our distinctive shared intentional psychology was the emergence of obligate collaborative foraging, especially collective hunting. For in the context of collective hunting the scale and the cost of failure of joint activities escalated, which required not only adjustments in our motivational capacities but also changes in our cognitive machinery to represent more complex forms of social interactions. In this section, I will briefly explain the evolutionary trajectory of shared intentionality as seen by supporters of the interdependence hypothesis. Although somewhat differing in their assessment of the role of big game hunting in the evolution of our social-cognitive capacities, the evolutionary account of shared intentional normative thinking that I will offer in this chapter will heavily build upon this model.

Tomasello and colleagues (2012) frame this idea in terms the interdependence hypothesis—the idea that throughout its evolutionary history the hominin lifestyle increasingly relied on mutualistic collaboration for survival and reproduction. In particular, the hypothesis postulates a two-step process in the evolution of human social cognition and cooperation:

> Step 1: Humans gradually turn more interdependent with one another in the context of foraging. Unlike any other primate, they become obligate collaborative foragers. In this context, they develop an increased interest in the well-being of their partners, but also new capacities for collaboration such as shared intentional capacities and cheating avoidance mechanisms not possessed by other great apes. As a result, humans start to display a form of normative cognition in discrete interactions with particular individuals, i.e., when we are engaged in and, therefore, directly affected by those interactions (Darwall, 2006).

Step 2: As human groups became larger, our shared intentional capacities scaled as a function of competition with other human groups. At this stage, the understanding of our social environment extends to the groups themselves (instead of just particular individuals) and activities are represented in an agent-independent way. Shared intentionality transforms into collective intentionality. For example, joint goals in simple dyadic interactions scale up into collective group goals for large-scale activities among unfamiliar individuals, including warfare, the construction of ceremonial complexes, and trade. As a result, humans start creating, following, and enforcing group-wide social norms, conventions, and institutions, which are used as ethnic markers.

Importantly, the timing of the shift between these two stages is not completely clear. Tomasello and others do not say much regarding this issue. But there is little evidence for an increase in focal band size much before the Holocene, but around 10 kya, at the Pleistocene-Holocene transition there is evidence for increased meta-group complexity, increasingly close relations between groups, and the formation of ethnolinguistic complexes (Flannery & Marcus, 2012; Seabright, 2004; Sterelny, 2013). As a result, there are reasons to suspect that the transition from one evolutionary step to the other was relatively recent.

Tomasello and colleagues (Tomasello, 2015; Tomasello & Vaish, 2013; Vaish & Tomasello, 2013) also apply these steps to the ontogeny and the phylogeny of moral thinking. However, most of the empirical work carried out by Tomasello and colleagues focus on social interactions that are normatively structured, rather than on morality itself. For example, studies about children’s awareness of the normative structure of conventional games and their understanding of social norms (Rakoczy, 2008; Rakoczy, Brosche, et al., 2009; Schmidt & Tomasello, 2012). Certainly, there are reasons to think that there is a substantial overlap between this type of normative cognition in general and moral thinking in particular since some other studies look at children’s normative understanding of collaboration, helping, sharing, or fairness, which are prototypical moral situations (Tomasello & Vaish, 2013; Vaish, Carpenter, & Tomasello, 2010; Vaish, Missana, & Tomasello, 2011).

Comparative analysis of cooperation in both great apes and humans support the idea that these sequences of stages roughly describe the evolution of human cooperation and social cognition. This view is supported by comparative analysis between great apes and humans (see technical box 4.1). For example, human cooperation is notably different from the type of cooperation we see in other great apes, for humans but not apes have the capacity to represent join intentions (Fletcher et al., 2012; Tomasello & Carpenter, 2005; Warneken et al., 2006) and cooperative communication (Duguid, Wyman, Bullinger, Herfurth-Majstorovic, & Tomasello, 2014; Melis et al., 2009; Warneken et al., 2006). Despite differences in sharing behavior between both *Pan* species (Hare & Kwetuenda, 2010; Hare et al., 2007; Tan & Hare, 2013), humans are the only great ape that reliably share the spoils of collaborative efforts (Bullinger, Melis, & Tomasello, 2011; Hamann, Warneken, Greenberg, & Tomasello, 2011; Melis, Hare, & Tomasello, 2006b; Rekers, Haun, & Tomasello, 2011; Warneken, Lohse, Melis, & Tomasello, 2011). Our species is also the only one that exerts extensive partner choice and partner control (Engelmann, Herrmann, & Tomasello, 2012; Melis, Altrichter, & Tomasello, 2013; Melis, Hare, & Tomasello, 2008; Melis, Schneider, & Tomasello, 2011; Warneken & Tomasello, 2013). Moreover, we are the only species that seems to coordinate action through joint commitments (Gräfenhain et al., 2009; Gräfenhain, Carpenter, & Tomasello, 2013; Greenberg, Hamann, Warneken, & Tomasello, 2010; Hamann, Warneken, & Tomasello, 2012).

Technical box 4.2. Comparative analysis of cooperation in great apes and humans

Experimental evidence suggests that chimpanzees and bonobos are motivated to engage in triadic activities and social games if they require little coordination—even spontaneously with a human partner (MacLean & Hare, 2013; Pika & Zuberbühler, 2008). Human children, however, go a step further by coordinating their plans of action to allow role reversal (Pika & Zuberbühler, 2008, p. 210). Unlike great apes, human children encourage their playmates through nonverbal communicative gestures even when their playmates are not necessary for achieving the goal of the game. They also distinguish between a partner who is unwilling to play and a partner who is unable to continue, even though the behavioral outcome of the partner is the same in both conditions (Warneken et al., 2012). By 3 years of age, human children seem to understand joint commitments and either make an effort to honor those commitments or apologize for breaking them (Gräfenhain et al., 2009). Although Tan and Hare (2013) found that bonobos voluntarily share food with a recipient even when they could have monopolized it, only human children are more likely to divide up their rewards equally if they obtained these rewards by working collaboratively than by working individually (Hamann et al., 2011). For example, Schmidt & Sommerville (2011) showed that 15-month-old infants expect resources to be distributed equally among recipients, and Geraci and Surian (2011) further showed that when 16-month-olds see one distributor being fair toward a recipient and another distributor being unfair toward the same recipient, they expect the recipient to approach the equal distributor. In addition, at around 5 and 6 years of age children are able to anticipate being judged and then behave in a way that increases positive evaluations of themselves. Engelmann et al. (2012) found that, unlike chimpanzees, human children steal less from an imaginary peer recipient, and tended to help that recipient more, if a peer was observing them. In other words, only humans seem to have a basic concern for their own reputation such that they manage the impression they are making on others by avoiding having norms applied to them.

On Tomasello and colleagues' view (Tomasello, 2014, 2015; Tomasello et al., 2012; Tomasello & Vaish, 2013), these differences in cooperation likely evolved in the context of collaborative foraging due to the heavy cognitive demands it imposes in terms of coordination, distribution of resources, and commitment. Great apes do not forage collectively in the sense of Tomasello and colleagues since they do not need it to survive and, consequently, they never evolved the appropriate motivational and cognitive machinery to do so. This can be well exemplified by stag hunt situations. In a simple stag hunt interaction with free partner choice, two individuals each have the option of pursuing either a hare or a stag. The former is a low-risk but low-quality resource while the latter is a riskier but highly valuable one. If one chooses to hunt a valuable stag, then one must have the cooperation of the partner in order to succeed when one could have got a hare by oneself. But chimpanzees do not rely on hunting to survive. They go after the valuable stag only if the risks are low and manageable without complex coordination. It has been shown, for instance, that hunting rates decline as prey encounter rates decline and the chimpanzees have to spend more time and energy finding situations where the prospects of success are good (Watts, 2012). Moreover, comparative studies have shown that in a stag hunt game situation, when risks are low (e.g., because the hare is of low value) and information is cheap (e.g., because the partner's behavior is observable), both chimpanzees and human children are able to successfully coordinate on the higher value stag around 90% of the time. If the risks increase and observing the partner is more difficult, in contrast, chimpanzees are less successful, whereas children compensate through communication (Duguid et al., 2014).

According to the interdependence hypothesis, the expansion of grassland ecosystems in Africa due to global climatic change and the resulting fragmentation of forest at the end of the Pliocene (Bobe & Behrensmeyer, 2004) might have led to a significant radiation of terrestrial monkeys (Jablonski & Leakey, 2008), who would have competed with humans over terrestrial resources. As a consequence, it is argued that humans began to rely on food resources that required complex collaboration and coordination for their procurement—e.g., collaborative hunting of large game and collaborative gathering of embedded plants. This type of foraging required the coordinated efforts of multiple individuals and their special knowledge and skills (Hill, 2002; Kaplan, Hill, Lancaster, & Hurtado, 2000). Under those conditions, individuals with increased skills for collaborative foraging would have had an adaptive advantage. Thus, humans gradually became obligate cooperative foragers.

Chimpanzees do hunt in groups (Boesch, 1994; Boesch & Boesch, 1989), and even bonobos have been reported to hunt monkeys in the wild (Surbeck et al., 2009; Surbeck & Hohmann, 2008). But collaborative hunting of the type that Tomasello and colleagues have in mind requires planned coordination, rather than just emergent coordination (Knoblich, Butterfill, & Sebanz, 2011). Planned coordination requires agents to share goals, and to be committed to them. Although it is difficult to know when ancestral hominin foragers began to exhibit this capacity, representing one’s and one partner’s role, often in interchangeable ways, should have been also critical at some point. We know, for instance, that extant hunter-gatherers have a capacity for complementary role taking and role reversal. If targeted hunting emerged around 1.7 mya as some researchers have argued (Bunn, 2007; Bunn & Pickering, 2010), this would mean that hunting would have been a planned rather than a mere opportunistic activity. Emergent coordination, in contrast, is fundamentally a consequence of perception-action couplings that make individuals act in similar ways and be responsive to each other, regardless of joint plans.

Paleoanthropological evidence also indicates that facultative bipedalism, which allowed our ancestor to walk for long distances but not to run, emerged very early in our lineage (Lovejoy, Latimer, et al., 2009; White, Asfaw, et al., 2009). In fact, our early ancestors had none of the adaptations to vertical climbing, forelimb suspension, and knuckle-walking (Lovejoy, Latimer, et al., 2009; Lovejoy, Simpson, White, Asfaw, & Suwa, 2009). This would have precluded both looking for protection by quickly climbing to the trees or the type of hunting that we sometimes see in chimpanzees. Thus, given the local predatory fauna (Hart & Sussman, 2005; White, Ambrose, et al., 2009), foraging alone would have been a very dangerous task. So foraging and scavenging in a group for safety likely came first, creating the conditions for the gradual evolution of more planned and coordinated foraging. Collaborative foraging would have been not only necessary in our lineage but also a much more complex form of planned coordination that went beyond mere emergent coordination (Knoblich et al., 2011).

When our ancestors had to forage collaboratively with others or starve, they have to face new cognitive challenges, at the same time that old selective pressures in the social domain increased. As these foragers have to rely more on foraging stag-like resources for sustenance, they not only had to deal with coordinating collective action at a larger scale but also they had to develop a bulletproof commitment with the hunting group. Hunting is risky and sometimes deadly. But it was likely necessary for sustenance. Consequently, our hominin ancestors would have had to be focused and committed to their collective deeds, even under adverse or detrimental conditions, or at the risk of having joined a potentially untrustworthy or mistrustful partner. They have had to choose partners wisely and being noted as a good potential partner too. They would have had also to share the resources in a way that minimizes conflicts and stabilizes cooperation over time—not only provisioning mating partners and offspring but also sharing with the unrelated group members involved in the hunt. Certainly, historically known foragers are not groups of close relatives and chimpanzee and bonobo groups are not groups of close kin either, but collaborative foraging might have begun in extended family groups of early hominins. Reproductive cooperation, which I will discuss in section 4.7, could have been important in this process. For once fathers are able to recognize their offspring, there is more scope for kin-mediated cooperation as a partial scaffold for early cooperation.

This type of environment favored the evolution of particular adaptations for social cognition such as sharing the goal of the activity among several individuals and representing interchangeable roles in that context (Tomasello, 2014). It would have also strengthened other capacities already existing in our lineage, e.g., skills for partner choice and control (Darwall, 2006). Moreover, collective hunting would have amplified our sharing tendencies in order to motivate and stabilize cooperation. This form of foraging, then, should have significantly enhanced our empathic capacities and highlighted our capacity to care about the well-being of our partners. For when collective hunting turned obligate, individuals became increasingly interdependent with one another.

Sometimes, advocates of the interdependence hypothesis frame the first step of their model in terms of dyadic interactions. But there is no reason to assume that all social interactions were indeed dyadic. For example, according to the interdependence hypothesis, shared intentionality evolved in the context of collaborative foraging, but there is no reason to think that collaborative foraging was actually dyadic. There is an important difference between the number of people engaged in a task and the type of psychology that they deploy to carry out that task. As I understand their proposal here, advocates of this view think that the psychology of collaborative foraging was adapted to dyadic interactions, rather than arguing that those interactions were carried out by actual dyads. As I will argue in the following section, even if this interpretation is incorrect, foraging activities such as big game hunting would have needed a form of shared intentional psychology adapted to more collective social interactions. They require a we-mode of representation that enables us to see collective interactions in an increasingly agent-independent way as explained in chapter 2. They were not conceived as interactions about you and me but instead as interactions about a more abstract plural subject, which is typically verbalized using the plural pronoun ‘we’.

The interdependence hypothesis gives us few details about the transition from the first to the second step of the model. But the basic dynamic of the transition is relatively simple and explanatorily valuable. In the view of Tomasello and others (2012; see also Sterelny, 2014), the capacity for joint action through shared intentionally increased as groups become larger than the relatively modest groups of early hunter-gatherers. As human groups became larger and more frequent in their encounters with other groups, they also started to face increased intergroup competition. Certainly, it is difficult to determine when groups began to grow in size or when conflicts between them started to increase, but they did over time. As a result, selective pressures derived from intergroup competition could have made humans increasingly interdependent across a wide range of collaborative activities—a process that perhaps took place very late during the early Holocene transition to complex hierarchical societies (Sterelny, 2016). But as social groups became larger, the challenges also became greater. Individuals had greater chances to free ride in those groups. Encounters with strangers and one-shot interactions became more common, so keeping track of previous interactions with a number of partners also became more difficult.

In this type of highly cooperative societies with a tribal structure comprising smaller bands, interdependence was crucial but much less face-to-face. The emergence of more complex forms of division of labor made the grounds of our practical demands more complicated to understand. If, say, two partners are hunting a stag, roles are more interchangeable and easy to grasp to each other. But division of labor as the one we see, for instance, in the late Pleistocene (and even more in the Holocene) requires a more abstract understanding of what is normatively required and of the reasons for those obligations. If in the past our ancestors were interdependent with particular others for foraging, well into the Holocene they became interdependent with everyone within the group as a whole—even though we probably do not know most of the people of the group and we do not understand their role in the group. In these larger groups, social interactions relied more on some form of group membership dynamic structured by social norms, rather than on our personal histories of past interactions with particular others. Increasing intergroup competition would have made also crucial to recognize one's group members. In those contexts, one needs to know who is a trustworthy ally as much as one need to be recognized as one. This made it essential to conform to the ways of the group in different social dimensions such as behavior, speech, and appearance. They signal relevant social information such as ethnic identity and status that is costly and difficult to acquire. Learning to speak Māori as a native is hard to fake for someone who has not been born into the culture. It signals to insiders that I am one of them at the same time that it makes me an outsider to those coming from a different ethnolinguistic background. Therefore, in larger groups, when cooperation with strangers became more frequent and intergroup conflict more common, social norms could have helped to identify reliable social partners.

Social norms also helped to foster coordination in larger groups by making the group ways part of the agents' common ground, i.e., the knowledge, beliefs, and assumptions that agents share at a certain point which facilitate collective action because they take them for granted and not subject to (further) discussion. Coordinating complex tasks and distributing the social benefit produced by those activities was a far more complex problem than distributing the meat within a group of hunter-gatherers. Moreover, to the extent that normative thinking is a form of shared intentionality, people would have been not only motivated to follow social norms. They would have been also very stable cooperators, relatively resilient to free-riding. Social norms would have been enforced on all the group members, including oneself, through negative emotions such as feelings of guilt and shame (see Boehm, 2012; Mameli, 2013).

I think human psychology was indeed transformed by this transition but changes in normative cognition were differences of degree rather than kind. Norms became more explicit and more agent-independent but they exploited an ancient sensitivity to (less explicit and less agent-independent) norms. This likely facilitated social changes of doubtful moral standing, e.g., the rise of slavery and ideologically-driven inequality (Sterelny, 2012b). This phenomenon could have occurred well after the Pleistocene-Holocene transition. Kim Hill, for instance, has suggested that in extant forager metapopulations, members of those groups do all know one another, perhaps as a consequence of the fluidity of band membership. He reports that foragers have extensive knowledge of their immediate band as well as those in their local groups. Members of the Ache metaband community seem not to have problems tracking group members of an around thousand members, despite considerable spatial dispersion (Hill, 2012).

As a result, it might be that the history of human normative cognition is mostly described by the type of simple normative cognition described in the first step in Tomasello and colleagues' model, or perhaps through some transitional form between the two stages (see technical box 4.2). Some of these social standards would have been at least initially local, tribal norms and customs, but the type of normative cognition in place during the Holocene was much more collective and abstract—a distant relative of the type of normative cognition we would expect to see in ancestral hunter-gatherers. One could speculate, for instance, that once collective hunting was complex enough to require norms about the roles of the activity and the fair distribution of the social product emerged, they could have extended toward more complex forms of social organization as the one seen in forager metapopulations, and then even more toward more institutionalized forms of cooperation. Sex and reproduction would be a case in point. Once groups became interdependent, managing conflict became even more important, and sex is an obvious conflict flashpoint. Practices such as bride-price and other forms of regulated mate exchange, including institutionalized marriage, would have been very important. Overall, as the groups grew larger in size due to an influx of people from nearby locations, and their boundaries became more abstract, and social interactions more complex and ill-defined, normative thinking became much more abstract, universal, and agent-independent.[22]

---

[22] It might be important to point out here that making judgments from an agent-independent point of view is not the same as using those judgments to set ourselves apart from others. A common reaction to encounters with diversity is that human groups become more insular, reinforcing group boundaries and seeking to preserve their own cultural identity while demonizing others. However, even when a person

Technical box 4.3. Group-mindedness and the Pleistocene-Holocene transition

The second step in Tomasello and colleagues' (2012) might be much more complex than assumed by the model. On the one hand, the kind of group-mindedness that it presupposes seems to be closely related to the emergence of symbolic material culture as witnessed in the archaeological record, perhaps around 75 kya (Henshilwood & Dubreuil, 2011). One reason is that evidence of symbolic behavior is usually interpreted as a form of signaling one's group membership or individual status. On the other hand, the later emergence of farming and sedentary life at the Pleistocene-Holocene transition seems to be more closely connected to the social and demographic changes stressed by the model, for only then competition between groups intensifies and cooperating with strangers becomes a real issue (Seabright, 2004; but see Bowles, 2009; Kelly, 2005). I take these steps to be idealized transitions built upon the motivational and cognitive machinery of the first step, which scale up with demographic and social complexity. On this view, although less abstract and stable, the basic machinery for group-mindedness was already present in place with the emergence of shared intentional states such as shared goals and intentions, which gradually extended to ever larger groups.

In particular, the challenges of a cooperative life in larger groups required a capacity to understand and represent abstract social norms, e.g., norms that apply to everyone in the group regardless of their role or identity. It would have required a motivation to comply even at the cost of one's personal interests, and a capacity to enforce those standards. In other words, this later step would have led to some form of agent-independent normative cognition insofar as humans start to follow and enforce abstract, group-wide social norms that apply to everyone more or less equally. Of course, these capacities are fallible, which is compatible with the levels of squabbling and conflict documented in the archeological and ethnographic record (Allen & Jones, 2014; Bowles, 2008, 2009; Fry & Soderberg, 2013; Lahr et al., 2016). Yet they would have kept conflict sufficiently under control for cooperation to occur in increasingly larger groups.

---

thinks that the people from the neighboring village are barbarians because they eat a certain taboo food, for instance, that person is still making an abstract and agent-independent judgment, namely that it does not matter whether one is a member of the relevant social group or not, eating that taboo food is wrong.

4.3. Did meat make us moral?

I will focus in this section on the first step of Tomasello and colleagues' (2012) evolutionary scenario since I think the distinctive normative dimension of some special class of shared intentional states emerged as a consequence of the selective pressures on collaborative hunting. On the view I am defending, collaborative foraging and hunting made us able to entertain joint mental states that are characterized by a distinctive gradient of generalizability, intrinsic motivation, and punitive attitudes. I take these mental states to be normative because they give rise to a certain form of social rationality. This form of rationality generates social expectations about what one should do in a given situation as well as expectations about what others should do to achieve a shared goal (e.g., when collaborating in hunting) and because we enforce them in ourselves and others (e.g., when excluding partners who do not meet these expectations).

This is not a minor feature of Tomasello and colleagues' evolutionary account. For example, another prominent view about the role of hunting in normative cognition such as Christopher Boehm's 'conscience-evolution hypothesis' (Boehm, 2008, 2012; Mameli, 2013; see also Gibbard 1990, pp. 66-67, 139-140) focuses on the affective and motivational mechanisms underlying norm compliance. However, his account does not explain the type of social expectations that these mechanisms are supposed to regulate. For him, other species such as domestic dogs and chimpanzees can be good learners of rules, though they lack the appropriate affective machinery that characterizes human conscience (see, for instance, Boehm, 2012, pp. 20-23). In contrast, I think that there are important differences in the way we represent norms and social standards, for we represent them as joint mental states.

It can be argued that social expectations may trigger punitive attitudes towards those who do not meet those expectations without these social expectations being truly normative. For it seems possible for an agent to have this kind of social expectations without thinking that others have an obligation to meet them. But not all normative mental states are judgments about normative obligations, just as not all moral judgments concern demands (see chapter 2). Moreover, there is a sense in which collective hunting made us agents who can entertain interpersonal normative obligations. As I will explain in the next few sections, hominin hunters equipped with short-range weapons were likely able to understand what they were rationally required to do in order to achieve their individual goals. The crucial contribution of shared intentionality to normative cognition was not to insufflate mental states with some primitive *virtus normativa* but to provide the cognitive and motivational infrastructure to understand what they were collectively required to do in order to achieve joint goals in the context of collective hunting.[23]

This is not only a view about normative guidance within a cohesive small group of people, where everybody takes one another to share the same expectations of themselves and others. It can also give us an account of the role of norms in intergroup interactions by helping us to draw and bridge the social boundaries between groups. For social norms can vary in their scope. When a norm applies to a particular group of agents, one has to represent the norm in a way that is less agent-independent and more closely linked to the particulars of individuals, their social roles, and the groups they live in. If one knows the rules of *Pitz* (a traditional Mesoamerican ballgame often held as a ritual event featuring human sacrifice), one applies its rules only to those who play *Pitz*. A thought like "We all think that those who are playing *Pitz* should do thus and so", for instance, could restrict the scope of the plural subject to those who are acquainted with the game. Another role is to bridge the social boundaries by enabling normative judgments that one applies across groups. For example, when one thinks that human sacrifice is morally reprehensible no matter one's cultural identity or fondness for pre-Columbian sports. Importantly, the shared intentionality framework is able to capture these negative judgments since they explain why we tend to share these normative expectations with others aiming them to join our normative standards. Of course, this does not mean that those normative mental states are always successfully shared, but rather that these are mental states that one is intrinsically motivated to share with others. Successful or not, the cognitive and motivational mechanisms deployed are the same—i.e., the mechanisms that support sharing intentional states.[24]

As we will see in the next sections, this account of the representation of social norms allows explaining different motivational and cognitive aspects of normative judgments. For example, the fact that some of these normative judgments are represented as being objective demands because they are represented as shared beliefs

---

[23] I am grateful to Philip Pettit for pointing me to this Molieresque metaphor, which I use here in a different sense than originally intended (Pettit, 1990).

[24] Notice that this claim is weaker than saying that normative judgments just are shared intentional states. As explained in chapter 1, shared intentional normative thoughts are a special subclass of normative thoughts, which means that not all normative thoughts are shared intentional normative thoughts.

about aspects of the social world that any agent can mutually attend and could agree to. However, as I will develop in more detail in chapter 5, much of this account of normative judgments will not focus on the perceived objectivity of normative demands, but rather on a different set of features such as their perceived inescapability and authority independence—features that only partially overlap with the idea of objective demands.

4.3.1. Social expectations

Hunting in great apes usually lacks visible organization or coordination, but there seem to be some exceptions. For example, in the Taï Forest, monkeys are extremely agile and the canopy is continuous (Boesch & Boesch-Achermann, 2000; Boesch & Boesch, 1989). As a consequence, the chances of a successful hunting are very small without any coordinated behavior. However, even in those cases, coordination is emergent rather than planned (Knoblich et al., 2011). Chimpanzees surround the prey by taking the most opportune spatial position in order to maximize their chances at catching the monkey. They are mutually responsive to one another’s spatial position, similar to the hunting behavior we see in wolves and lions, but there is no agreement on a joint goal or real division of labor involved (Cheney & Seyfarth, 1990; Tomasello & Call, 1997). There is joint action but no shared intention.

Generally speaking, joint actions can be defined as any form of social interaction where two or more individuals coordinate their actions to bring about a certain change in the environment (Sebanz, Bekkering, & Knoblich, 2006). Carrying a table, playing a piano duet, creating and maintaining social institutions are typical examples of joint and collective actions in culturally modern humans. Individual actions are cognitively different from joint actions when sharing goals and intentional states are essential for understanding coordinate behavior. Coordinated behavior in the context of joint actions can be either emergent, as I take hunting in chimpanzees from the Taï Forest to be, or planned, as an in modern humans (Knoblich et al., 2011).

In emergent coordination, joint action emerges through rather automatic mechanisms that couple perception and action, such that they cause individuals to act congruently. For example, a group of people lying on the grass in a park may quickly start running to the nearest bandstand when it suddenly starts to rain (Searle, 1990). Each person has the intention of running to the bandstand looking for shelter, but these intentions are individual rather than joint—similar to the type of collective behavior that is triggered when chimpanzees chase monkeys in the Taï Forest. Coordination would be emergent, i.e., it would arise from similar motor routines and cues (Marsh, Richardson, & Schmidt, 2009; Spivey, 2007) and it would not require any joint goal or plan.[25]

However, coordinated behavior may also be driven by shared representations of goals and tasks. A group of actors in a play could perform exactly the same action as in the above example, but the cognitive mechanisms necessary to perform the action would be significantly different. In the play, the actors would have a shared or joint goal of the form "We intend to do *x*", but in the former case, each person would have an individual intention of the form "I intend to do *x*". This is why they are shared intentional states (Searle, 1995, 1998; Tuomela, 2003). These intentions require some representation of other agents' mental states—although the extent to which other agents' mental states are co-represented may vary from case to case (Vesper, Butterfill, Knoblich, & Sebanz, 2010). Full-blown joint actions, as described here, require a joint goal. But not all planned coordination is a real joint action because planned coordination does not require a joint goal of the form "We intend to do *x*".

In a minimal sense, planned coordination requires specifying the outcome of a coordinated action, one's role in the activity, and some awareness that the outcome can only be brought about with the support of another agent (Knoblich et al., 2011). However, representing the outcome of a coordinated action is not the same as representing that outcome as a joint goal. The individual goals of two agents may overlap, the agents may coordinate their actions through a careful representation of the partner's role in the activity, but the action would not be truly joint. Since both agents have individual goals, once the goal of one of them is achieved, the activity stops regardless of the other agent's intentions. Partners want to achieve their individual goals. Perhaps these goals can only be achieved with other peoples' help. But the fact that those others were able to achieve their goals is not part of the agent's own goal—the fact that my choice of action is counterfactually sensitive to the actions of others does not make it a joint goal.

In this line of thought, for instance, Warneken, Gräfenhain, and Tomasello (2012) have argued that, in order to test whether social activities are structured by joint goals, non-human apes must encourage their partners when they stop playing their role in the activity even when those partners are not necessary for achieving the individual

---

[25] This does not mean, however, that mechanisms of emergent coordination are not important to planned coordination. Various sources of emergent coordination identified in the literature have proven to be key facilitators of joint action (Knoblich et al., 2011).

goal of the activity. Subjects also have to be able to distinguish between partners who are unwilling to play from partners who are unable to do it, even if the partner's behavioral outcome is the same—otherwise, subjects may be just viewing their collaborative partners as mindless social tools. Certainly, chimpanzees and bonobos have been shown to be motivated to spontaneously participate in joint activities with humans (see MacLean & Hare, 2013), but it is still not clear whether they coordinate their actions through joint goals.

In addition, there is also no evidence that any other non-human ape coordinates their actions via shared task representations. These representations are control structures that allow agents to engage flexibly in joint actions by specifying in advance the individual roles in the activity. They govern key cognitive processes that enable coordination such as monitoring and predicting (Knoblich & Jordan, 2003; Pacherie & Dokic, 2006). But as Melis and her colleagues (Melis et al., 2006b) have shown, although chimpanzees seem to understand the instrumental value of others to achieve certain goal (since they tend to select those partners who are more useful in experimental settings), they do not need to have detailed representations of the conspecific's actions. For example, they may understand that they can only get the food with the help of a conspecific, and they may also represent the goal of obtaining food by pulling a rope in an experimental apparatus. But they do not need any complex shared task representations. In this case, chimpanzees only show a capacity for minimally planned coordination. Similarly, for instance, experiments involving social games in chimpanzees and bonobos (MacLean & Hare, 2013; Pika & Zuberbühler, 2008) have shown that neither chimpanzees nor bonobos seem to able to coordinate their plans of action to allow role reversal (Pika & Zuberbühler, 2008). If they were able to do so, this would be evidence of shared task representations in *Pan*.

Thus, the central idea here is that shared intentional states provide the kind of cognitive sophistication that allow humans to represent social expectations that apply reciprocally to other individuals or the group as a whole. This is not a simple predictive expectation such as the one we may have when we expect rain on a cloudy day. Great apes may have expectations about their physical environment or expectations about what other agents could do in a given situation (Krupenye, Kano, Hirata, Call, & Tomasello, 2016). When we share the goal of a collective activity, this generates expectations about what we and other group members should do in order to achieve that goal. The idea that these expectations are normative is not new. The semantics of deontic modals has been linked, for instance, to the Bayesian notion of expectation

(Cariani, 2016). If we have a shared goal, it might be crucial for us to share those expectations too—especially, when those activities are complex or risky. Real collective activities such as the type of interactions involved in collective defense, scavenging, or collective hunting very likely required representing social expectations about what other individuals could do in a given situation in order to achieve a joint goal. Hominin hunting was more complex than chimpanzee hunting. Hominins are only able to coordinate the kind of collective action that leads to reliable and safe hunting success, including the distribution of the social product, if they are able to share their shared goals and to form stable normative expectations regarding the actions of their partners. Hominin cooperation required shared intentionality because agents had to be able to count on each other, as opposed to merely calculate with each other.

But even significantly less complex activities might have also contributed to the emergence of the shared intentionality syndrome. In simple cases of joint attention, for instance, expectations about what a partner can see or is not able to see are formed. These expectations are monitored in order to reestablish common ground between both individuals when those expectations are broken. Importantly, the motivational component of these mental states behaves like a gradient. For example, in simple cases of joint attentional engagement common ground is built but the motivation to maintain it might not be robust, e.g., one of the partners can easily lose interest in the shared scene. In other cases, however, both components should have to be stable enough to go hand in hand. This could have been so in many socioecological scenarios such as collective defense and scavenging, but I will try to argue that this was clearly the case at least in the context of collective hunting of big game. For hunting in forager social worlds seems to be a source of pride and achievement of prestige (Hawkes, O’Connell, & Jones, 2001), which ameliorates defection problems. Hunters like hunting, even though this activity may be risky.

In sum, collective hunting of big game arguably scaled up the cognitive demands on our shared intentional capacities since at some point it likely required more planning and division of labor. This generated complex expectations about others that required extensive monitoring in order to adjust behavior efficiently to meet the group’s shared expectations. At the same time, collective hunting of big game very likely imposed significant failure costs when others did not act as expected. In those cases, losing motivation for the activity and distraction were costly. Thus, as we will see in sections 4.3.2 and 4.3.3, agents likely had to couple those expectations with a robust motivation to participate in the activity and to correct and punish others when they did not do their part. Collective hunting was also motivationally demanding.

4.3.2. Social rationality

Even more important, shared intentionality gives rise to a form of social normative thinking. Purely egocentric I-intentions involve a form of instrumental rationality, which may be considered normative. Instrumental rationality is linked to basic normative mental states to the extent that this form of rationality is an important component of practical rationality. The former requires from the agent a capacity to deliberate and adopt suitable means to his/her ends, while the latter requires from the agent a capacity to deliberate about what the agent ought to do given the agent's circumstances, goals, and available information. For this reason, the relation between practical rationality and deontic modality has received increasing attention in fields such as formal semantics (Cariani, 2016; Charlow, 2016; Lassiter, 2016). These I-intentions can even support certain forms of cooperation when cooperation is prudential. But shared intentionality is not quite like instrumental or prudential rationality. They are not just about my own ends as an agent. Their normative demands are not exactly about the instrumental relation between an agent and his or her goals. They have a collective, social flavor. They are not instrumentally normative in that sense, but rather socially normative. Actions are required not because I intend to get meat for myself, but rather because we are intending to hunt.

As we have seen in section 4.2, chimpanzees cooperate to hunt monkeys in the Taï Forest, although the best interpretation of that cooperative activity is that each individual is trying to get the prey by himself. For they are more successful together than they would be trying to get the prey alone. The result is a collective activity. But there are reasons to question whether chimpanzees have an intrinsic motivation to cooperate. Chimpanzee hunting behavior is opportunistic rather than intrinsically motivated—they participate in the hunt if the prospects are good and only because they do not rely on it for survival. For example, cooperative hunting is much less prominent than individual hunting in habitats with fragmented forests where hunting is more difficult (Stanford, 1998a) and experimental evidence consistently suggests that in foraging collaborative tasks they treat partners primarily as social tools (Call, 2009).

It is true that humans do not always cooperate. But we engage in relatively complex cooperative activities for their own sake even if an individual effort would have sufficed in other cases. For example, Gräfenhain and colleagues (2009) have shown that at a very young age, children not only act jointly with others in pursuit of an instrumental goal but also seem to be highly motivated to engage in joint activities just for the sake of acting together with another person. In play situations that could be played either alone or jointly with a partner, children as young as two years of age often reengage their adult partners even if they could easily achieve the desired outcome (e.g., getting access to an enclosed toy) by themselves. Moreover, around the third year of age, children begin to reengage their partner more often when the partner expresses his/her intention to share the activity by inviting the child to play the game together with the experimenter.

If the model of the baseline of social-cognitive capacities of early hominins I have offered in the previous chapter is correct, much of the motivational prerequisites of shared intentionality would have been at least readily evolvable from the genetic and development package inherited by the *Pan*/*Homo* LCA. But shared intentional states are not only intrinsically motivational since they also involve a degree of generalization and abstraction that they go beyond the egocentric representation of the agent’s goals and desires. Minimally, they require a significant other and to merge (metaphorically speaking) our own intentional states with those of our partner. As we will soon see, social interactions between infants and caregivers in the context of cooperative breeding were like that in a rather incipient way. Interactions between partners for collective hunting required a much broader social perspective.

I think joint intentions inherited this form of instrumental normativity and made it social. Collective big game hunting operates under this logic. It is a form of social or collective rationality (Tomasello, 2015). Likely, this source of normativity required a capacity for metacognition that enabled the monitoring of our own mental states—in particular, the representation of goals and beliefs about the potential transitions between different states of the environment that could lead to the desirable outcomes. It required, say, assessing my beliefs and desires to make instrumental decisions about what is the best course of action to achieve a certain goal and quickly updating those states when something does not go as planned. Just updating one’s beliefs and acting on them would not be enough since we would also need to monitor one’s states of certainty and uncertainty, seek new information in conditions of uncertainty, and recognize whether one has access to the information needed to solve the problem at hand. But by the time hominins were able to hunt big game and making stone tools, this is a plausible assumption. It might be true that metacognition is not present in all the primate order

(Basile, Hampton, Suomi, & Murray, 2009; Beran, Smith, Coutinho, Couchman, & Boomer, 2009; Fujita, 2009; Paukner, Anderson, & Fujita, 2006), but macaques and apes seem to have some metacognitive capacities (Call, 2010; Call & Carpenter, 2001; Couchman, Coutinho, Beran, & Smith, 2010; Hampton, 2001; Washburn, Gulledge, Beran, & Smith, 2010). If this is true, the crucial difference between humans and other non-human species regarding metacognitive capacities is not just a matter degree but also of kind, for we would be the only species able to track our own shared, we-mode intentional states.

A capacity for shared intentionality also significantly changed the affective dispositions that boost motivation and commitment to cooperative endeavors. In the model of early hominin social cognition offered in chapter 3 much of these dispositions were readily evolvable from the *Pan*/*Homo* LCA. But this baseline of social-cognitive capacities significantly changed with the emergence of shared intentionality. Joint intentions took the empathy-related infrastructure of our lineage and catalyzed sympathy. Sympathy is a concern for the well-being of others. It requires perspective-taking. But if perspective-taking were just a matter of great ape mindreading, it would be just paternalism. A mother, for instance, may assume that is up to her what is good for the younger because she takes herself to a have an influence on the youngster. If I genuinely join your intention to do something, my motives to help involve recognizing your perspective on your needs. When I join your intention to perform a certain action, my helping is at least in part a consequence of the fact that you are intending to perform that action, you struggle to succeed alone, and we can make of it a rewarding collaborative effort.

Joint intentions can also involve trust and commitment. Even in very simple joint tasks, 14- and 18-month-old infants try to reengage the partner through communicative gestures (Warneken et al., 2006; Warneken et al., 2007). 3-year-old children are more likely to attempt to reengage partners when they have explicitly accepted the invitation of an adult to play than when the adult just follows the child activity (Gräfenhain et al., 2009). Children’s understanding of joint commitments is complex, e.g., they are more likely to wait for their partners when they are unable to continue, more likely to spontaneously help them when needed, or to take over their roles if necessary (Gräfenhain et al., 2013). This type of commitment should have been especially important when joining these activities was costly and risky like in big game hunting, as we have seen in section 4.2 (Duguid et al., 2014).

Arguably, collective hunting and defense against predation relied on relatively sophisticated forms of communication due to the heavy demands on coordination of these activities. Big game hunters have to think ahead how to target potentially dangerous prey. It requires mastering a number of complementary tasks and, in some cases, task specialization. Hunting partners have to be committed to these plans and tasks. Often, though, they would have had to take high-stakes decisions under time pressure and partial information. Under these circumstances, communication would have been difficult, and the capacity to understand and meet partners' expectations would have been increasingly important. Successful hunting requires meeting the demands of these practical commitments. This does not mean that joint commitments create unbreakable bonds between agents, but rather that cooperation becomes more stable and, consequently, more likely to be successful, because the commitment to the task goes now both ways. Early hunters likely understood their mutual dependence in a cognitive way. They were able to think not only about what they needed to do to succeed but also about what they, as a group, had to do in order to pull off collective enterprises.[26]

Sharing the goal of hunting also put individuals on a roughly equal footing, i.e., they came to value their partners differently. For example, activities structured around joint goals require from the individuals involved staying engaged in the activity until partners get their share from the activity. If goals are shared and the goal is getting the meat, we should share the spoils at the end. Of course, successful collaboration and successful establishment and respect for regularities of division of the product will coexist with failures to coordinate and squabbles about division. Even lethal conflict is frequent in forager societies. However, although noisy and partially successful, our shared intentional psychology should have made collaboration much more robust and stable under these noisy conditions than mere rational deliberation about individual intentions.

---

[26] One reason why coercive scavenging preceded large game hunting is that it required a basic form of collective action—a halfway house between the demands of chimpanzee group hunting and those of planned ambush hunting. This form of scavenging does not require much role division or teamwork. When compared to hunting in chimpanzees, it requires only a bit more planning, e.g., deliberately carrying weapons and stones, as well as more commitment since it was potentially more dangerous. However, in this situation, agents primarily need to react on the fly, so mutual adjustment through mechanisms of emergent coordination as the ones seen in chimpanzees would have been good enough.

Sharing the meat was at least in part a matter of sympathy or concern for the partners' needs, and the increasing pattern in social tolerance that characterized our lineage, as explained in the model of the hominin baseline of social-cognitive capacities introduced in chapter 3. In this model, early hominins were characterized by increased cooperative and sharing tendencies beyond the hypothesized levels of a chimpanzee-like model of that ancestor. Nonetheless, sharing was also a matter of social rationality. Hunters could have operated to a certain degree under the logic of "I want meat", but reliable cooperation and coordination would have required from them to act under the plural logic of "We want meat". This shift would have been facilitated not only by the fact that successful hunting delivers meat in large packages so the costs of sharing were likely low but also by the early adaptations for peaceful sociality seen in early hominins which reduce conflict when distributing the spoils and consequently incentivizes cooperation. This plural logic would have meant that collaborative partners were not seen just as mere instruments to achieve our individual goals. Because it is also part of the joint goal that hunting partners will get some of the food, and all of them should be committed to precisely that goal. Chimpanzees usually abandon collaborative activities once they get their reward (Greenberg et al., 2010). But children often assist unlucky partners so that both get the reward in collaborative situations, even when there is nothing left to gain for them (Hamann et al., 2012).

Yet, away from the heat of the hunt, collaboration and sharing were likely also a matter of individual rationality. Hunters, and especially big game hunters, should have realized the instrumental importance of others in the activity. They would have had a clear idea of the negative social consequences of monopolizing the meat all for themselves; they should be aware that, in small groups, they rely on a few potential partners. Not sharing the spoils, would undermine cooperative partnerships that are essential. This was probably explicit rather than implicit knowledge—behavior mediated by prudential reasons that are related to the agent's interests. Cooperative ventures of this kind are fragile and often give no time for doubts and regrets, so there was likely a very narrow window of opportunity to learn these behaviors through, say, associative learning.[27] On top of that, shared intentional capacities would have allowed

---

[27] As we will see in section 4.4, cooperative breeding and its consequences for human life history would have played an important role in easing the learning trajectory of many essential cooperative skills. For children and juveniles would have had lots of practice in cooperation and cooperation failure. In particular, peer interaction during middle childhood could have played the role of preparatory step for

hunters to share mutual perspectives about the activity. They would have known that in order to maintain those relationships, cooperative partners have to be nourished and highly motivated. Otherwise, for everyone, the costs are high and the dangers are real.

However, even if hunting partners were able to understand sharing as a prudential course of action, it would have been an open problem to determine what type of distribution will keep partners cooperating. Rational maximization of returns is cognitively demanding since variables such as differential contributions to success and issues related to temporal discounting would be hugely important but difficult to assess. These complexities increase if interactions are repeated over time and the number of hunting partners is large. Also, even if distributing the meat was a computationally tractable problem, e.g., through the use of heuristics, the process would have been socially disruptive, unless the other group members would be expecting the same result. Heuristics about distributing the meat work well when they are common ground. Thus, these representations about how to distribute the social product should have been shared intentional states, i.e., common expectations that we share with others so that any group member can quickly update them according to circumstances.

4.3.3. Punitive attitudes

The account of the representation of social norms via our shared intentional infrastructure also helps to understand certain peculiar motivational aspects that are typically linked to normative judgments about the transgression of social norms. In section 4.3.1, I have explained that sharing mental states such as goals and plans generate shared expectations. In section 4.3.2, I have argued that sharing these expectations took the kind of instrumental normative thinking characteristic of great apes and transformed it into a form of social normative thinking about what one think we should do in the context of collective activities. In this section, I will show how failing to meet these shared expectations may trigger some form of punitive and corrective attitudes as the ones introduced in chapter 2.

I agree that, in the context of ancestral big game hunting, what counted as a fair distribution was a matter of joint regulation of meat distribution (Boehm, 1999, 2012). Protests about undeserved shares would be social feedback to tune standards of

---

more serious collaboration during adulthood since in this developmental period children can still rely on parents and other adults for protection and subsistence.

distribution when individuals are jointly motivated to share the meat in ways that encourage cooperation. With shared intentionality comes a way to understand the group standards of food distribution as a we-mode social expectation. Expectations about the division of spoils need to be congruent to make cooperation stable and to minimize conflict, but they are not just private expectations that turn out to overlap and coincide with those of others. These expectations have to be shared in terms of both their cognitive content and their associated motivational component.

When a group member feels mistreated, others have to be sensitive to this fact such that they can quickly keep conflict within bounds. Although certainly possible, it is not clear that once I got my expected share I will care for other agents' violated expectations for mere instrumental reasons—for that, agents would need to be cool-headed rational maximizers in typically heated situations that demand from them to react with anger or, at least, contempt. In contrast, if they shared those standards, they would be mutually committed to making others comply with them. A group member that does not share meat with another not only would be the target of the punitive attitudes of the excluded individual but the group as a whole—since they would be in a joint mental state of the form "We are sharing the meat". Not sharing the meat with someone is now everybody's business. This would have led not only to some cognitive representation about what would have counted as a fair distribution within the group but also to forms of mutual vigilance and regulation of the joint commitment to those standards.

Individuals would have also chosen their partners more carefully. Some models of the evolution of impartiality and fairness, for instance, see these traits as adaptations to environments in which individuals compete to be chosen and recruited in mutually advantageous cooperative interactions (Baumard, 2016; Baumard, André, & Sperber, 2013). This is very plausible in cases where hunting was supported by the use of projectile weaponry and perhaps even in relatively large groups of ambush hunters using short-range weapons. With the increased communicational skills that big game hunting presupposes, foraging partners would have cared about their reputation. Partners would react angrily against unresponsive partners, signaling what they perceive as an unfair treatment or as an intolerable mistake. These angry reactions would have engendered punitive attitudes that, in extreme cases, would have triggered direct physical aggressions. A group of hunters should have been sensible to these signals and reacted with guilt, shame, or remorse. Social selection in favor of angry-types would be expected in this situation since the willingness to punish others even in a one-shot games appears to maintain high levels of cooperation. If the members of a group are able to identify angry-type punishers, then angry punishment would have likely coevolved with cooperative preferences (Bruner, n.d.).

Guilt would have brought also individual fitness benefits. This could have been so by preventing transgression and by leading to reparative behaviors as some models on the evolution of apology seem to indicate (O’Connor, n.d.). This is a plausible conjecture. Although it is difficult to project the behavior of extant humans into earlier human evolution, there are reasons to think that guilt affects prosocial behavior in ourselves and others. Anticipated guilt in humans reduces the likelihood of social transgression (Svensson, Weerman, Pauwels, Bruinsma, & Bernasco, 2013), and increases prosocial behavior (Ketelaar & Au, 2003; Malti & Krettenauer, 2013; Regan, 1971), e.g., cooperation in social bargaining games (de Hooge, Zeelenberg, & Breugelmans, 2007; Nelissen, Dijker, & deVries, 2007). Guilt also leads to reparation, acceptance of punishment, and self-punishment (Nelissen & Zeelenberg, 2009; Ohtsubo & Watanabe, 2009; Silfver, 2007). Moreover, expressing guilt and remorse also reliably lead to a reduction of punishment (Eisenberg, Garvey, & Wells; Fischbacher & Utikal, 2013; Gold & Weiner, 2000).

As a result, affective reactions like anger and guilt would have been sensitive to the mode in which social expectations are represented. We-mode representations of these expectations are socially normative and increasingly more agent-independent—they are not just expectation about what a particular individual should do in a given situation, but about what anyone should do in a similar situation. These affective reactions would become then normative emotions, i.e., emotions that support social normative standards of behavior. Briefly, there would be joint representations of these standards, which in turn would be supported by affective processes that work as joint commitment devices.

4.4. Cooperative breeding

One problem with the traditional evolutionary account of shared intentionality is that it does not explain why many of these capacities emerge so early in ontogeny. This is somewhat paradoxical since much of the research on shared intentionality is developmental (Tomasello & Carpenter, 2007; Tomasello et al., 2005; Tomasello, Carpenter, & Liszkowski, 2007; Tomasello & Hamann, 2012; Warneken et al., 2006). Tomasello and colleagues (2012), for instance, point out that these traits went down in ontogeny because they require time for maturation, though they do not provide a systematic phylogenic rationale for the plausibility of this view. If developmental timing is any guide, though, the early emergence of human proclivities for mutual orientation and joint activities would point to selective pressures on infant sociality. For the early emergence of these capacities would suggest that they are adaptations to challenges faced at those early ages—or at least, they became so (see Hawkes, 2012, 2014; Hrdy, 2009b).

Following Hawkes (2012), I think that the key selective pressures that originated shared intentionality were linked to changes in our rearing environment. The extent to which cooperative breeding took place prior to *Homo erectus* is difficult to determine. But is likely that this trait was already a distinctive feature of *H. erectus*, around 1.9 mya. Moreover, although *H. erectus* had the use of stone tools and weapons, and likely the control of fire (Gowlett, 2016), it seems extremely unlikely that they could have survived in the open, dangerous environment of Africa during the Pleistocene unless they lived in cooperative multifamily bands. The evolution of obligate bipedalism, for instance, would have made us quite vulnerable when resting. *H. erectus* females, in particular, would have had a hard time supplying the nutritional requirements of their offspring and defending them from predation without the assistance of other family members—though this might depend on when we lost our hair and whether the thermoregulatory capabilities that came with bipedalism were enough to enable foraging in the hottest times of the day, reducing contact with potential predators.[28]

As Susan Hrdy (2009b) has argued, much of our empathetic skills could well have been a consequence of the peculiar way that children were reared in the genus *Homo*. At an early stage in the evolution of our lineage, our bipedal ancestors were increasingly cared and provisioned not just by their mothers but also by alloparents, i.e., any conspecific involved in parent-like behavior towards an individual that is not his or her offspring. According to her, this form of cooperative breeding first emerged among our upright primate ancestors, and further evolved during the Pleistocene in African *H. erectus*. This shift is believed to have brought about important social cognitive consequences. For receiving help from others alters maternal trade-offs and imposes

---

[28] Species better adapted to arboreality do not face those challenges since tree nesting and adaptations for tree climbing offer more safety from potential predators. In addition, great apes such as chimpanzees and bonobos depend less on underground tubers and, therefore, females can forage without leaving their infants unattended.

new selection pressures on infants and young children to engage actively their caretakers' attention and commitment.

By the time that the first evidence of power-scavenging of big game start to appear, around 1.7 mya, *H. erectus* had new ways of finding, processing, and digesting food needed to support larger bodies and brains. Meat was indeed increasingly important, but its availability was unpredictable. Long-term trends toward a cooler, drier climate leading up to the Pleistocene pressured *H. erectus* ancestors to supplement a diet with reliable but hard-to-procure underground tubers that plants in dry areas use to stockpile carbohydrates (Hawkes, 2014; Hawkes, O'Connell, Jones, Alvarez, & Charnov, 1998; O'Connell, Hawkes, & Blurton Jones, 1999). Savanna baboons are known to dig up corms, bulbs, rhizomes, and tubers that require significant strength and manual dexterity (Altmann & Altmann, 1970; Hall, 1962; Hamilton, Buskirk, & Buskirk, 1978; Post, 1982; Whiten, Byrne, Barton, Waterman, & Henzi, 1991; Whiten, Byrne, & Henzi, 1987). Also, some savanna chimpanzees are known to use sticks to dig out shallow tubers (Hernandez-Aguilar, Moore, & Travis, 2007; Lanjouw, 2002). Therefore, in such ecological context, it would not be hard to believe that early hominins may have done so as well.

There is evidence that starchy tubers were an important fallback food for early African hunter-gatherers. For example, hunter-gatherers such as the Hadza in Tanzania, rely on roots and tubers as an important part of their diet and, consequently, they have accumulated extra copies of the salivary amylase gene (*AMY1*) that is correlated with salivary amylase protein levels—an important enzyme in the digestion of starchy foods (Perry et al., 2007). Similarly, recent analyses of molar surface texture also suggest that early *Homo* adopted a diet that included underground roots (Ungar, Grine, Teaford, & El Zaatari, 2006; Ungar, Krueger, Blumenschine, Njau, & Scott, 2012; Ungar, Scott, Grine, & Teaford, 2010). On top of that, the use of fire, perhaps around 800 kya (Wrangham, 2009), would have made tubers and roots gathered by females even more digestible and nutritionally valuable and, therefore, still more useful and relevant for survival (Hrdy, 2009a).

As a consequence, plant foods gathered and processed by females would have provided incentives for food sharing as well as new opportunities for post-reproductive females to help their kin by sharing those resources. According to the grandmother hypothesis (O'Connell et al., 1999), for instance, unlike other primates, natural selection would have favored females who live long after menopause to help provision younger kin. Grandmothers exemplify well the fitness benefits of cooperative breeding for the adult since several studies have shown that help provided by grandmothers produces positive effects on grandchild survival and growth, which leads to greater inclusive fitness (Fox et al., 2009; Jamison, Cornell, Jamison, & Nakazato, 2002; Sear, Mace, & McGregor, 2000; Voland & Jan, 2002). This hypothesis accounts for the distinctive human postmenopausal longevity, later maturity, and shorter birth intervals seen in our lineage (Blurton-Jones, Hawkes, & O'Connell, 1999; Hawkes, 2003; Hawkes et al., 1998; O'Connell et al., 1999). Moreover, it would have also stressed the selective pressures for increased social learning, as many of these tubers require significant processing to detoxify them.

More importantly, this new rearing environment allowed humans to wean infants relatively early and reduce the time-periods between pregnancies. Unlike any other ape mother, human mothers can bear a new baby before the previous child is independent because they receive help (Hrdy, 1999, 2005, 2009b). Cooperative breeding would have dramatically shortened those intervals in a way that would have facilitated longer periods of maturation and larger brains. It is now known that *H. erectus* possessed short growth periods that were more similar to chimpanzees than living humans, but childhood was already fully extended around 200 kya with the appearance of *H. sapiens*, as some studies of tooth growth in early hominins show (Smith et al., 2015; Smith et al., 2007). Cooperative breeding, so the argument goes, would have allowed the offspring to grow up slowly and remain dependent on parents and other caregivers for more time. As a result, retention of childlike features could be favored as well as an extended period of learning and brain growth.

However, more offspring at a time would have meant increasing competition between young offspring for attention from caregivers, who must then monitor the whereabouts and intentions of their caregivers (Hrdy, 2005, 2009b). This part of the story is essential for my account of the early ontogenetic emergence of shared intentional capacities (for a similar suggestion, see Burkart et al., 2009; Hawkes, 2012; Hawkes & Coxworth, 2013). The ontogenetic niche of early *Homo* would have been different from the one we see in other great apes. In great apes, mothers rear infants one at a time without help. In these species, maternal care is continuous. All great ape mothers are in constant contact with her infant during its first six months of life. They nurse long and wean late. Chimpanzees wean on average after 5 years (Clark, 1977), while orangutans at 7.7 years (Galdikas & Wood, 1990). This strategy is believed to increase the survival of juveniles by, for instance, reducing the risk of infanticide, at the cost of slow future population growth (see Kennedy, 2005; Lovejoy, 1981).

Unlike great apes, human infants do not have their mother's undivided attention (Hawkes, 2012). This likely led to an increase in the risk of mortality of infants and youngsters. Humans living in natural fertility societies, including hunter-gatherers, wean on average around 2.5 years (Kennedy, 2005; Marlowe, 2005), increasing the risks to the child (Arifeen et al., 2001; Clemens et al., 1990; Rowland, Barrell, & Whitehead, 1978) but boosting population growth. In these rearing environments, children survival would depend on their differential capacity to engage and create commitment with potential caregivers. They would have been selected for developing specialized adaptations for eliciting parental attention, soliciting more food, or being held or carried. Thus, selection for increasing social skills to engage mothers and others in caregiving activities would have been strong.

At first, this new evolutionarily context would have led to some new social adaptations for emotional sharing and engagement in early human infants. Unlike other great apes, human neonates smile. It is well known that around six weeks of age they begin smiling to others as a form of emotional engagement, followed some weeks later by laughing. Great apes such as chimpanzees and bonobos only smile and laugh when they are tickled (Bard, 2012; Davila-Ross, Allcock, Thomas, & Bard, 2011). Around 2 months of age, human infants begin to engage in turn-taking sequences involving sounds and gestures that some people call 'protoconversations' (Reddy, 2015; Rochat, 2009; Trevarthen, 1979). Human children seem highly sensitive to this form of emotional feedback. In still-face experiments, for instance, when toddlers try to interact with a nonresponsive expressionless adult, they repeatedly attempt to get the reciprocal interaction back by smiling briefly at the adult. But when these attempts fail, infants usually withdraw, orienting face and body away from their adult partners (Tronick, Als, Adamson, Wise, & Brazelton, 1978). Experiments on gestural communication also show that when adults suddenly stop communicating with infants, they react in a similar way (Tomasello, 2008; Tomasello, Carpenter, et al., 2007).

Along with emotional sharing and engagement, human infants also develop a capacity to coordinate mutual gaze with caretakers. This capacity is not evident in wild infant chimpanzees (Plooij, 1984), although parallel attention has been reported in captivity (Tomonaga et al., 2004). This phenomenon has been sometimes attributed to the fact that the attention of multiple caretakers in captivity, including human ones, resembles the ancestral environments where grandmothers and other alloparents started to provide independent mothering (Hawkes & Coxworth, 2013; Hrdy, 2014). It appears then that some important precursor of this trait was already present in our ancestors, but it was later on that this infrastructure became richer and more complex via adult social selection due to a change in rearing environments (Hrdy, 2009b).

These traits are sometimes understood merely as adaptations for bonding, but there are reasons to think that they also extend in development as adaptations for basic communication and social learning. For example, at around 9 months of age, human infants also begin to understand others as intentional agents (Tomasello et al., 2005), i.e., agents with goals and perceptions that guide their behavior—similar perhaps to other great apes but at a much younger age (Wobber et al., 2014). When a system geared to a particular developmental period meet new systems geared to a different developmental stage, this produces novelty. Then, along with the emergence of basic mindreading capacities, the sharing of emotions extends to the sharing of attention to external events. Joint attention is perhaps the most basic form of shared intentional states. It requires not only active tracking of the gaze direction of other individuals but also active triadic engagement by, for instance, showing and offering objects to one's partner with an expectation that this will lead to a positive sharing of emotions about these objects. These behaviors are not present in great apes (Tomasello & Carpenter, 2005; Tomonaga et al., 2004).

Joint attention could have emerged out of basic capacities to track other individuals' targets of attention by means of some recursive structure of the form "I know that you know that I know that *p*". Although this interpretation is likely far-fetched and over-intellectualized if understood in terms of explicit mindreading, current theories of mindreading often distinguish implicit forms of mindreading which do not make demands on executive function (Heyes & Frith, 2014). This form of mindreading could then be present in early childhood (Onishi & Baillargeon, 2005), but also in great apes (Krupenye et al., 2016). This suggests that the mechanisms that mediate explicit mind reading are different from those controlling implicit mind reading, which are phylogenetically ancient and perhaps genetically inherited. However, studies of twins (Hughes et al., 2005), people with hearing impairments (de Villiers & de Villiers, 2012; Meristo, Hjelmquist, & Morgan, 2012; Pyers & Senghas, 2009), and children from Western (Lohmann & Tomasello, 2003) and non-Western societies (Seeger, 1981) suggests that many aspects of explicit mindreading might well be a culturally inherited skill. As a result, infants might well be equipped with neurocognitive mechanisms that yield accurate expectations about behavior in the form of automatic or implicit mindreading, whereas explicit mindreading is a culturally inherited skill transmitted by verbal instruction (Heyes & Frith, 2014).

Yet basic capacities for shared intentionality may not necessarily be a consequence of recursive forms of implicit mindreading. These capacities could have also emerge as a new cognitive phenotype to form we-mode representations of the form “We know that *p*”, i.e., knowledge states that are represented as shared by a certain group of individuals, but which cannot be reduced to recursive mindreading. In this latter view, joint, collective mental states are ascribed to single individuals rather than groups—i.e., they are intentional states held by individuals but which make fundamental reference to a collective formed in conjunction with the other individuals (Searle, 1990). In fact, it is possible that the capacity for shared intentionality in modern humans relies on the coexistence of alternative paths that lead to the formation of these two varieties of shared intentional states. One of these paths can be recursive, while the other is irreducible to a plural mode of representation.

As a conjecture, it is possible that joint attention in early *Homo* infants could have relied initially on some form of implicit, but recursive form of mindreading, even if humans were able to evolve at some point a capacity to form irreducible we-mode representations. One reason to prefer this explanation is that populations usually contain selectable quantitative variation in existing traits—e.g., infant chimpanzees have basic capacities for gaze following and parallel attention (Okamoto, Tanaka, & Tomonaga, 2004; Okamoto et al., 2002; Tomonaga et al., 2004). In this line of argument, evolutionary change would work more easily by modifying existing traits rather than creating new ones from scratch.[29]

Real joint attention, which occurs in all sensory modalities, requires from infants to be aware that caregivers are attending to the same state of the environment as they are. Joint attention is in this sense a special form of mindreading. It requires sharing attention to goal-directed, intentional actions. For attentional states are intentional states, i.e., the agent’s focus of attention is directed to a target As in the case of parallel attention, infant’s and caregiver’s focus of attention should be directed to the same

---

[29] It is increasingly clear that chimpanzees and bonobos possess some form of mindreading (Call & Tomasello, 2008; Karg, Schmelz, Call, & Tomasello, 2015), including a capacity to track other individual’s target of attention. However, the above explanation also relies on the assumption that mindreading capacities regarding other individual’s focus of attention were present in early *Homo* infants. The existence of basic capacities for gaze following and parallel attention in infant chimpanzees might be too meager evidence for the existence of mindreading capacities, i.e., whether chimpanzees attribute referential intent and visual experience to their partners or merely follow gaze direction to specific locations (see, for instance, Okamoto et al., 2004, p. 243).

object. The crucial difference is that in cases of joint attention not only both individuals are aware of their own attentional states but also they are able to track their partners' focus of attention, such that both understand they are attending the same target, giving them an incipient form of bird's-eye view. This could have been achieved through some form of recursive mindreading of the form "I know that you know that I am attending to object *x*". Or perhaps in a more implicit and perceptual way such as "I see that you see that I am attending to *x*", for shared attention does not need to be an abstract and amodal form of representation. In both cases, this would allow the child to make a mental comparison, so to speak, between his/her own perceived targets and the targets of the partner's focus of attention.

Truly joint attention would have also required the infant to engage caregivers just for the sake of sharing with them that attentional state. Without parents by their side, infants would need to engage other adults in order to receive alloparental care from them. Initially, they could have used joint attentional skills to take advantage of already existing adult bonding and affective mechanisms for their own ends. But potential alloparents would be more conditional, less responsive, and less motivated caregivers, even between individuals from an extremely socially tolerant lineage (see chapter 3). They could just ignore their demands or provide attention to other infants. Under those conditions, children would have needed to engage the caregiver's attention for the sake of the engagement, instead of just monitoring the adults' actions. This engagement probably involved not only simultaneously attending to the same object but also likely some form of visual referencing of the partner by the infant, i.e., a look to a caregiver's face or eyes about an object or event in the infant's and caregiver's mutual focus of attention (Carpenter & Call, 2013; see also Bakeman & Adamson, 1984). They would have needed to register that the caregiver was also looking to them, and then to manifest some sort of affective contact with the adult (Hobson & Hobson, 2007).[30]

Initiating episodes of joint attention would be very important for the infant. Cooperative breeding would have forced infants to create as much attachment as possible with others because alloparental care is more conditional and distributed across many infants in the group. Gestures and eye contact would be important under those conditions to direct others' attention to objects, to events, and to themselves, just for the sake of sharing with them interests and pleasurable experiences that foster attachment

---

[30] It is also very likely that once selective pressures for cooperative breeding took place, adult individuals would have also been selected for increasing responsiveness to children's attempts to initiate joint attention engagement.

and care. Chimpanzees, for instance, can follow a human's gaze to an out-of-sight location behind a barrier and use gestures to beg from a human most often when the human can see them. But there is no evidence that they can spontaneously share attention and other experiences with members of their own species (Tomasello & Carpenter, 2005).

There seem to be markers of this process in our anatomy. The human eye has unique features, such as a high contrasting sclera. This trait could have evolved to provide information about other people's attention and boosting social cognitive development (see, for instance, Baron-Cohen, 1995; Kobayashi & Kohshima, 2001; Tomasello, Hare, Lehmann, & Call, 2007). In particular, these features of the eye would allow the infants' saccades to be easily observed by caregivers and, therefore, act as elicitors of contingent social feedback (Mundy & Jarrold, 2010). For example, shifting visual attention would enable infants not only to actively select information from their environments but also to elicit social-behavioral responses from others, e.g., parental smiles, vocalizations, or gaze shifts. Given the appropriate motivational machinery, parents and caregivers could pick-up and show an out-of-reach object to the child, when the infant shifts attention to the object. Similarly, if a child shifts attention to the caregiver's eyes, the infant could receive a vocal or physical response.

Similarly, basic forms of cooperative communication could have also been a consequence of the early emergence of these shared intentional skills, for they could have emerged prelinguistically in the form of pointing and pantomiming. Great apes will sometimes point for humans in an imperative way, e.g., when they want an out-of-reach object (Leavens & Hopkins, 1998). But human infants across different cultures point declaratively from around their first birthday (Callaghan et al., 2011; Salomo & Liszkowski, 2013; Tomasello, Hare, et al., 2007). At that crucial developmental stage, human infants become highly motivated to share attention and interest to different objects and events by offering, showing, and pointing referentially to external objects and events.

To sum up, all these uniquely human infant behaviors would have evolved in the context of sibling and peer competition (Hrdy, 2006, 2009b). For in a cooperative breeding context, adults are rewarded by sharing emotions, interest, and attention to objects and events with infants. As a result, infants compete to provide adults with this positive affective feedback to secure care and attention. This creates a process of social selection on infants for increasing shared intentional skills (Crook, 1972; West-Eberhard, 1979). In this view, shared intentionality started as an ontogenetic adaptation

(Alberts & Gubernick, 1984), i.e., as part of a sequence of specialized changes that enabled infants to survive the ontogenetic niche created by their parents. However, as we will see in the next section, these early mechanisms of joint attention and basic communication very likely also coevolved as mechanisms to facilitate social learning, and perhaps even teaching, when adult activities required skill that demanded prolonged periods of preparation and maturation.

4.5. Evolutionary development

Recent models of the role of ontogeny in the evolution of human cooperation take into account the tradeoffs that emerge during the transition from childhood to adulthood (Tomasello & Gonzalez-Cabrera, 2017). As I have argued above, shared intentional states were initially geared to dyadic interactions between infants and caregivers. But once they emerged, they were likely retained in adulthood because they would transform social interactions in a species characterized by its prosocial behavior. A heterochronic process of juvenilization could have played an important role here, along similar lines of the self-domestication hypothesis introduced in chapter 3 (see also Hare et al., 2012). There is a good evolutionary scenario to argue in favor of such a process here, for cooperative breeding facilitates prolonged periods of maturation and collaborative activities themselves would have benefited from those shared intentional skills that first emerged as ontogenetic adaptations.

These changes would have been deep and complex, with important downstream consequences for human development. Konner's (2010) detailed comparison of apes and humans developmental trajectories, for instance, documents a series of changes in developmental timing that are similar to those described by Hrdy (2009b) and Hawkes (2003, 2014), e.g., increased life span and time to sexual maturity, longer gestation, rapid brain growth, and juvenilization of some anatomical and behavioral features. But he also includes the appearance of middle childhood as a new developmental period between infancy and puberty.

Cooperative ventures such as pair bonding, cooperative breeding, collective defense, and more cognitively taxing activities such as collective hunting and foraging would have been transformed by extending these basic shared intentional capacities into adulthood. On my view, cooperative breeding came first. These rearing environments selected for very basic capacities of shared intentionality. For example, infants and alloparents could mutually engage in episodes of shared attention and shared communication through pointing and perhaps pantomiming in situations in which initially only the infant would be motivated to join the adult's intentional states and share their own mental states. Once these basic capacities for shared intentionality were exploited in these social contexts, they triggered a positive feedback loop for increasingly joint, collective intentional states. I think that collective hunting was the decisive step in this direction. If very basic forms of shared intentional states based only on recursive forms of mindreading were ecologically viable in less motivationally and cognitively demanding dyadic contexts, they would have required a particular mode of representation that would allow them to form shared intentional states automatically and effortlessly in high-stake contexts such as collective defense, scavenging, or collective hunting. To put it briefly, cooperative breeding would have initially triggered the shared intentionality syndrome. But these basic capacities would have subsequently extended into adulthood transforming our foraging practices, which in turn would have selected for more complex forms of shared intentionality.

The tamer picture of the *Pan*/*Homo* LCA I have proposed in chapter 3 would have been crucial for this purpose. For if early hominins were indeed tolerant of conspecifics cofeeding, they would be especially well suited not only for cooperative breeding but also scavenging since it would reduce the social stress generated by eating together around a carcass. Chimpanzees who rank high in food tolerance tests, for instance, have shown to be more collaborative and willing to share the fruits of their collaborative efforts than intolerant chimpanzees (Melis et al., 2006b). As we have seen, bonobos are more tolerant around food. In comparison to chimpanzees, they cooperate more and share more food items that are easily monopolizable (Hare et al., 2007). Therefore, increased tolerance around food in the *Pan*/*Homo* LCA would have given substantial leverage for the emergence of hominin species which actively shared the spoils of collaboration, for there is evidence that tolerance is reasonably evolutionarily labile (Hare & Tomasello, 2005).

In a context as the one proposed by Tomasello and colleagues (2012), where global cooling and a drying would have caused an unprecedented radiation in terrestrial monkeys, scavenging could have been a reasonable strategy since they would have competed with *Homo* for many plant foods. Hunting smaller prey would have been also ecologically possible, but this type of hunting is also widespread among modern terrestrial primate species. Another available option would have been to procure meat from large mammals, which is riskier but could have been initially obtained through scavenging (Binford, 1981, 1987; Blumenschine, 1986; Blumenschine, Cavallo, &

Capaldo, 1994; Bunn & Blumenschine, 1987; Bunn et al., 1980; Potts, 1984; Shipman, 1986).

Rudimentary projectile technology would have made possible to drive predators from their kills, e.g., by throwing rocks and clubs at them. Even this simple technology would significantly reduce the risk associated with the activity, taking advantage of dangerous predators that suddenly become easier stationary targets. For a predator, it may be better to surrender a kill and try it again than being injured by a group of scavengers. Scavenging was likely the precursor of collective hunting. But aggressive scavenging is a cooperative, stag hunt-like activity that requires no complex coordination or role differentiation—although it requires commitment and tolerance around food.

Around 800 kya scavenging turned into active collaborative hunting (Dubreuil, 2010b), although others like Bunn (2007) believe that hominins were already hunting large mammals by around 1.6 mya. Collaborative hunting is cognitively more taxing than scavenging primarily because it increases the cognitive demands on planned coordination. It also likely required better tools (especially weapons) since hunters are not just driving rivals away. Certainly, it still requires a capacity to share the spoils of the hunt and to avoid free riders. But early *Homo* would have already developed many social cognitive skills that made possible such a demanding collective activity, including joint attention and basic communicative skills.

Ontogenetic adaptations such as joint attention and communicative gestures are usually seen as having immediate adaptive value while contributing little to individual's chances of survival and reproduction of the individual during its adulthood. These ontogenetic changes help children to survive in their current ontogenetic niche. But ontogenetic adaptations could have easily become essential skills for adults too. This would have been the case in contexts such as collective hunting (Tomasello et al., 2012). In fact, on the view I am defending, one distinctive contribution of collective big game hunting would have been to enhance our emerging capacities for shared intentionality over a gradient of cognitive complexity for planned coordination—at the same time, it would have continued expanding our baseline capacity for social tolerance and impulse control. Another is to increase the demand for different forms of partner control that are fueled by distinctively human forms of punitive attitudes, which likely increased the costs and payoffs in situations of partner choice and led to an increased sensitivity to the reputational effects of collaboration via social selection of collaborative partners (Boehm, 1999, 2007, 2012).

Crucially, then, the transition toward active collaborative hunting was largely leveraged by the interplay between the selective pressures that led to cooperative breeding in humans as well as organized, goal-oriented, collective hunting. Cooperative breeding increases sibling and peer competition between infants. But as children grow older, they start spending more time among peers since even under conditions of cooperative breeding most of the adult's attention would be redirected toward younger individuals. As a result, the frequency and reliance of peer interaction gradually increases. The fitness benefit is first and foremost about the particular developmental periods involved, but it may extend to later developmental periods if traits developed at one period turn out to be good preparation for challenges faced at a later developmental period. This is a plausible assumption. Middle childhood is sometimes called the 'forgotten age' due to the predominant focus in development psychology on early childhood and adolescence (Mah & Ford-Jones, 2012). But peer interaction during this period, e.g., in the form of play, could have helped as a preparatory step for more serious and risky collaboration during adulthood. Then, assuming random variation in time of onset, natural selection could have favored individuals with a more premature development of adult-like shared intentional capacities, including their distinctive punitive attitudes and concomitant sensitivities for partner choice.

Of course, during the transition to a middle childhood, children are still dependent on adults. They continue to be engaged in a process of sibling and peer competition for care and attention, including against younger individuals who naturally receive most of the attention. As a result, engaging caregivers becomes progressively more complex as they focus their attention to the youngsters. But at the same time, children also become more physically apt to help and collaborate. Learning how to make themselves useful to adults in their everyday activities would be then a possible way to gain such care and attention. Younger children are not of much help to adults, but older ones (starting from 7 years of age in extant natural-fertility populations) may contribute to the overall productivity of the group as to generate an intergenerational interdependent economy (Kramer, 2005, 2010). This creates, in turn, a safe epistemic niche for the child to rehearse their collaborative skills and to learn about adult activities even without explicit teaching involved.

The skills developed in the child's interaction with adults are then redeployed in the context of peer collaboration. During middle childhood, a child has to learn how to make decisions with others with little or no adult supervision. They are typically low-risk decisions, such as where, when, and what to play, in environments that still afford social exploration. Adult individuals would be concerned with defense and childcare for the youngsters but they will likely not interfere with children businesses. In this context, it would have been important to learn how to treat others fairly or in a mutually satisfactory manner; not only because after certain age adults do not mediate peer interactions but also because they are the onset of the social relationships that are essential for survival during adulthood. At that stage, children have the opportunity to interiorize the social benefits of sharing resources with others and to learn how to be a reliable playing partner who keeps his/her commitments to joint activities. Moreover, the transition to middle childhood open a window of opportunity to learn the social cost of peer reprisal in low-risk contexts where failure to cooperate is not lethal.

Recent studies on impression management show how sensitive modern human children are to the reputational effects of their social behavior on peers. Young children, but not other great apes, seem to be concerned about others’ evaluations of their cooperative and prosocial tendencies because they appear to adjust their behavior based on their prediction of how others will assess this behavior (Engelmann et al., 2012). They seem able to make reputational judgments about others as cooperators, based on direct or indirect evidence, i.e., they know that they are being judged by others and, consequently, they modify their behavior to affect those judgments. Of course, it is difficult to project features of extant human populations back into ancestral ones. But giving the child psychology that we see in extant human populations, it indicates that these features could have ‘migrated down’ in development as a preparation for challenges faced at a later developmental period—in particular, if longer periods of maturation of essential social skills can indeed benefit complex adult activities such as collaborative foraging.[31]

5-year-old human children, for instance, have been shown to share more and steal less when they are being watched by a peer than when they are alone. In contrast, chimpanzees behave the same regardless of the condition. Previous research has shown that at this age children first engage in second-order mental reasoning (Frith & Amodio, 2006; Sullivan, Zaitchik, & Tager-Flusberg, 1994) of the form “I am thinking about what you are thinking about me” (Banerjee, 2002), which is a central cognitive prerequisite of self-reputational behavior. The prediction in these studies is that subjects would help more and steal less when are being observed if they were concerned with their self-reputation as collaborative partners. In one of these studies, for instance,

---

[31] I owe this metaphor about traits migrating up and down in development to Michael Tomasello.

preschool children have to solve a task that requires more stickers than the ones they receive in an envelope from the experimenter. Under this condition, solving the tasks requires the subject to take some of the stickers that are left in the experimental room but which the subject is explicitly told are intended for another child who would play the game later on. Consistent with the hypothesis that children care about their reputation, if they try to solve the task when there is a novel peer in the room, they are much more prone to refrain from taking the stickers than when there is no peer in the room.

Overall, experimental results indicate that 5-year-old human children (but not chimpanzees) share more and steal less when they are being watched by a peer than when they are alone. In the stealing task, for instance, subjects stole in 4% of all cases in the observed condition and in 24% of all cases in the unobserved condition. In the helping task, subjects helped in 11% of all cases in the unobserved condition and in 28% of all cases in the observed condition. In contrast, in the helping task, chimpanzee subjects helped in 34% of all cases, for the observed condition, and 36% of all cases in the unobserved condition. Similarly, in the stealing task, they stole in 20% of all cases, for the observed condition, and in 23% of all cases in the unobserved one.

When we developed this sensitivity to social reputation is an important but open question. However, the proposal is that at some point in our evolutionary history (initially in play, but perhaps in other contexts as well) young children started to cooperate with peers or risked suffering the fitness consequences of not doing so. The primary selective pressures would have proceeded mainly upstream from adulthood and adolescence to childhood, enhancing their sensitivity to the reputational effects of their own cooperative behavior. Arguably, this capacity did not evolve in the context of child-adult interactions since these relationships are hierarchically structured and, therefore, much more sensitive to covariables such as authority and fear of punishment.

This is important because, during adolescence and sexual maturity, peer interaction becomes increasingly socially taxing in an evolutionary scenario such as the one envisioned by the interdependence hypothesis, where one must collaborate with one’s peers for basic subsistence. And probably of importance for fitness too since the social capital that children build during childhood in the form of networks of allies and friends will endure into adulthood, making a non-trivial difference in terms of their fitness. If there were selective pressures for obligate collaborative foraging, as the interdependence hypothesis postulates, then those individuals who were best able to collaborate with their peers would have had an adaptive advantage. It is in this context that it has been speculated that those individuals who had already developed especially powerful skills and motivations of shared intentionality, beginning in infancy and continuing throughout childhood, would already have a special stock of skills needed for cognitively and motivationally activities such collaborative foraging (Tomasello & Gonzalez-Cabrera, 2017).

The result of this process should have been some kind of ontogenetic preadaptation, i.e., an ontogenetic adaptation selected for at an early developmental period that turns out to be useful for a later development stage—especially when individuals reach sexual maturity and, therefore, not only survive but reproduce. This process creates an evolutionary feedback loop of downstream and upstream selective forces in the emergence of shared intentionality and its distinctive normative dimension (see figure 4.2).

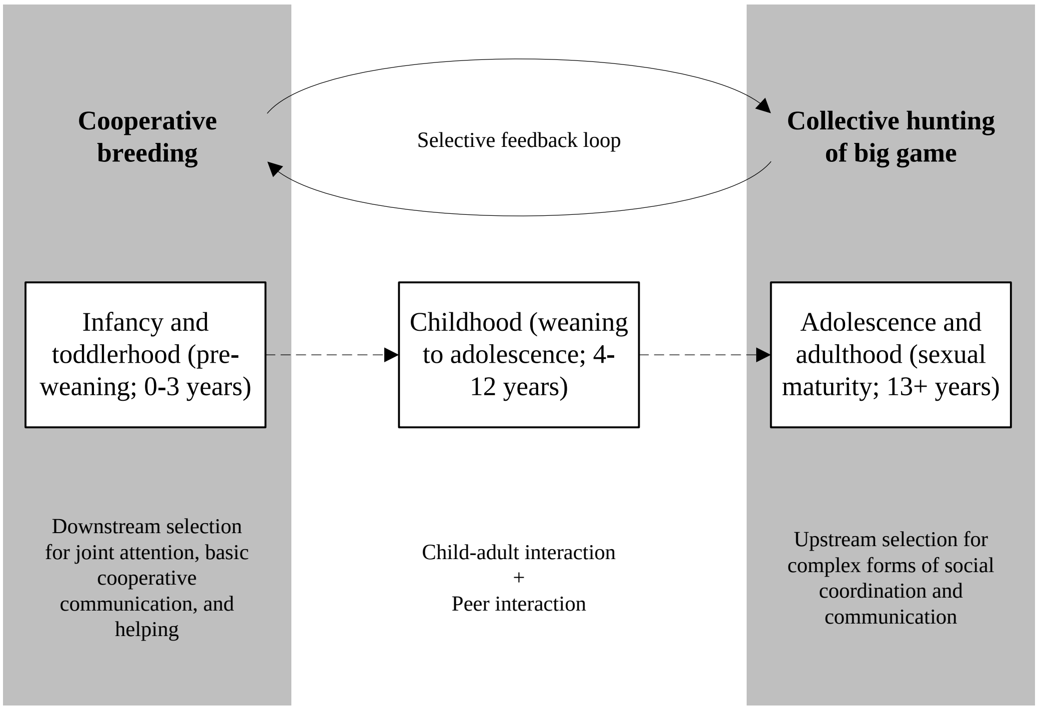

Figure 4.2. Selective feedback loop between cooperative breeding and collective hunting of big game. Cooperative breeding generates downstream selection in infancy and childhood for joint attention, basic forms of communication such as pointing, and increased helping skills. These basic capacities subsequently extend into adulthood, given the advantages they provide to coordinate adult activities, especially collaborative foraging. In turn, selective pressures stemming from increased interdependence in adult activities created demands for more complex adaptations for coordination and communication that extended downstream into childhood as a preparation for adult social interaction. This includes an initial understanding of social normative demands as a function of selection for displaying punitive attitudes toward those who do not meet our joint social expectations. The developmental trajectory is represented here with dashed lines.

Initially, cooperative breeding generates downstream selection in infancy and toddlerhood for basic capacities of shared intentionality such as joint attention, communicative pointing, and other helping skills. These basic capacities subsequently extend into adulthood, due to the selective advantages they provide at the cost of relatively little or no disruption in the overall cognitive development of the organism. In particular, given the advantages these skills provide to coordinate adult activities, especially collaborative foraging, selective pressures stemming from increased interdependence created demands for more complex adaptations for coordination and communication that extended downstream into childhood as a preparation for adult social interaction. More importantly, as part of their preparation for adulthood, children began to be progressively sensitive to the social normative demands of their social environment as a result of selection for displaying punitive attitudes toward those peers who do not meet their joint social expectations. Traces of this process can be observed in the increasing awareness of the child about their own social reputation.

In sum, if the above model is right, there were selective pressures for infants' ontogenetic adaptations for joint attention, cooperative communication, and helping, to 'migrate up' because they were useful for children and adults in the context of ecological pressures for collaborative foraging. But ontogenetic adaptations geared toward a particular developmental period are also extended downstream in development as a natural consequence of random variation in age of onset and selective pressures acting on adjacent developmental periods (Tomasello & Gonzalez-Cabrera, 2017).

4.6. Conclusion

Following my discussion in chapter 3, I have argued here that the more socially tolerant model of the *Pan*/*Homo* LCA I offered in that chapter is compatible with the idea that early hominins hunted in groups (see also Pickering, 2013). The tamer picture of our early ancestors I offered in that chapter played a crucial role in the present chapter in two different ways. One was to provide a cognitive baseline for my lineage explanation of normative guidance. Another was facilitating the gradual expansion of early hominin cooperative capacities since collaborative foraging would be implausible in a very aggressive and dominant lineage. However, I think that collaborative hunting is only part of the story of the evolution of normative guidance.

Certainly, the traditional interdependence hypothesis of the origins of shared intentionality points to collaborative foraging, especially collective hunting, as the main driver in the evolution of our shared intentional capacities (Tomasello et al., 2012). But this explanation leaves open the question of why these capacities emerge so early in ontogeny. I have argued instead that cooperative breeding was the initial driving force in the evolution of our share intentional capacities (see Hawkes, 2012, 2014) while the distinctive normative dimension of some shared intentional states emerged largely as a consequence of the selective pressures on collaborative hunting. For obligate collective hunting requires shared intentional states that regulate collective activities in a way that is distinctively normative, i.e., they are mental states that are generalizable in scope and intrinsically motivational but also able to engender punitive attitudes (Sripada & Stich, 2007).

Finally, I have put forward a model of the role of ontogeny in the evolution of normative cognition. This model takes into account the tradeoffs that emerge during the transition from childhood to adulthood in order to integrate insights from both the interdependence hypothesis and the cooperative breeding hypothesis (Tomasello & Gonzalez-Cabrera, 2017).

Chapter 5. Sharing our normative worlds

In chapter 4, I have argued that shared intentionality emerged from the interplay between the selective pressures that led to cooperative breeding in humans and collective hunting. On this account, selection did not only act upon adult cognitive capacities but rather upon the entire life cycle. Now, in this chapter, I will connect these issues with a particular ontogenetic account of this capacity. In particular, building upon usage-based models of language acquisition (Bybee, 1985; Givón, 1979; Langacker, 1987), I will argue that the process of norm acquisition is the result of domain-general mechanisms of pattern-recognition, intention-reading, and affective processing. These mechanisms allow children to produce normative judgments that are generalizable, intrinsically motivational, and which trigger punitive attitudes against norm violators. Moreover, some of these normative judgments are moralized in a prototypical way often engendering moralized emotions like guilt and shame (but also less studied positive emotions) which contribute to these judgments being perceived as inescapable and authority independent (Mameli, 2013). As we will see in this chapter, this model leads to important philosophical consequences for one of the more heated debate in contemporary moral psychology, namely the debate on moral nativism (Dwyer, 2007; Hauser, 2006b; Hauser, Young, & Cushman, 2008a, 2008b; Joyce, 2006; Katz, 2000; Mikhail, 2007; Nichols, 2005; Prinz, 2009; Sterelny, 2010).

In the view I have been defending, although moral judgments are not a unified class of mental states, shared intentionality can help us to define a particular kind of social normative thinking. Moreover, some prototypical moral judgments may emerge as a particular subclass of these mental states, for as I will argue in this chapter, emotional and affective dispositions not only play a key constitutive role in normative cognition but also they help us to single out a particular subset of prototypical moral judgment that are perceived as inescapable and authority independent (Mackie, 1977).[32] In particular, what I want to show in this chapter is how shared intentionality enables the

---

[32] It might be important to remember that this is a claim about the cognitive psychology of normative thinking and its origins, not about the semantics of moral language. The cognitivist/noncognitivist divide is difficult to bridge in metaethical discussions about the semantics of moral language because it is difficult to come with an appropriate formal semantic model (Schroeder, 2008). However, one does not face the same challenges in psychology where formalizing semantic content is not a theoretical goal. Many mental states such as emotions typically (though not necessarily always) involve both cognitive and conative states.

representation of mental states of the form “*x* is morally required” in an agent-independent way, and how they are prototypically moralized in virtue of their affective and motivational component.

For this purpose, the rest of this chapter is divided as follow. In section 5.1, I will explain the philosophical motivations behind debates about the ontogeny of normative cognition through views of moral nativism. In section 5.2, I will contrast these views with one that extends the ontogenetic account of normative cognition introduced in chapter 4. In section 5.3, I will introduce a model of norm acquisition that builds upon usage-based theories of language development in cognitive linguistics. Certainly, the model does not depend on the validity of any particular approach to language development, but they help us to understand how the cultural transmission of social norms is possible through the motivational and cognitive infrastructure of our shared intentional psychology. In section 5.4, I will argue that this model allows us to understand how some normative judgments are perceived as inescapable and authority independent by selectively tuning the affective processing mechanisms involved in norm acquisition and norm execution. Finally, in section 5.5, I will draw some philosophical consequences from this discussion by arguing that at least not all of our moral cognition is built upon alleged innate moral information, which gives additional support to the fragmentationist view of moral judgments introduced in chapter 2.

5.1. Moral nativism

In chapter 2, I have argued that the class of moral judgments shatters into different prototypical classes of moral judgments such as judgments with Turiel’s signature moral pattern and judgments about inescapable and authority independent demands. I argued that these fragmented class of prototypical moral judgments could by clustered, nonetheless, around a special form of norm psychology that stems from our distinctive capacities for shared intentionality. I proposed in that chapter that the lineage explanation of shared intentionality could help us to explain the emergence of shared intentional normative judgments, which in turns fractionates into these different clusters of prototypical moral judgments, each of which might be a culturally relative and acquired through the interaction with our particular cultural environments. However, as pointed out at the beginning of the thesis, some dominant views in moral psychology (Dwyer, 1999; Hauser, 2006b; Mikhail, 2011; Mikhail et al., 1998) argue that moral cognition is not only a well-defined natural kind in the psychological sciences, but also that moral thinking is innate. At least to an important extent, these views offer a quite different view than the one I try to defend in this dissertation—a discussion that I have postponed until this chapter. The aim of this section is primarily expositive, for I want to explain moral nativism in one of its most prominent versions, the so-called ‘linguistic analogy’, and provide the philosophical motivation behind debates about the ontogeny of normative cognition, which I will address in this chapter.

The lineage explanation I have put forward in this thesis integrates facts about the evolutionary and developmental trajectory of shared intentionality and its distinctive normative dimension. This account connects normative thinking to domain-general mechanisms of shared intentionality that arguably played an important role in other aspects of the human uniqueness syndrome such as human ultra-cooperative tendencies (Tomasello, 2009b), enhanced mindreading skills (Call, 2009), and human-unique capacities for linguistic communication (Tomasello, 2008). I understand domain specificity here as the view that specialized learning devices support some aspects of cognition. This position is closely linked to modular views of the mind. But since domain specificity does not commit to all the assumptions of traditional Fodorian modularity such as informational encapsulation, the former can be considered as a more general version of the thesis of the modularity of mind (Fodor, 1983). Thus, the mechanisms of shared intentionality that support the kind of social normative thinking I propose in this dissertation are domain-general. For they do not just enable acquisition of norms and the development of the capacity to make social normative judgments, but also enable a number of different cultural skills that children normally acquire through social learning—more notoriously, as I will argue in section 5.5, many key aspect of their capacity for linguistic communication. To put it briefly, there are reasons to think that the basic cognitive and motivational infrastructure underlying shared intentionality is highly entrenched in human development. But many aspects of human cognition derived from those skills, including the kind of normative cognition I have been discussing in this thesis, are the result of those cultural environments that our shared intentional capacities help to construct (Tomasello et al., 2005). In this chapter, I will focus on this latter point.

Roughly understood, psychological nativism is the view that certain cognitive capacities are innate. The idea of innate traits can be explained via genetically canalized developmental systems that take development to the same endpoint from many different environmental starting points (Ariew, 1996, 1999, 2007). However, both the appropriate account of innateness and its extent is controversial (Griffiths & Machery, 2008;

Mameli & Bateson, 2011), but I will set aside this issue to focus on a particular family of nativist views that stem from the Chomskyian model of language development. Nativist approaches are popular in the contemporary debates on human cognition mainly due to the prominence of linguistic nativism. It not surprising, then, that one of the most prominent nativist views of normative cognition has borrowed different elements from it. Building on Chomsky's (1957) ideas, researchers such as John Mikhail, Marc Hauser, and Susan Dwyer, for instance, have vigorously argued that humans possess a universal moral grammar similar to the one put forward by Chomskyan linguistics (Dwyer, 2007; Hauser, 2006b; Hauser et al., 2008a, 2008b; Mikhail, 2007). The key argument in favor of the existence of a universal grammar in linguistics (and moral psychology) is the so-called 'poverty of stimulus argument', i.e., the idea that the environment does not contain enough information to enable a learner to acquire the linguistic competence typically displayed by young children—at least, not by applying mere domain-general learning mechanisms to environmental inputs (see Botterill & Carruthers, 1999; Cowie, 1999; Laurence & Margolis, 2001).

Roughly speaking, the Chomskyan picture of language acquisition was forged as a reaction to the Skinnerian model of language development. According to Skinner (1957), children learn a language by deploying a variety of simple learning mechanisms—primarily, mechanisms of instrumental conditioning and principles of induction that allow stimulus generalization. But Chomsky (1968, 1980, 1986) famously argued that there are many abstract grammatical principles to which children lack any appropriate evidential access. Children's language access consists of only a few series of discrete individual utterances. For example, understanding the relationship between indicative and yes-no questions requires identifying the organization of lexical items into subsentential constituents that cannot be captured as a set of simple lexical patterns (Chomsky, 1980). Hence, there should exist a specific language faculty with innate domain-specific information that is recruited in first language acquisition: an innate universal grammar which is common for all natural languages and that is common to all stages of child language development (Pinker, 1984).

Of course, languages differ considerably from one another, but those differences are a matter of parametric variation. On the principles-and-parameters approach (see, for instance, Chomsky, 1981, 1982) the syntax of a natural language is described through the interaction of both general principles and specific parameters which are either turned on or off depending on the particular language we acquire from our cultural environment. As pointed out by Dąbrowska (n.d.), the list of relevant parameters differs substantially from one linguist to another. Pinker (1994, p. 112) thinks that there are only a few, Fodor (2001, p. 734) suggests that there might be around 20, Roberts and Holmberg (2005, p. 541) argue that they might be around 50-100 of them, and some others have thought that might be hundreds of them (Kayne, 2005). There is no canonical list of parameters. For example, Baker (2001) discusses 10 different parameters, while Fodor and Sakas (2004) list 13, with few parameters in common between both lists (Haspelmath, 2007; Tomasello, 2005). And even Chomsky (1993, 1995, 2001) himself has abandon this model in favor of a minimalist approach.

In this line of thought, Hauser and colleagues have argued that the child's moral judgments emerge from a universal moral grammar consisting of various abstract moral principles. Specific moral judgments are just the consequence of the interaction between these abstract moral principles and specific cultural environments through a process of moral learning—a process in which cultural environments set the parameters of the universal moral grammar to the specific values of one's community (2006b, p. 43). As in the case of linguistic acquisition, the key argument in favor of moral nativism is the poverty of the stimulus argument. On this version of the argument, the capacities evident in moral cognition are acquired in a manner that goes beyond the information available in the child's local learning environment. For example, John Mikhail (2009) has notably argued that the moral environments to which children are exposed are too impoverished to explain how children understand the moral/conventional distinction (Nucci et al., 1983; Turiel, 1983, 1998). So, the nativists argue, the moral/conventional distinction involves innate, domain-specific information.

Another example comes from Hauser (2006b) study on the so-called 'principle of double effect', i.e., the idea that it is more acceptable to cause a harm for a greater good if that harm is a foreseen but unintended side effect than intentionally causing such a harm as a means to the same end. People often distinguish between actions that produce foreseen but unintended consequences, and actions that bring about those foreseen consequences in an intended way. This distinction is morally relevant, but Hauser argues that this and other moral principles are not only culturally universal but also typically opaque (if not completely inaccessible) to the agents through introspection. Thus, the idea is that in the right kind of conditions (e.g., when we face particular types of moral dilemmas), these principles reliably produce moral judgments that are fast, automatic, and non-conscious, in a way that systematically depends on tacit principles that are built-in in our cognitive machinery. In the same way, for instance, we can explain how children reliably make moral discriminations such as those that we see in the literature about the moral/conventional distinction.

I think that a closer look at the parallels between language and moral thinking can help us to deal with the poverty of stimulus argument.[33] For the kind of parallel that I will draw in this chapter illustrates how children can acquire complex normative principles and capacities by deploying similar learning mechanisms to those deployed to acquire many of the complex aspect of language that Chomskian nativists used to think were innate. In particular, I will argue that evidence from developmental psychology largely supports a model of normative development that is in stark contrast with the one proposed by moral grammarians. This model builds upon an alternative linguistic analogy which rejects the idea of an innate universal grammar (see, for instance, Bates & MacWhinney, 1989; Bybee, 1985, 2010; Croft, 2000; Givón, 1979; Goldberg, 2006; Hopper, 1987; Langacker, 1987; Tomasello, 2003). According to this model, children's understanding of the moral/conventional distinction is prototypical and culturally relative because it is a capacity acquired through social learning. That is, we learn from our cultural environments that prototypical cases of moral transgressions have the Turiel's signature moral pattern discussed at the beginning of this dissertation. As I will argue, even core moral principles, such as the principle of double effect, appear gradually and in a piecemeal fashion through mechanisms of intention-reading, pattern-recognition, and affective processing that are deployed in the child's everyday interactions with others.

5.2. The ontogeny of human social norm psychology

In the previous section, I have introduced moral nativism in the version advocated by moral grammarians. For moral grammarians, moral cognition builds upon an innate faculty of syntax-like moral principles. Given this approach, moral grammarians pay relatively little attention to the developmental trajectory of these capacities. Core moral principles are innate and cultural environments set the parameters of this universal moral grammar to the specific values of one's social group. Moral grammarians think that all moral judgments derive from this core moral faculty. But on the view I offer in

---

[33] Of course, the poverty of the stimulus argument could be rejected without embracing any particular model of language development. One could think that a similar model offers an appropriate picture of normative development, even when it does not accurately account the development of language.

this thesis, the capacity to make at least some prototypical moral judgments (i.e., judgments with Turiel's signature moral pattern and judgments about inescapable and authority independent demands) stems from an ontogenetically robust form of norm psychology that builds upon our cognitive and motivational infrastructure for shared intentionality. In sections 5.3 and 5.4, I will argue that this form of norm psychology enables us to entertain shared intentional normative thoughts that can acquire the characteristic features of these prototypical classes of moral judgments through learning and socialization. However, in order to understand how this form of norm psychology gives rise to these prototypical classes of moral judgments we need to understand how this psychology develops in ontogenetic time. As a result, I will focus in this section on explaining my account of the developmental trajectory of our capacity for shared intentional normative thinking—a view that builds upon the ontogenetic account introduced at the end of chapter 4.

Developmental trajectory is important. Discussions about built-in cognitive capacities and acquired ones try to make sense of our distinctive developmental trajectory. But as we have seen in chapter 4, the developmental trajectory of human normative cognition is expected to be highly intricate. Basic forms of shared intentional states initially emerged as ontogenetic adaptations for early childhood in a cooperative breeding context. This is the case of joint attention and declarative pointing. These capacities were subsequently extended in ontogeny because they were co-opted by adjacent developmental stages, including adulthood, since they were plausible useful for coordinating behavior in foraging contexts. When shared intentional states extended into adulthood, these basic skills for shared intentionality meet the selective pressures of adult activities such as collaborative big game hunting. Big game hunting required increased motivation and commitment to shared goals as well as careful monitoring and enforcing of those joint intentions. The coupling of these punitive attitudes with the gradient of generalizability and intrinsic motivation of shared intentional states was the origin of social normative thinking.[34]

---

[34] In small-scale societies, ethnographic evidence suggests that the regulation of cooperation by punishment is infrequent and often low-cost. However, this does not mean that punishment is completely absent, e.g., in the form of ridicule or ostracism (Boehm, 1999; Gurven, 2004). In addition, punitive attitudes as defined in chapter 2 have been reported among forager societies, for they complain or become upset when shares are not returned (Henrich et al., 2006). They also likely play an important role in partner choice, which according to theoretical models help to solve cooperative dilemmas without direct punishment (Aktipis, 2004; Barclay, 2006; Noë & Hammerstein, 1994).

However, I have also argued that selective pressures stemming from collaborative foraging pushed down the development of these social skills into earlier developmental stages. The demands of more complex forms of communication and commitment required increasing preparation and maturation, which favored the emergence of these traits in progressively earlier developmental stages in which collaborative activities are typically low-risk. As a result, the development of human social normative cognition was compressed by both upstream and downstream selective pressures, i.e., those that favored the retention of basic capacities for shared intentionality into later developmental stages and those that favored the earlier emergence of more complex forms of shared intentionality and its distinctive normative dimension. This section will focus on the developmental trajectory of this capacity.

Several studies of cooperation in human children suggest that the above model of the developmental trajectory of normative cognition is roughly right. For example, from very early on children also seem to be much more prone to collaborate and to express sympathetic concern toward others than any other great ape. Experimental evidence shows that 12-month-old children point at an object to help an adult find it (Liszkowski, Carpenter, Striano, & Tomasello, 2006). 18-month-old children show sympathetic concern for others and systematically attempt to comfort them when they get hurt (Eisenberg & Miller, 1987; Vaish, Carpenter, & Tomasello, 2009; Zahn-Waxler, Radkeyarrow, Wagner, & Chapman, 1992). They also understand other people’s goals and help them to achieve those goals, e.g., by getting out-of-reach objects or removing obstacles for them (Warneken & Tomasello, 2006). 2-year-old children are also able to share resources with others even at some cost to themselves (Brownell, Svetlova, & Nichols, 2009; Svetlova, Nichols, & Brownell, 2010). In fact, children are not only more prone to share resources with other but also it has been shown that collaborative activities systematically promote equal sharing, even at the cost of children’s own resources (Hamann et al., 2011). This picture is very similar to the one it has been offered in the previous chapter since it relied on helping and sharing with others in the context of collaborative activities that are mediated by the sharing of goals and attention.

Also, children from 2 years of age seem to display sympathetic concerns for others. Studies on child sympathetic responses typically measure arousal by relative changes in pupil dilation. This might be controversial. But for quite some time, it has been argued that the activity of the sympathetic nervous system correlates with changes in pupil dilation (Loewenfeld & Lowenstein, 1993; Lowenstein, Feinberg, &

Loewenfeld, 1963). In particular, it has been shown that positive or negative situations that require the organism's attention typically trigger the activation of this system (see for example Levenson, 2003). In humans, for instance, even the anticipation of events that are emotionally salient increase pupil dilation (Bradley, Miccoli, Escrig, & Lang, 2008; Nunnally, Knott, Duchnowski, & Parker, 1967; Partala & Surakka, 2003). Unusual social interactions (Gredeback & Melinder, 2010) and representations of impossible physical events (Jackson & Sirois, 2009) increase pupil dilation in children younger than 1 year of age. Sympathetic concern is important because it suggests that children do not display these behaviors out of purely instrumental reasons.

In this line of research, it has been also found that 2-year-old children's sympathetic arousal is similar when they themselves help others and when they see a third party helping others (Hepach, Vaish, & Tomasello, 2012). In both cases, sympathetic arousal has shown to be different from that displayed when the person is not being helped at all. Overall, these results show that young children do not require that they perform the behavior themselves and thus get credit for it in order to be motivated to help others. They seem to have a genuine concern for the welfare of others—children display signs of wanting others to be helped, regardless of whether they themselves are the ones who provide the help. Of course, sympathetic arousal does not necessarily translate in effective behavior and, therefore, it does not mean that we will display prosocial behavior all the time. But it points to an intrinsic motivational component that is characteristic of our shared intentional psychology in collaborative situations that require a modicum of coordination of our mental states.

Importantly, other great apes can share, help, and even display certain sympathetic concerns for other. But at this point in children's development, helping and sharing are supported by a different set of psychological mechanisms. For example, if a 12-month-old child points at an object to help an adult find it, this gesture depends on cognitive and motivational capacities for shared intentionality—more precisely, in the form of joint attention.[35] What the child is doing when pointing is directing the adult's attention to the objects by using a gesture in order to initiate a joint attentional interaction. The gesture is effective insofar as there is a common ground with the adult in the form of a joint attentional frame. The child is not only experiencing the same thing at the same time but also knowing together with the adult that they are

---

[35] It might be worth to emphasize again that attention is an intentional state. Therefore, shared attention is a form of shared intentionality (Carpenter & Call, 2013).

experiencing the same thing at the same time. So, even these very simple activities in which children engage in their early developmental stages are real collaborative activities instead of mere group activities—at least in the sense that social coordination is achieved by a different way to represent the interaction.

This becomes clearer in cases involving role reversal which is facilitated by a bird’s-eye view’ of the social interaction as introduced in chapter 2. There is evidence that 12- and 18-month-olds are able to engage in some basic forms of role reversal, e.g., dyadic, body-oriented role reversal imitation (Carpenter, Tomasello, & Striano, 2005). Very early in development, children then seem to be able to coordinate simple complementary roles with an adult partner. But role reversal requires from the children that they conceive both roles in the same representational format, such that they become interchangeable. They need to attend to both roles simultaneously and then combining both roles into a single mental representation that comprises the whole activity from an external viewpoint, which allows the reversal of roles in tasks with complementary actions (Fletcher et al., 2012). But this is not true of chimpanzees (Fletcher et al., 2012; Tomasello & Carpenter, 2005).

Young children seem also intrinsically motivated to engage in joint activities, which is another important component of shared intentionality. 18-month-olds and 2-year-olds systematically try to reengage the partner in order to continue the activity (instead of continuing the activity by themselves) when a cooperative activity breaks down, e.g., when a partner suddenly stops doing his/her role (Warneken, Chen, & Tomasello, 2006). This is true even when the child’s partner is not necessary anymore to complete the task. A similar behavior is shown even in 14-month old infants who seem to make some effort to reengage passive partners when performing simple coordinated activities with an adult (Warneken & Tomasello, 2007). Thus, children seem to be motivated by the joint activity in itself, rather than using the partner as a mere social tool to achieve their own goal (Warneken et al., 2012).

Moreover, experimental evidence also suggests that children around the age of 3 years honor joint commitments and expect that others do so as well, to the point that they often acknowledge when they have not honored those commitments (Gräfenhain et al., 2009; Hamann et al., 2012). This strongly suggests that from very early on humans develop a form of normative cognition particularly geared to dyadic interactions: 1- to 3-year-old children possess a type of normative cognition based on joint intentions, personal relationships, and social emotions, that makes them able to be fair and to feel sympathetic with particular others. Hence, this developmental stage resembles the first step proposed by advocates of the interdependence hypothesis, which I have argued can be understood as preparation for adult activities due to upstream selective pressures that allowed once typically adult skills to migrate down in human psychological development.

As time progresses, children's normative cognition also changes. By 3 years of age, children seem to coordinate complementary action roles with others through an agent-independent conception of those roles. Shared intentionality enables us to share goals with respect to some external event and being motivated to cooperate in their realization. It allows us to form intentions of the form "We want to do this together". Thus, very early in life, children begin to coordinated and complement action roles—e.g. the adult experimenter holds a box and the infant inserts a toy. They begin to share intentions or action roles in pursuit of a goal in a way that suggests that each partner is aware of both roles as in "You hold the box, I insert the toy", which facilitates mutual help and role reversal.

As explained in chapter 2, a representation is said to be 'agent-independent' if its content does not specify any agent—i.e., if the child understands that those roles can be variably filled by different agents over time. The development of this way to understand social interactions is gradual. While 3-year-old children form and use agent-independent representations of action roles, 4-year-old children may, for example, use these representations flexibly for episodic memory and future deliberation in planning their own and others' actions (Rakoczy, Gräfenhain, Cluver, Dalhoff, & Sternkopf, 2014). The increasingly agent-independent way in which tasks and roles are represented is consistent with the evolutionary account offered in the previous chapter. For on this view, our shared intentional capacities became more complex along a cognitive gradient of generalizability that is supported by an increasing capacity to form agent-independent representations of social roles.

In the previous chapter, I argued that with an increasing capacity to form agent-independent representations, we should also expect to see a more collective form of shared intentionality. Consistently with this, children about 3 years of age seem to begin to be able to understand their social interactions in terms of interactions with groups as opposed to a mere collection of people with various dyadic relationships. Younger children's social responses typically differentiate familiar from strange individuals. But at the age of 3, children start to respond to the groups themselves, based on social cues such as linguistic accent (Kinzler, Dupoux, & Spelke, 2007), physical and behavioral resemblance (Dunham, Baron, & Banaji, 2008). At this age, young children seem to understand that part of being a member of a group is to commit themselves to the ways in which the group does things. They seem to understand that becoming a group member involves accepting the social norms of the group as the evidence I will review in the paragraphs below seem to indicate. Certainly, toddlers follow social norms but they may understand them as mere imperatives. They may agree to the imperative requests of adults but they do not seem to understand social norms and how they work since they do not actively enforce norms until the age of 3 years (Rakoczy, 2008; Rakoczy, Brosche, et al., 2009; Rakoczy et al., 2008; Rossano, Rakoczy, & Tomasello, 2011; Schmidt & Tomasello, 2012; Vaish, Missana, et al., 2011).

There are many prudential reasons to conform to social norms, but 3-year-old children feel compelled to enforce them on others. At that age, children seem to begin to identify themselves with their groups by endorsing the explicit and implicit normative standards of the group and enforcing them selectively to ingroup members only (Schmidt, Rakoczy, & Tomasello, 2012; Schmidt & Tomasello, 2012). For example, recent studies on norm enforcement in young children show that 3-year-olds view norms of conventional games as limited in scope to ingroup members, i.e., children spontaneously protest more against norm transgressions committed by ingroups than outgroups (Schmidt et al., 2012). In total, 9 of 16 children (ingroup condition) versus 3 of 16 children (outgroup condition) uttered forms of protest, such as critique (e.g., “That is not how it is done!”) or normative tattling (“Look, he’s doing it wrongly!”), which were indicative of children’s disapproval of the transgression. However, this parochialism did not extend to prototypical cases of instrumental (e.g., using a music box in the wrong way) or moral transgressions (e.g., destroying other people’s property) where children treat transgressors equally. In all the instrumental tasks in the study, children spontaneously protested against the transgressors about half of the time regardless of the transgressor’s group affiliation, which demonstrates that children were neither indifferent to nor afraid of the outgroup individual.[36] A similar result was obtained in the case of moral transgressions.

Overall, enforcement of group norms seems to fit naturally on the trajectory of increasing group identification. For example, it has also been shown that 4- and 5-year-olds display loyalty to the group by being less likely to reveal the secret in ingroup

---

[36] In the instrumental tasks, the experimenter performed an action such as activating a music box announcing to the child and the puppet “We do it like this!” Then, the child and the puppet could reproduce the action but the latter made an instrumental mistake by, for instance, trying to activate the music box by pushing it instead of turning its crank.

conditions than in outgroup conditions (Misch, Over, & Carpenter, 2016). They remain loyal to their group even when telling the secret would be materially advantageous, e.g., by obtaining additional rewards in the form of stickers. In the context of norms, children as young as 3 years of age seem to recognize the norms of the group and endorse them rather than merely responding to groups based on social cues such as linguistic accent or physical and behavioral resemblance. Once the children are able to identify and enforce the norms of their social groups they are able to make shared intentional normative judgments—i.e., judgments that are characterized by a gradient of generalizability, intrinsic motivation, and punitive attitudes, as explained in chapter 2.

Importantly, young children do not need natural pedagogy (Csibra & Gergely, 2009), normative language, or explicit instructions from adults—a practice that is less common in some societies (Lancy, 1996). Of course, all those cues help the normal development of children’s normative cognition. But they only need to see that the adults apparently expect things to work a certain way to jump to a normative interpretation of an action or activity (Casler, Terziyan, & Greene, 2009; Schmidt et al., 2011a). When an adult individual models an activity and shares the activity with the child, the child not only imitates adult actions but also infers normativity that the child then shares and enforce on others. In the absence of normative language and other pedagogical cues, for instance, experimental evidence shows that young children interpret adult actions as normatively governed by using intention-reading mechanisms that collect social-pragmatic information—e.g., when an adult acts as if he/she recognized novel objects and knew exactly what to do with them . However, children seem to infer normativity from those cues as expressed in their verbal and behavioral protest, critique, or correction (which sometimes includes explicit teaching) making use of normative language as in “No! It does not go like this!” or “You must use this!” (Schmidt et al., 2011a). Children’s normative learning is also not completely blind but rather biased toward adults and reliable role models. This result is expected in a model as the one I have proposed in the previous chapter, which predicts that shared intentional normative thinking is initially deployed in the context of adult interactions rather than children’s peer interaction. For example, it has been shown that from very early on children rely more on subjects that have proved to be reliable in the past (Rakoczy, Warneken, & Tomasello, 2009), but also the pick-up normative cues especially from adults (Rakoczy et al., 2010).

In addition, children also seem to be promiscuous normativists—i.e., individuals prone to make fast generalizations about the normative structure of their social environment. Children often jump to normative conclusions not only without any linguistic cue but also without any intentional or conventional cues. For example, a recent study has shown that 3-year-old children spontaneously infer the presence of social norms even when an adult had done nothing to indicate such a norm in either language or behavior (Schmidt, Butler, Heinz, & Tomasello, 2016). In this study, each child was placed in a room with an experimenter and a puppet. While the puppet sat quietly, the adult would use a collection of objects to play a simple game (e.g., pushing a ball with a hammer or fitting pegs into disks). The experimenters used objects from which the child could infer a conventional purpose (e.g., a human-made object with a hook) and objects that suggested no conventional purpose (e.g., a branch that happened to be usable as a hook). In one condition, the experimenter framed it as a demonstration, making eye contact and asking the child to pay attention. In the other two conditions, the experimenter acted indifferent as they played or pretended to have come up with the game by accident (e.g., by saying “oops” as they performed the action). Later, the puppet asked to participate in the game, but performed an action that was different from the action performed by the adult.

The results of this experiment showed that children were able to infer social norms even from a single intentional action. More strikingly, they protested more (e.g., seizing the objects for a demonstration, calling to the adult to intervene, or using normative language to explain the right way to perform the action) when the action was pedagogical or intentional than when it was accidental, regardless of whether the objects involved in the game had a conventional purpose. In other words, children’s default assumption seems to be that adults’ actions are not random but instances of a normatively structured action type. A similar hypothesis, for instance, has been argued about children’s teleological thinking, i.e., the tendency to view all kind of entities as designed for purposes (Kelemen, 1999). This does not mean, of course, that children have these norms innately instantiated in their brains for all the arbitrary contexts in which apparent normative behavior is displayed, but rather than they are prone to make inferences about the normative structure of adults’ actions—especially when adults display signs of familiarity with the object or situation.

Children’s reactions to violations of norms in game contexts appear to be quite similar to their reactions to violations of prototypical moral norms such as those that prohibit harming others. In both cases, children seem to be intrinsically motivated to follow those norms and they display punitive attitudes. Children also strongly object to norm transgressions of conventional games, and they protest using normative language, e.g., “No! It does not go like this!” (Rakoczy et al., 2008). However, children do seem to differentiate conventional norms from moral norms on the basis of their scope. It has been shown that young children enforce prototypical moral norms equally on all violators while they enforce game norms only on members of their own cultural in-group (Schmidt et al., 2012). Prototypical moral norms are represented by the children as robust, agent-independent norms, whose transgressions require minimal pragmatic cues to be detected—the children would only need to recognize the mere intention of harming others. Toddlers certainly respond when adults tell them things such as, “We don’t hit other children.” But at 3 years of age children can easily generalize these prohibitions in a normative way.

5.3. Norm acquisition

I have provided in the previous section a picture of the developmental trajectory of shared intentional normative thinking. This picture progressively builds upon the infrastructure of our shared intentional capacities beginning with simple capacities for shared attention and declarative pointing to shared forms of collaboration and commitment that trigger normative judgments which children share with other people. When the child is able to understand collective activities by forming shared goals and intentions, they are able to entertain social expectations about the activity that entail shared commitment. Thus, around the 3 years of age, children begin to enforce those normative judgments in themselves and others in a form of shared intentionality which is increasingly collective. For these judgments are not just mere expectations about what one specific partner should do in order to succeed in a certain task but judgments about what any individual should do in order to engage in the activity.

However, although children have the capacity to make normative judgments and to pick up the norms that structure their social environments, they learn these norms from others. In this section, I will sketch a model of norm acquisition that helps us to deal with the poverty of stimulus argument of section 5.1 and makes sense of the ontogenetic trajectory of human social norm psychology explained in section 5.2. In this model, children acquire social norms primarily from their interactions with adults and seemingly competent models. Certainly, they have a capacity for shared intentional normative thinking but they rely on mechanisms of norm acquisition to make those judgments. According to the model of norm acquisition I want to propose, information stemming from social interactions is processed by three systems, each of which plays an important role in the acquisition of social norms as shared intentional states: intention-reading mechanisms, pattern-recognition mechanisms, and affective processing mechanisms (see figure 5.1).

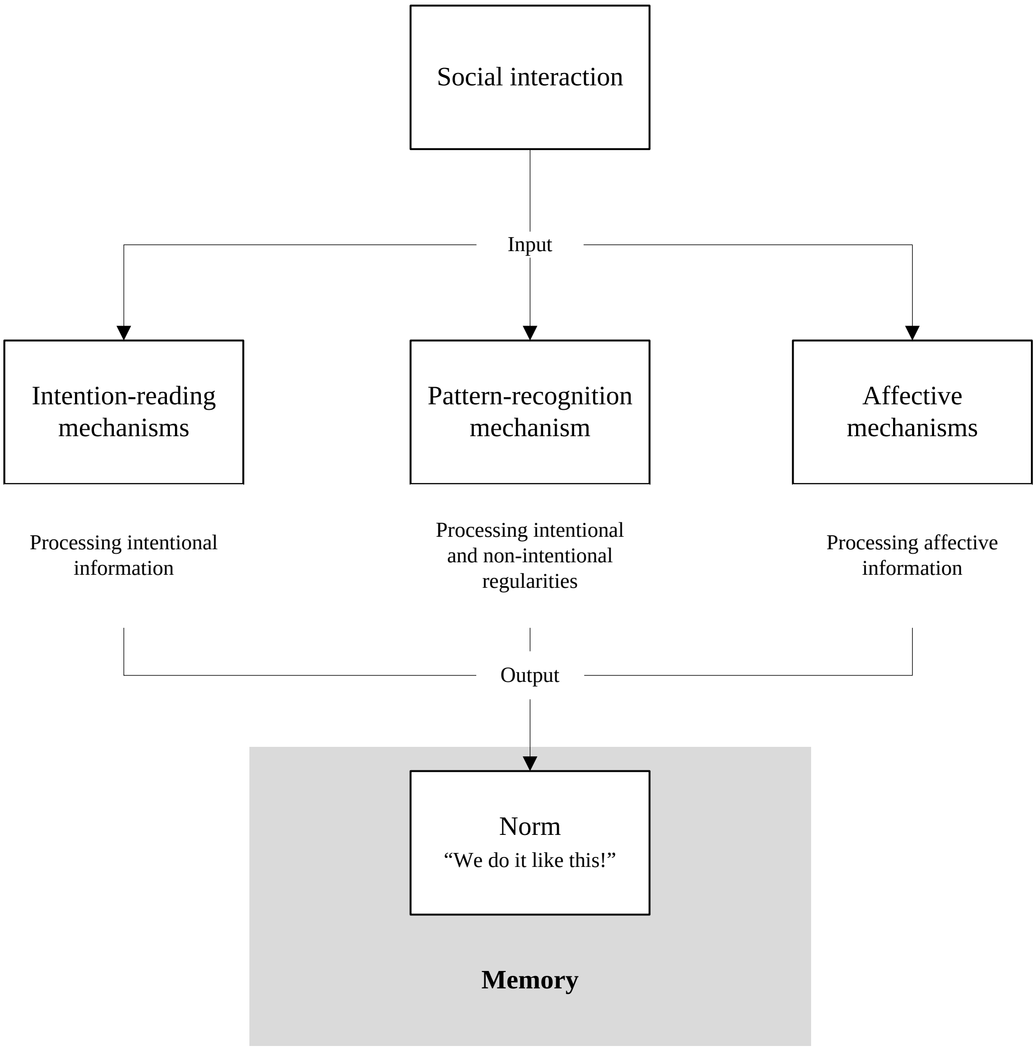

Figure 5.1. Model of norm acquisition. Information stemming from social interactions is processed by a mechanism of intention-reading, pattern-recognition, and affective processing. This information is hypothesized to be integrated somewhere in the brain to form representations of norms that are stored in memory.

This information is conjectured to be integrated somewhere in the brain to form representations of norms that are stored in memory. Intention-reading mechanisms process intentional information. Roughly speaking, they allow the child, for instance, to engage in basic mindreading tasks and to share attention with others. Pattern-recognition mechanisms help to process information about intentional and non-intentional regularities in children's social environments. Pattern-recognition is paradigmatically a cognitive process that allows the child to recognize, for instance, patterns of discontent and disapproval in their local social environments. It helps them to form perceptual and conceptual categories and analogies that are essential to navigate their normative environments. Affective processing mechanisms, in turn, play a crucial but not exclusive role as motivational mechanisms. Roughly, they work as mechanisms of valence in moral decision-making, but also as perceptual biases that allow the children to focus on the right kind of stimuli as well as to communicate the right kind of attitudes that are associated with entertaining normative judgments. Granted, there is an important overlap between some of these mechanisms, especially because empathy-related processes include intention-reading mechanisms. But the categorization of these mechanisms does not need to be clear cut.

This model builds upon usage-based theories of language development since the basic mechanisms deployed for language and norm acquisition are largely the same. The most important difference between these two learning processes is the role that emotional and affective mechanisms play in norm acquisition. For as moral sentimentalists have argued, emotional and affective processing is fundamental for our understanding of moral thinking (Hume, 1740/1978; Smith, 1759/2002). Thus, a hypothesis that requires further investigation is that these processes allow us to understand how some shared intentional normative judgments display the signature moral pattern proposed by Turiel and colleagues (Nichols, 2002, 2004), as well as why they are perceived as inescapable and authority independent (Mameli, 2013), by selectively tuning the affective processing mechanisms involved in norm acquisition and norm execution.

Generally speaking, one advantage of linking normative cognition to the emergence of our shared intentional capacities is that it allows us to provide a more comprehensive view of human uniqueness. For it focuses on the evolution and development of a suite of cognitive and motivational mechanisms that lead to a set of interconnected symptoms that are characteristic of the human syndrome. As a result, it gives us a theoretical ground to speculate about the relationship between normative cognition and language, although certainly not of the kind that moral grammarians have in mind. On this view, human normative competences do not depend upon principles analogous to those of Chomsky’s universal grammar (Chomsky, 1968, 1980, 1986), but rather on domain-general mechanisms which play a central role in building our shared intentional cognitive infrastructure, including at least some important aspects of human linguistic communication. As explained in section 5.1, I will argue that these mechanisms are domain-general in the sense of not being specialized mechanisms for norm acquisition.

5.3.1. Acquiring social norms as shared intentional states

Just as a reminder, on the view I am defending in this thesis, social norms are represented as shared intentional states that are characterized by a gradient of generalizability, intrinsic motivation, and punitive attitudes. Generalizability refers to the scope of the norms and the degree of abstraction in which they are represented. Hominin shared intentional capacities were initially deployed in dyadic interactions between infants and caregivers. But if the model offered in the previous chapter is right, the distinctive normative flavor of shared intentionality likely emerged in social contexts such as collective hunting of big mammals, which is a cognitively taxing activity in terms of planned coordination. Part of the reason why I think this is a crucial part of the story is because hunting big mammals is not a dyadic but a complex collective activity. As a result, this created pressures for modes to represent the activity in a more agent-independent way.

There is an important class of normative judgments are not about particular individuals, but about generic agents placed in generic situations and fulfilling generic roles. Joint activities such as collective hunting were likely much more computationally demanding than other forms of social interactions in that respect—although others activities such as collective defense, scavenging, or even the fire making and maintenance. Recursive mindreading may explain, for instance, the jointness of basic intentional states in early hominins, e.g., the jointness of certain mental states that are shared between infant and caregivers. But they are hardly computationally viable in contexts such as hunting, where decisions sometimes need to be taken quickly, under significant stress and time pressure, and at a high cost for both the individual and the group. Thus, the representational machinery deployed in those contexts should have been characterized by a gradient of generalization and abstraction.[37]

---

[37] The above scenario assumes that contexts like hunting were rather stereotyped and hence there can be norms about who should do what and when. It is difficult to determine to what extent this is a realistic assumption. However, hunting practices are highly stereotyped in chimpanzees (Boesch, 2002) and the evidence for the earliest ambush hunting by hominins and the pattern of animal movements found in the Kenya Rift suggests a similar picture (Kübler, Owenga, Reynolds, Rucina, & King, 2015). In the latter

Intrinsic motivation refers to the motivational force an agent feels when entertaining a normative thought. This is a characteristic feature of shared intentional states even in its most basic form—e.g., in both joint attention episodes and declarative gestures there seems to be an intrinsic motivation to join the attentional states of others and share information with them (see, for instance, Carpenter & Call, 2013; Gómez et al., 1993). We enjoy doing things with others and we naturally enjoy sharing with them attention, believes, desires, and other intentional states. Unlike great apes, we find sharing these states generally rewarding. So, like other shared intentional states, the norm is treated by the agent as an end in itself, rather than as a means to other ends. They exhibit a special motivational force that differs from mere instrumental motivation. The implicit normative mental states that regulated early collective hunting in humans, for instance, should have been somewhat robust—they would have predisposed individuals to stick to certain social standards even when there was little prospect for instrumental gain, reciprocation, or reputation building.

Hunting of big game is dangerous and fairly unpredictable. In those environments, hunting partners need to be skilled, brave, quick decision makers, but most importantly, they need to be trustable and reliable. There would have been plenty of time for thoughtful consideration of prudential reasons before going to hunt. They would have had also many opportunities for partner choice. However, in the midst of a hunting trip, letting others down is not a good decision. Thus, even if they needed to know when to give it up as a bad job, they should have been intrinsically motivated to hunt and act together. It is true that after hours of finding nothing, hunters would need to decide when to quit, and in doing so they would have made an individually reasonable choice. But these decisions would probably have had a similar collective flavor in order to reduce conflict between partners. Hunters would have had a strong commitment that shields them from hurting their long-term interests by unnecessarily shrinking the pool of potential partners or taking the risks of direct reprisal and reputation damage. Hunters should have been very committed to the task even when the chances of being detected falling short of these standards were small.

Of course, intrinsic motivation is not the same as overriding motivation, which means that in some other cases, intrinsic commitment could fail due to instrumental motivation to free-ride, risk aversion, fear, among other reasons. However, if early

---

case, the site location in relation to limited animal routes in the area suggests that hominins were aware of animal movements and used the location for ambush hunting during the Lower to Middle Pleistocene.

hunter-gatherers displayed an intrinsic motivation to comply with the normative standards of the collective activity as I have argued, they would be motivated to meet those standards over and above mere instrumental motivation. Positive emotions such as pride, praise, and admiration would have also played an important role here. Yet the extent to which these emotions play a role in supporting normative-guided behavior is difficult to determine since much of the research in normative cognition has been focused on negative emotions under the assumption that punishing and enforcing norms are the hallmarks of normative thinking. I share this assumption. For as it has been defined in this dissertation, social normative thinking entails a tendency to enforce social norms.

All shared intentional states are characterized by some gradient of generalizability and intrinsic motivation. But not all shared intentional states trigger the kind of punitive or corrective attitudes that are distinctive of the class of normative mental states that were selected for in social context such as collective hunting, i.e., shared intentional normative judgments. These punitive attitudes refer to the particular motivation to punish those (including ourselves) when one think that they have violated a social normative standard. Norm violations typically trigger punitive emotions like anger and guilt (when the norm violation is committed by oneself) and punitive behaviors like criticism, exclusion, or physical harm from most people within a social group, which are specifically directed at norm violators (Brown, 1991; Roberts, 1979; Sober & Wilson, 1998). Thus, punitive attitudes are especially important in the present account of normative thinking because they are an independent feature of shared intentional states. Some shared intentional states involve punitive attitudes, but punitive attitudes are not a distinctive feature of shared intentional states. In contrast, all shared intentional states are more general and abstract than basic I-mode intentions. Also, all shared intentional states are intrinsically motivational in the sense that they bring about a motivational force in virtue of their joint nature.

These punitive attitudes point out to an important aspect of norm acquisition. For in the model I have been developing throughout the thesis, many normative judgments, including prototypical moral judgments, could arise from a particular interplay between agent-independent representations of norms and emotional dispositions. Children may acquire, for instance, representations of norms about behaviors that trigger the disgust system (Kelly, 2011). As we will see in the next section, the disgust mechanism transforms the representation of a shared normative standard into a form of non-conventional normative standard, which is significantly different from the type of representations of conventional norms (e.g., the norms of conventional games) that the child acquires early in life. These emotional dispositions are to a large extent learned. Certainly, affective responses can be understood in some cases as unconditioned responses in the sense that they can occur without any learning or conditioning. From the point of view of Pavlovian conditioning, disgust reactions may be very rigid (gape face, nausea, a sense of oral incorporation or contamination, quick withdrawal, etc.), but the range of elicitors has to remain flexible in order to make the system robust enough to environmental change. Thus, even disgust responses are highly sensitive to mechanisms that adapt these responses to circumstances. Food aversion learning is a typical example. Foods that make an animal sick cease to elicit approach responses when they are encountered again (Garcia, Hankins, & Rusiniak, 1974). Rats and other species can share information about safe foods through odors (Galef, 1996; Galef & Wigmore, 1983). The link between these affective responses and our social environment is even less hardwired.

5.3.2. **The mechanisms of acquisition**

Humans share normative standards in the sense of sharing intentions of the form specified in the previous section. But in the view I want to propose, normative judgments and its distinctive set of emotional dispositions for norm compliance and punitive attitudes emerge in development as a result of the acquisition of shared normative standards through domain-general mechanisms of intention-reading, pattern-recognition, and affective processing—including empathy-related processes such as emotional contagion, perspective-taking, and sympathy (see table 5.1).

Intention-reading mechanisms	**Pattern-recognition mechanisms**	**Affective mechanisms**
• Joint attention—an ability to share attention with others through different sensory modalities • Capacity to follow the attention and gesturing of others • Understanding of communicative intentions—a capacity to actively direct the attention of others to distal objects by using non-linguistic gestures • Cultural learning—a capacity to learn by imitation the intentional actions of others	• Statistical learning • Categorization—a capacity to form perceptual and conceptual categories of similar objects and events • Schema formation—a capacity to form sensory-motor schemas from patterns of perception and action • Analogy—an ability to create analogies based on similar functional roles	• Emotional contagion—a capacity to feel the same emotional states of other individuals • Emotional perspective-taking—the capacity to grasp other individuals' emotional states through effortful, top-down cognitive processes • Sympathy/other-regarding concerns—the ability to entertain an other-oriented emotional reaction such as concern or sorry • Other emotional systems (e.g., the disgust system)

Table 5.1. List of basic cognitive and motivational infrastructure for norm acquisition.

Importantly, the above mechanisms closely resemble a usage-based model of language acquisition (see, for instance, Bybee, 1985; Givón, 1979; Langacker, 1987). I will come back to the relevance of this feature of the model in section 5.5 when discussing the linguistic analogy put forward by moral grammarians. But for the moment it might be worth pointing out that intention-reading skills lead to the emergence of shared intentional states that in turn make possible agent-independent normative cognition while pattern-recognition skills allow the children to find normative regularities in their local environment. This is similar to the role they play in language acquisition since intention-reading mechanisms facilitate the understanding of communicative intentions while skills of pattern-recognition help the child to find the regularities that emerge from the use of linguistic constructions. However, most models of the development of early linguistic competencies differ from the present model of normative development since the latter emphasizes the role of affective processes in both the perception of normative social standards, normative motivation, and the communication of associated behavioral dispositions (e.g., the motivation to punish norm violators). In the proposed model, emotional, empathy-related processes are co-opted to reliably match up the appropriate pro-social responses with the right situation, which is represented by the child in an agent-independent, we-mode way, as his/her shared intentional capacities emerge.

All these mechanisms play an important role in norm acquisition. For example, as pointed out by moral grammarians, automatic normative judgments may implicitly require processing information about the goals, intentions, and consequences of actions (see, for instance, Hauser, 2006b; Hauser et al., 2008a; Mikhail, 2007). Information about intentional facts is processed by our intention-reading mechanisms, not only when acquiring normative standards but also (very likely) when executing them to produce normative judgments. Intention-reading mechanisms start to develop early in life, around 9-12 months of age (Tomasello, 1995) and have been extensively reported in the developmental literature (see Bakeman & Adamson, 1984; Bates & MacWhinney, 1979; Corkum & Moore, 1995; Tomasello, 1998; Tomasello, Kruger, & Ratner, 1993). We have the capacity to share attention with others and to follow, for instance, their visual attention and their pointing toward objects and events. We have built-in skills for understanding the communicative intentions as well as a capacity to direct others' attention through non-linguistic gestures. We also learn by imitation from the intentional actions of others. We are not blind to intentional facts.

In addition, as in the case of language, we are well-equipped to recognize patterns and regularities, not only in speech but also in other domains (Conway & Christiansen, 2001; Gentner & Markman, 1997; Gomez & Gerken, 1999; Marcus, Vijayan, Rao, & Vishton, 1999; Pelucchi, Hay, & Saffran, 2009; Piaget, 1952; Rakison & Oakes, 2003; Ramus, Hauser, Miller, Morris, & Mehler, 2000; Saffran, Aslin, & Newport, 1996; Schneider, 1999). These mechanisms also emerge early in development and seem to be very ancient, for some of them have been found in other primates, e.g., as mechanisms to distinguish individuals of the same species (Yovel & Freiwald, 2013). Pattern-recognition mechanisms include the capacities for statistical learning, categorization, schema formation, and the ability to create analogies would have been also relevant to understand non-linguistic regularities in the social domain. For example, prototype-matching is a mechanism of pattern recognition that describes the process by which a sensory unit registers a new stimulus and compares it to the prototype, or standard model, of the stimulus. They do not require a perfect match between the incoming stimuli and the stored concept in the brain, but they are a highly efficient mechanism to maximize similarity in fast and unconscious ways. For this reason, prototype-matching can play an important role in social cognition since they help us to deal with social stimuli such as faces (Solso & McCarthy, 1981) or personality traits (Niedenthal & Cantor, 1984). These mechanisms would work well given the prototypical nature of moral issues such as inflicting physical harm on others or making unfair distributions of collectively obtained resources.

Each of these mechanisms may act independently, but they often interact with each other to understand our normative environments. Pattern-recognition mechanisms, for instance, interact with other learning mechanisms that help to scaffold normative development. Given the distinctive set of intention-reading mechanisms we have, the regularities that we find in the social domain would not be simple physical or behavioral regularities but rather regularities about intentional, goal-oriented facts. These facts may work as the primary data of children's normative learning, and they can be the input of pattern-recognition mechanisms as well. So, our capacity to recognize basic social and intentional patterns in our daily interactions with others would give us cognitive leverage to make the type generalizations that seem to be puzzling to some (see, for instance, Dwyer, 2007; Hauser, 2006a, 2006b; Hauser et al., 2008a, 2008b; Mikhail, 2007). Pattern-recognition mechanisms would allow us then to make generalizations about what in our social environments would be considered cruel, harmful, unfair, or just simply immoral. Moreover, since a capacity to recognize patterns and regularities in our social environments is often fast, automatic, and non-conscious, these generalizations from particular exemplars and prototypical cases could make intuitive judgments about new cases similarly fast, automatic, and hardly accessible through introspection.

These mechanisms not only help us to acquire norms but also the conceptual building blocks of normative and prototypical moral thinking. The idea that moral concepts and intuitive moral judgments are generalizations from exemplars is not new (Goldman, 1993; Johnson, 1993; Park, 2013; Sterelny, 2010; Stich, 1993). On this view, we would not only produce intuitive judgments involving thick moral concepts, e.g., judgments about what courage, cruelty, truthfulness, and kindness are (Geertz, 1975; Ryle, 1971; Williams, 1985), or perhaps concepts such as harm, justice, or rights that preoccupy social domain theorists, but also about metaethical concepts such as the concept of moral judgment or morality itself. Both kinds of concepts would be acquired through exemplars and paradigmatic cases. If, as discussed in chapter 2, some moral judgments are actually clustered around a gradient of properties such as inescapability and authority-independence (Mackie, 1977), this pattern would also emerge from pattern-recognition mechanisms and its interplay with the intention-reading and affective machinery that is distinctive of our lineage.

Similarly, since these mechanisms support the acquisition of these core concepts, these mechanisms may also help us to judge prototypical moral issues. For example, decisions about how to distribute certain resources may be judged as fair on the basis that they resemble prototypical or exemplary fair decisions. Certainly, the experimental tasks involving distribution of stickers that are common in the empirical literature (Gräfenhain et al., 2013) are not anything like the type of moral issues that adults have to deal with in everyday life, but the structural similarities between peer interactions between children and adult activities may still work as input for our pattern-recognition mechanisms. They facilitate social comparison. Similarly, judgments may be considered more or less similar to paradigmatic cases of moral judgments in our cultural environment. Certainly, the extent to which people explicitly engage in tasks of classification of judgments and norms into moral and nonmoral types is unknown, but there is unequivocal evidence that adult subjects do engage in this cognitive tasks to some extent (Levine, Rottman, Davis, Stich, & Machery, n.d.). In fact, recent studies suggest that starting around 6 years of age children are able to engage in intuitive folk metaethical judgments (Schmidt, Gonzalez-Cabrera, & Tomasello, 2017).

5.3.3. Norm classification

I have argued in the previous section that mechanisms of pattern-recognition, intention-reading, and affective processing support different aspects of norm acquisition such as detecting normative regularities in one’s social environment, learning core normative and prototypical moral concepts, and executing them to make possible judgments about specific moral issues such as fairness. But I have also suggested that they may also help us in cognitive tasks such as categorization in which different classes of norms and normative judgments are distinguished. For categorization is a psychological process in which complex concepts are matched to a target item by checking whether its definitional or prototypical constituents apply to the target. Norm classification is part of the process of norm acquisition. As in the case of meaningful linguistic structures, we do not simply store norms as discrete chunks of information but we also learn to apply the norm when appropriate. We acquire norms in a way that carries information about where, when, and how use these norms. This helps to store information in ways that make memory retrieval fast and efficient, but it also clusters information that helps to categorize norms into, say, moral and nonmoral kinds.

Sometimes the distinction between different types of norms seems to be very implicit. As we have seen in chapter 2, across many cultures children seem to be able to distinguish moral from conventional transgressions from a very young age when these judgments are clustered around the properties proposed by Turiel and colleagues (Nucci & Turiel, 1978; Nucci et al., 1983; Turiel, 1983, 1998), but they seem to do so based on the responses of parents, teachers, and peers (Smetana, 1997). Children acquire the pattern of moralization of and from their social environments. In those cultures, caregivers tend to respond to perceived moral transgressions by drawing the child's attention to the consequences of the transgression. They (sometimes explicitly and sometimes implicitly) link the child's actions to other people's distress, explaining how the victim feels so that the child can understand the suffering inflicted. This may then be used by the parents to inculcate a healthy sense of guilt in their children.

This aspect of norm acquisition takes place not only when the child is the one who commit the transgression. Young children are exposed to songs, fables, and children's stories that praise good behavior and disapprove bad behavior. Storytelling usually exposes the child to their local moral environments and the consequences of actions, and the emergent norms (if not explicit) are more or less stable across a number of scenarios. In contrast, caregivers are more likely to respond to perceived conventional transgressions by reminding the child the type of norms that apply in the context in question (Smetana, 1984, 1985). This process requires the set of mechanisms that I have outline in this section. As in the case language acquisition, it requires understanding adults' communicative intentions, finding the normative regularities that are made salient to the child through storytelling, and perhaps most importantly, understanding the emotional perspective of both the storyteller and the characters in the story.

As I will expand in the next section, the distinction between different types of norms and normative judgments sometimes relies on crucial differences at the level of affective processing mechanisms. Affect is a highly salient aspect of children's social experiences that contributes to the early development of moral thinking and moral concepts. Affective information may be used to understand, differentiate, and remember moral and other types of social events (Arsenio, 1988; Arsenio & Ford, 1985; Arsenio &

Kramer, 1992). Young children tend to judge moral events as affectively negative and conventional transgressions as affectively neutral. Affect ratings are not only highly correlated with judgments about the obligatoriness and alterability of moral events, but also helps to remember those events more easily (Arsenio & Ford, 1985). As a consequence, differences in the tendency of prototypical moral and conventional events to elicit emotional arousal may lead to a differential encoding of those events. Highly arousing transgressions may be considered immoral in part because they are more affectively salient than less arousing events. Elementary school-aged children employ information about situational affective consequences (e.g., whether actors or victims were happy, sad, angry, fearful, or neutral) to correctly infer whether initiating events were moral, conventional, or personal (Arsenio, 1988). In general, social expectations about the emotional consequences of different social situations have been found essential to differentiate moral from conventional transgressions in the social domain tradition, and such expectations have been found to increase in complexity with age (Arsenio & Kramer, 1992).

In many cases at least, children reactions to violations of game rules (Rakoczy, Brosche, et al., 2009; Rakoczy et al., 2008) seem to be very similar to their reactions to transgressions of prototypical moral norms such as those that cause actual harm (Rossano et al., 2011; Vaish, Missana, et al., 2011), but they seem to be able to differentiate these norms in terms of their scope (Schmidt et al., 2012). They apply game rules only in the appropriate contexts and only to the appropriate social group because they can understand that adults expect things to work a certain way (see, for instance, Casler et al., 2009). Children also interact with peers and they have to coordinate and to negotiate resource distribution with them. Some demands would be local and highly context-sensitive while others will be more robust and counterfactually stable. Thus, children may use these cues about context-sensitivity of other children’s protests to determine what transgressions are conventionally wrong in the sense of Turiel and others.

To put it in the context of shared intentionality framework, skills of shared intentionality allow the children to represent joint activities from an agent-independent point of view in which all roles are interchangeable. These skills seem to develop very early in life. For example, after only a few demonstrations of the game (only one demonstration for a 2-year-old child) children seem to have a third-person representation of the interaction that allows them to take either role in the game (Warneken et al., 2006). A capacity to represent social interactions in this agent-independent way is very helpful for the child in order to evaluate their own behavior and to understand that others evaluate them in the same way that they evaluate others. But until that point in their development, child normative cognition is more trained to deal with interactions with caregivers rather than with groups of peers.

The type of shared intentional states that they enforce on others, if they do at all, is basically restricted to particular individuals and, therefore, those shared intentional states are less norm-like. Once social interactions become more collective, children have to face different social challenges that seem to be accompanied by changes in the way they represent their social environments. Some shared intentional states then start to be clearly enforced, and sometimes in context-specific ways, while others end up being robust and stable across different possible social situations.

5.4. Moral sentimentalism

I have begun this chapter by explaining in section 5.1 the view that the capacity for making moral judgments stems from an innate moral faculty in the form of syntax-like moral principles. However, in sections 5.2 and 5.3, I have been trying to build a picture of the development of human social norm psychology and the mechanisms deployed to acquire social norms which differ from the one offered by moral grammarians. On this view, we are able to make shared intentional normative judgments that we are inclined to share with other people in virtue of our shared intentional psychology. But we acquire these norms from our social environment. I have also suggested that these shared intentional normative judgments somewhat resemble the signature moral pattern proposed by Turiel and colleagues, as explained in chapter 2. In this section, I will develop this idea in more detail by arguing that the tuning of the affective processing mechanisms that support shared intentional normative thinking may explain the ramification of this kind of normative thinking into prototypical classes of moral judgments—in particular, judgments with Turiel's signature moral pattern and Joyce's judgments about inescapable and authority independent moral demands.

So far, I have focused on some general capacity for normative thinking via our shared intentional psychology but there are sufficient reasons to think that there is a substantial overlap between the type of normative cognition that the above learning mechanisms scaffold and moral thinking. In particular, the above picture of norm acquisition is largely compatible with one according to which at least an important cluster of prototypical moral judgments, such as judgments with Turiel's signature moral pattern and Joyce's judgments about moral demands, are the result of the execution of affective mechanisms and the cognitive infrastructure of shared intentionality, which supports norm representation. On this view, many of the prototypical aspects of moral norms discussed in chapter 2 (e.g., the inescapability and authority independence of prototypical moral demands) are explained by this affective gradient and the role it fulfills in norm compliance and enforcement. For example, on Shaun Nichol's (2002, 2004) 'norm with feelings' approach, disgusting transgressions are able to produce the signature moral pattern postulated by social domain theorists (Turiel, 1983). Disgusting transgressions tend to be considered less permissible and more serious than affectively neutral transgressions because they carry both the averseness of being transgressive and the additional aversive component of causing disgust. This means that disgusting transgressions activate further our affective mechanisms (in addition to the perceived violation of a norm), which makes us judge those transgressions more seriously.

Similarly, disgusting violations might be considered to be more authority independent or objective. If the relevant authority changes or suspends a norm, then that would suffice to stop judging transgressions to that norm as an offense.[38] However, even if that is the case, disgust reactions to the violation of the suspended norm will continue. These reactions then will make the perceive wrongness of those actions much more robust against authority contingencies. When the norm is suspended, disgust reactions continue being elicited by actions that are perceived disgusting, regardless of the authority. Those reactions will work as error signals, and these error signals will warn us about the violation of the suspended norm as if it were still in place. Normative judgments that are closely associated to disgust cannot be suspended easily. Even if the host of a party decides that spitting in one's glass is OK, the action or the mere idea of the action may still be enough to provoke a disgust reaction.

The view I am proposing is a version of Nichol's theory, but one in which social norms are acquired, represented, and executed as shared, we-mode intentions. Despite these differences, both views share a number of commonalities. In both views, norms lead to moral judgments. They do so via the link between those norms and affective processing. This link is learned and acquired rather than innate. Moreover, the Turiel's signature moral pattern is seen in both views as a byproduct of that link. There is indeed

---

[38] Perhaps in some cases we could still tend to judge these actions as somewhat wrong, e.g., as a result of prolonged conditioning. But the point here is that the judged wrongness of the action will be more authority contingent than in cases where the action is a disgusting violation.

a critical period in child development during which the child's affective mechanisms interact with information about normative prohibitions in the form of shared intentional states to form what Nichols' call 'non-conventional normative theory'—which social domain theorists (wrongly, I think) identify with the whole class of moral judgments.

This critical period begins around the age of 3 years with the emergence of the child's agent-independent normative cognition reviewed in section 5.2 (Tomasello, 2015; Tomasello & Vaish, 2013; Vaish & Tomasello, 2013). Of course, if normative judgments are a form of shared intentional states they are hybrid mental states with both a cognitive and a motivational component—with the motivational component fueled by affective mechanisms. Yet the key developmental difference would be the particular way in which these normative judgments (in the form of agent-independent representations of shared social standards) are wired to our affective mechanisms—e.g., in a way that they trigger the particular punitive attitudes we see in the child. Some norms might be linked to disgust reactions while other might be linked differently. But in all those cases, it is this particular association with our affective mechanisms what explains how agent-independent representations become norms that are spontaneously enforced by the child—even in the robust, authority-independent way that resembles the Turiel's signature moral pattern. In other words, it is this link with our affective mechanisms that explains many of the numerous recipes of paradigmatic moral judgments: the generality, seriousness, and authority independence of Turiel's and colleagues moral responses (Nucci et al., 1983; Turiel, 1983, 1998), and as I will argue in this section as well, Mackie's (1977) inescapability and authority-independence.

An explanation for this signature pattern is important in the context of this chapter because moral nativists have argued that the moral/conventional distinction involves innate, domain-specific information (Mikhail, 2009). More precisely, using the mechanisms of norm acquisition that I have introduced in the previous section, I want to explain why there seems to be a robust moral/conventional distinction. The thesis I have been defending in this dissertation is that while there are reasons to think that the kind of norm psychology that our shared intentional infrastructure affords is deeply entrenched in development, the prototypical pattern of moral judgments I have discussed in chapter 2 is culturally relative and a function of the way in which norms themselves are acquired. As a result, nativism of the kind advocated by moral grammarians should be wrong, which also implies that either moral cognition has no relation whatsoever with linguistic competence or the relation between these two phenomena is different from the one moral grammarians have proposed. But although moral nativist theories have been important and deeply influential in the literature, there are reasons to think that an alternative non-nativist model is preferable. As Nichols (2005) has argued, moral cognition is shaped by innate biases that come in the form of affective mechanisms. But affective systems shape normative cognition in ways that are neither innate nor specific to a certain domain.

Generally speaking, Nichols's views are framed within an important philosophical tradition, known as 'moral sentimentalism', according to which emotions and affective dispositions play a key role in understanding moral thinking. A large number of psychological findings emphasizing the role of emotions in moral thought support this view (see Haidt, 2007; Hauser et al., 2008b; Inbar, Pizarro, Knobe, & Bloom, 2009a; Nichols, 2004, 2005; Pizarro, 2000, 2006; Prinz, 2007, 2009). In general, sentimentalists think that emotions provide the distinctive motivational force that characterizes human moral thinking. More specifically, it has been argued that affective responses to actions with certain sorts of consequences may actually explain the distinction between moral and conventional norms in the context of Turiel's classic paradigm (Nichols, 2002, 2004). Transgressions of norms prohibiting such actions evoke Turiel's signature moral pattern while transgressions of norms governing actions that do not trigger these affective responses fail to evoke such a pattern. In general, experimental evidence indicates that responses to Turiel's dimensions of moral judgments (generality, seriousness, authority independence, and type of justification) are somehow mediated by affective responses (see, for instance, Kelly et al., 2007; Nichols, 2002, 2004; Quintelier et al., 2012; Schnall et al., 2008; Wheatley & Haidt, 2005).

In this view, generality, seriousness, authority independence are aspects of both norm compliance and the punitive attitudes associated to noncompliance. As a result, one should expect emotional and affective phenomena to play a role in norm compliance and in our attitudes toward norm violators. To begin with, although the motivational mechanisms behind norm compliance have not been rigorously studied yet, it is possible to speculate that emotional and affective process play a role in compliance motivation. For example, Chang and colleagues (2011; see also Koban et al., 2013) have shown that social emotions, such as guilt, arise when subjects deviate from social expectations and that they work as 'social error' signals that serve as the primary motivation to conform or adapt behavior. These findings indicate that a particular neural system (previously implicated in expectation processing) plays a critical role in assessing self-reported subjective feelings of guilt that in turn motivate cooperative behavior in experimental situations such as the ultimatum game.

But we have yet to learn how to tune our affective responses in order to motivate behavior in the adaptive direction. What counts as an appropriate response in one context is not the same in another, which means that error signals are also context dependent. Thus, we need to learn when and how much, for instance, we should feel guilty in a certain situation, in order that anticipated guilt would motivate behavior in the adaptive direction. As our social environments become more complex, social roles also become more diverse and specialized. But arguably, in some contexts at least, the mechanisms of norm compliance would need to be more counterfactually robust. For example, normative demands about keeping promises might be more robust across different types of social interactions than normative demands such as those related to appropriate behavior at weddings and funerals. They often have ritual significance and hence have a high affective load. Yet they are also quite context-specific, so displays of extreme joyfulness or sadness are only acceptable depending on particular ritual observances and factors such as social proximity. As a result, we need to develop a more robust motivation across the former type of scenarios such that we are always motivated to keep our promises more or less independently of the specific situation. Similarly, emotional responses would need to be tuned to the perceived severity of norm violations such that the motivation to comply is more reliable when the costs of violating the norm are higher. Therefore, we would need to learn what, when, and where normative demands are higher across these diverse social contexts via the selective tuning of the emotional and affective responses that support norm compliance.

Moreover, emotional and affective phenomena also play an important role in acquiring the punitive attitudes associated to noncompliance. Children’s behavior is clearly normative when they punish other for transgressing those standards, but they have to learn where and how to do it. Social norms play an essential role regulating an individual’s social interactions with others in the general direction of coordinated collaborative activities and punitive attitudes help in this role. There is, of course, a tendency to romanticize the role of norms in fostering human social life. Consequently, there is also a risk of overlooking the negative role that normative guidance plays in organizing our social lives (see, for instance, Greene, 2013; Sterelny, 2012b). Norms fulfill the role of coordination devices but they do not necessarily maximize cooperative outcomes since punishment can also stabilize very maladaptive norms (Boyd & Richerson, 1992). The expansion of symbolic behavior and social norms in humans very likely played a role in entrenching social hierarchy—e.g., by creating a form of tunnel vision (Seabright, 2004). But the stability of these forms of organization in large human groups also relied on institutionalized punishment.

Crucially, empirical evidence also indicates that emotional and affective responses are involved in the generation of a punitive motivation directed at those who violate norms, including oneself. Thus, norm violations are closely linked to the experience of certain affective responses including anger and contempt against norm violators, and the experience of strong motivations to punish the transgressor (Haidt, 2003), or even shame and guilt when one has violated those rules (Mameli, 2013; Vaish, Carpenter, & Tomasello, 2011, 2016). Therefore, emotions seem to play an important motivational role in both norm compliance and normative-based punishment.

Building on the idea that emotional and affective processing play a key role in norm compliance and the acquisition of the punitive attitudes associated with noncompliance, what I would like to argue now is that the selective tuning of one's affective responses could also help us to explain why some shared intentional normative judgments could have a pattern that resembles Turiel and colleague's moral signature pattern. These judgments would give rise to this prototypical class of moral judgments by learning to adjust our emotional responses to perceived transgressions in a way that they support both norm compliance and punitive attitudes toward norm violators. But this special form of norm psychology may also ramify into judgments about prototypical moral demands in the sense of Joyce and others (see Joyce, 2001; Mackie, 1977; Mameli, 2013).

Let me begin with the former. Emotional and affective processing may explain the apparent robustness of the signature patterns found Turiel's moral/conventional studies. In one of his experiments, Nichols (2002) found that transgressions that generate a strong affective response (disgust in this case) are distinguished from those transgressions that are affectively neutral. These transgressions were distinguished along all the four moral/conventional dimensions proposed by Turiel and colleagues. Transgressions that generated strong disgust responses were judged by the subject as less permissible, more serious, less contingent on authority, and they were also more likely to elicit non-conventional justifications in comparison to affectively neutral conventional transgressions. In a second study, Nichols (2002) also found also that subjects with high disgust sensitivity are more likely than subjects with low disgust sensitivity to judge a disgusting violation as serious and authority independent. Thus, the affective responses played an important role in prompting individuals to treat disgusting violations as non-conventional, which indicates that responses to the Turiel's moral/conventional task are mediated by affective responses.

These experiments indicate that the disgust system generates the signature pattern of Turiel's moral responses when subjects are questioned about the permissibility, seriousness, authority contingency and justification of the norms. As a consequence, Nichols' studies provide evidence that moral-like responses to those questions can be induced by affective reactions. In an experiment, Wheatley and Haidt (2005) shown that when participants are hypnotically induced to feel a flash of disgust when they encounter an arbitrary word such as 'often' or 'take', they judge moral transgressions described in vignettes containing those words more severely than in those vignettes in which those words are absent. Similarly, Schnall and colleagues (2008) have shown that participants make more severe moral judgments depending on participants' disgust reactions and sensitivity to their own bodily sensations. These findings suggest that at least some prototypical set of moral judgments may be grounded in affective responses, i.e., disgust is not just an affective factor that influences moral responses but rather the psychological explanation for why these judgments exhibit the Turiel's signature moral pattern in the experimental situation. Disgust influences the perceived severity of the transgression, which in turn is one of the features of the signature moral pattern.

There are reasons to think that distress in others produces intense affective responses. These responses emerge early in development (Nichols, 2001; Zahn-Waxler et al., 1992), and they can appear even in the absence of overt emotional signals through mechanisms of affective perspective taking and other relevant contextual factors (Hepach, Vaish, & Tomasello, 2013; Vaish et al., 2009). In the context of the moral/conventional task, harm scenarios elicit affective responses, and these responses produce judgments with the signature moral pattern described by Turiel. Therefore, it is possible to argue that both norms prohibiting disgusting behavior and norms prohibiting harmful behavior belong to the same class of normative judgments—the so-called class of 'norms with feeling' (Nichols, 2002). In Nichols words: "Violations of norms with feeling are judged as less permissible, more serious, and less dependent on authority than conventional normative violations. In addition, the level of affective response has a significant effect on the extent to which subjects distinguish norms with feeling from norms without" (Nichols, 2002, p. 233).

This body of empirical data also suggests that conventional transgressions often do not evoke the typical signature conventional pattern. Nichols (Nichols, 2002, 2004) studied reactions to a particular class of transgressions that do not involve harm, injustice, or violation of rights, such as rules of etiquette. But norms that prohibit disgusting behavior (e.g., not to snort and spit into the glass of water one is using) evoke one or more of the features of Turiel's signature moral pattern when compared with norms that prohibit non-disgusting behavior (e.g., not to drink soup out of a bowl).

Turiel's results must be interpreted with caution. Nisan (1987) reports that although Kibbutz children and urban secular kids in Israel give the standard responses for both moral and conventional transgressions, in traditional Arab villages all rules (including those that are prototypically conventional) are taken to be authority independent and generally applicable, which are core features of typical moral responses only. Similarly, Haidt (1993) study shows that low socioeconomic status participants from Brazil and the United States judge all conventional transgressions as authority independent and generally applicable. More recently, Kelly and colleagues (2007) have shown that people may judge that norms proscribing harmful actions that are not of the schoolyard variety (which have been systematically used in previous studies) are neither authority independent nor general in scope. Similar results were found when alleged confounds (Fraser, 2012) were removed (Quintelier & Fessler, 2015; Quintelier et al., 2012).

As Kelly and Stich (2008) have shown, the above results then indicate that the regularities found in the moral/conventional task are neither pan-cultural nor as developmentally entrenched as initially thought. This makes it easier to explain the apparent robustness of the moral/conventional distinction via domain-general mechanisms. In these experiments, participants not necessarily exhibit all the features of the respective signature patterns. Also, these response patterns are not reliably evoked by particular types of transgressions. On the one hand, transgressions involving harm, injustice, or violation of rights trigger the signature conventional pattern (Kelly et al., 2007; Quintelier & Fessler, 2015; Quintelier et al., 2012). On the other hand, transgressions that do not involve harm, injustice, or violation of rights can actually trigger the prototypical signature moral pattern (Haidt et al., 1993; Nichols, 2002, 2004; Nisan, 1987).

Moreover, the tuning of emotional and affective responses may also help us to understand how some shared intentional normative judgments ramify in prototypical moral demands that are perceived as inescapable and authority independent, as defined by some philosophers (Joyce, 2006; Mackie, 1977; Mameli, 2013). Normative demands are not uniform. They sometimes require increasing conformity and more stringent punitive attitudes toward norm violators in some particular social contexts more than others. As a result, affective responses have to be tuned to these high demands of norm compliance and punitive attitudes. As we have seen in chapter 2, one can use Mackie's (1977) idea of inescapability and authority-independence to single out a particular class of prototypical moral judgments about moral demands. On this view, there is a privileged class of normative judgments that: "[...] involves a call for action, or for the refraining from action, and one that is absolute, not contingent upon any desire, or preference or policy or choice, his own or anyone else's" (Mackie, 1977, p. 33). More precisely, these demands are inescapable when they are considered as applicable to all the individuals regarding of their ends, desires, or preferences, and authority independent if its normative force is not derived from someone's approval or endorsement. This distinguishes a particular class of prototypical moral demands from mere practical demands such as "Keep promises!" or "Shut the door!" that are expressed in the imperative mood. This class of prototypical moral demands, and only this class, are experienced as inescapable and authority-independent.

Other particular features could also be ascribed to this form of moral demands. Matteo Mameli (2013) has argued, for instance, that judging that something is morally required is characterized not only in terms of their inescapability and authority independence but also in terms of meriting. Normative judgment requires not only certain emotional reactions to violations; it also requires such reactions to be seen as appropriate or deserved. In this way, Mameli argues that there is a distinctive class of normative judgments of the form "*x* is morally required" that imply the deployment of a particular set of emotional dispositions:

(i) A disposition to feel anger or contempt toward those that have not met a practical demand *x*.
(ii) A disposition to feel guilt or shame when one has not met *x*.
(iii) A disposition to feel anger or contempt toward those who do not manifest dispositions (i) and (ii).
(iv) A disposition to feel guilt or shame when one does not manifest dispositions (i) and (ii).

Mameli thinks that it is because of the robustness of dispositions (i)-(iv), and how such robustness plays out in counterfactual thinking, that we feel that those moral requirements are applicable to individuals irrespective of their ends, desires, preferences, and so on. In his view, moral judgments of the form "*x* is morally required" express a practical demand in favor of an action which is seen as persisting even when its violation (or imagined violation) advances the interests of the violators or even when such an action is against the will of a particular authority. This feature of moral judgment is explained in terms of the robustness of dispositions (i)-(iv), i.e., the capacity of these dispositions to trigger those emotions regardless of the agent's beliefs about the ways in which violations can advance the interests of violators or the potential approval or endorsement of these violations by the relevant authorities. For example, we would consider the demand for keeping promises to be inescapable and authority independent because we are robustly predisposed to feel shame when we break a promise and contempt towards those who break their promises across a wide range of counterfactual scenarios involving promises. These dispositions explain why we perceive some demands to be moral, in the prototypical sense of being inescapable and authority independent, since the counterfactual robustness of those dispositions is the hypothetical cause of these psychological features. Similarly, meriting may be explained in terms of the second-order emotions against people who are unresponsive or not sensible enough to the violation of a norm, e.g., a person who does not feel anger or contempt against norm violators.

Moral demands that also include meriting seem to single out a special variety of moral demands that is not cross-culturally robust. For example, there is no ethnographic evidence of second-order punitive attitudes in forager societies (see, for instance, Boehm, 2012, pp. 206-209). So, whether people in these societies experience second-order emotions against people who do not display punitive attitudes is unknown. It could be the case that this variety of moral demand is sensitive to our foraging ecology. Moreover, this variety of moral demand seems to be less developmentally delayed with respect to other features of prototypical moral judgments. For example, although children are able to distinguish between authority-dependent and authority-independent norms from about 3 years of age (Smetana & Braeges, 1990), the available evidence suggests that an understanding of emotions such as guilt begins at around the age of 6 (Nunner-Winkler & Sodian, 1988).

If the above analysis is correct, the selective tuning of one's affective responses help us to explain why some shared intentional normative judgments could ramify into both judgments with the Turiel's moral signature pattern and judgments about prototypical moral demands in the sense of Joyce and others (see Joyce, 2001; Mackie, 1977; Mameli, 2013). More generally, the above results give us further reasons to be skeptical about the idea that moral cognition is a homogeneous domain, as I have argued more extensively in chapter 2. There seems to be no common property (or set of properties) that moral judgments share. As some have argued, morality is a very heterogeneous cognitive domain, and moral judgments only share some kind of ‘family resemblance’ between members (see for example Nado et al., 2009; Parkinson et al., 2011; Sinnott-Armstrong, 2008; Sinnott-Armstrong & Wheatley, 2012, 2013; Taylor, 1978). There are different varieties of moral judgments, some of which follow the pattern found by many in the Turiel’s moral/conventional paradigm or the pattern of inescapability and authority independence that some philosophers ascribe to moral demands.

Certainly, it is possible to argue that there is some broad and generalizable cluster of normative judgments along the lines of Nichols’ theory of norms with feelings (see also Kelly & Stich, 2008; Sripada & Stich, 2007), and within this broader framework we can try to capture some interesting varieties of moral judgments. In this line of argument, we can understand moral sentimentalism as a theoretical attempt to single out particular varieties of moral judgments. In this view, moral sentimentalism may explain some aspects and features of prototypical moral judgments, but they will not necessarily provide an account such that it explains all moral judgments and only them. This is what we should expect from the fragmentationist picture of chapter 2. For prototype- and exemplar-based moral thinking rejects the idea that the domain of the moral has sharp boundaries.

5.5. Language and norms

Following the proposal outlined in chapter 2, I have argued in the previous section that shared intentional normative judgments ramify in distinguishable clusters of prototypical moral judgments such as judgments with Turiel’s signature moral pattern and judgments about inescapable and authority independent moral demands. As explained in chapters 3 and 4, while our shared intentional capacities and its distinctive normative dimension reliably emerge in development due to the crucial role they played in human social evolution, our capacity to make moral judgments is more diverse and culturally labile. At least in some important cases, making moral judgments is a cultural result of the particular ways in which we share shared intentional normative judgments. In particular, I have argued that we can account for the psychological features of these prototypical moral judgments via the strategic tuning of the affective machinery that is associated to this particular form of social norm psychology. However, as I will argue in this section, these mechanisms are domain-general since they are also mechanisms for language acquisition in usage-based approaches in cognitive linguistics. The fragmentation of shared intentional normative judgments into different subclasses of prototypical moral judgments is then learned and acquired, while the mechanisms that support their acquisition are domain-general. As a result, an important chunk of our moral judgments does not necessarily builds upon innate domain-specific information that is recruited in moral development as some prominent moral nativists have argued.

Usage-based models of language acquisition are in stark contrast with Chomskyan models of language acquisition. Similarly, a usage-based model of norm acquisition could also challenge some versions of moral nativism as the one offered by moral grammarians in section 5.1. On this view, young children begin with no moral abstractions at all, or with only very few of those abstractions, e.g., the type of discriminations reported in the Turiel's moral/conventional task (Nucci & Turiel, 1978; Smetana & Braeges, 1990; Turiel, 1983) or the different varieties of moral demands which are of central interests for some philosophers (Joyce, 2006; Mackie, 1977; Mameli, 2013). Core moral concepts such as the concept "*x* is morally required" would also be explained in the same way.

The analogy between usage-based models of language acquisition and norm acquisition is theoretically motivated. In these models, linguistic constructions are meaningful while Chomskyan syntactic structures do not carry meanings themselves. As a result, usage-based models of language acquisition are models of intentional communication. Normative thinking is not just a decision-making mechanism in which we keep for ourselves the normative states that govern individual behavior. We also share those normative states in way that is meaningful to others—a process that played a fundamental role in taking hominins with increase social tolerance and enhanced prosocial skills, as explained in chapter 3, to ancestors able to coordinate social behavior in complex and efficient ways through the sharing of normative mental states, as explained in chapter 4. It is the sharing of these mental states what explains the emergence in our lineage of our distinctive social norm psychology. But while most of this infrastructure might be phylogenetically old and ontogenetically entrenched, the specific normative states that we share are not determined by these psychological mechanisms alone. This infrastructure facilitates cultural transmission. So, one reason to expect this analogy to work is that not only normative thinking but also intentional communication builds upon our shared intentional cognitive infrastructure (Liebal, Behne, Carpenter, & Tomasello, 2009; Moore, Mueller, Kaminski, & Tomasello, 2015; Tomasello, 2008).

Moreover, there are independent reasons why one should think that moral thinking resembles intentional communication more than syntax. If the representations of core moral norms were grammar-like rules, they would not be belief-like mental states, they would be inaccessible for the agent, and rarely the subject of public scrutiny if at all (Sterelny, 2010). While the above model may explain how we may produce fast, automatic, and non-conscious moral judgments in the same way that we can interiorize many of the idiosyncratic features of natural languages in usage-based models of language acquisition, the model of the moral grammarian lacks a similar analogy to explain the interaction between moral grammar and reflective and conscious moral cognition.

Although moral grammarians accept the existence of conscious, explicit, and reflective moral beliefs (see, for instance, Dwyer & Hauser, 2008 as a reply to Dupoux & Jacob, 2007), grammatical competence does not depend on explicit reasoning about norms or beliefs about grammatical correctness—at least not typically among native speakers. Whatever our representations of correct grammatical structures are, they are not belief-like mental states. They are unconsciously computed. As a consequence, syntactic rules and, even more so, the principles of a universal grammar are not the subject of public scrutiny. When they are indeed discussed and examined, it is typically a job for the linguist. In contrast, if the representation of norms and moral beliefs are supported by our shared intentional infrastructure, they are social expectations that are not only entertained by the agent but also intentionally shared. They are the focus of our shared social attention.

Alleged moral universals such as the principle of double effect are very different from what we would expect if they were like, say, Chomsky’s projection principle (Chomsky, 1981). Moral principles can be questioned and they are the subject of further reflection and thought. The principle of double effect may be implicit and cognitively opaque. It can operate quickly and unconsciously in the background of our decision-making process. But these principles are not utterly inaccessible. They can, in fact, be grasped through ordinary language. It was introspection, for instance, what allowed Aquinas (13th c./1975, II-II, Q. 64, Art. 7) to originally formulate the principle when discussing the permissibility of self-defense. The principle was spelled out in the familiar terms of foreseen consequences, intentional actions, and omissions because, if anything, the principle of double effect is a meaningful moral principle. Therefore, it is not surprising that we learn meaningful moral norms in a similar way that we learn meaningful linguistic constructions.

Moreover, a usage-based model of norm acquisition helps us to deal with the poverty of stimulus argument, since it explains how the child may acquire normative principles by deploying a similar set of learning mechanisms as those deployed in the case of language acquisition. The poverty of the stimulus argument typically addresses the impoverished Skinnerian picture of learning and social interaction. But norm and language acquisition is facilitated by a number of much more complex learning mechanisms. Intention-reading, pattern-recognition, and affective mechanisms start to develop early in life. From a cognitive point of view, all these mechanisms are domain-general as they are not specific mechanisms of norm acquisition.

These mechanisms are redeployed in a particular way in order to understand shared normative intentions in a similar way that they are deployed for learning, understanding, and sharing communicative intentions. We are able to join the intentions of others and we share ours with them. We share both a particular we-mode of representation and a particular motivational component. This facilitates human social environments to play a more active role in the successful transmission of culture, including social norms. The poverty of stimulus argument is often formulated as a problem of negative evidence since linguistic strings are not produced and then labeled as ungrammatical. But even if this is the correct picture of language acquisition, children systematically receive negative evidence in the normative domain. As some social domain theorists have argued, the learning environment of children is highly scaffolded. For example, caregivers usually respond to perceived conventional transgressions by reminding the child what norms apply in the situation in question (Smetana, 1984, 1985).

However, the content of norms and their scope could be initially learned from adults and other caregivers without need of pedagogical cues or normative language (Schmidt, Rakoczy, & Tomasello, 2011b) because the shared intentional infrastructure that is built into our intention-reading machinery emerges quite early in ontogeny and is specially prepared for child-adult interactions. This makes the understanding of adult’s social expectations when context-sensitive rules are introduced less demanding (Casler et al., 2009). Children comprehension of norms is more constrained by problems of emotional control than deficits in intention-reading capacities. As we have seen, children’s reactions to violations of game rules (Rakoczy, Brosche, et al., 2009; Rakoczy et al., 2008) are typically similar to their reactions to prototypical moral transgressions, e.g., those that cause harm (Rossano et al., 2011; Vaish, Missana, et al., 2011). At a younger age, children can afford overreacting because social interaction is a highly channelized child-adult interaction. But with time and increasing peer interaction, they have to learn to self-regulate emotions and to tune their affective responses appropriately.

In sections 5.3 and 5.4, I have argued that the kind of norm psychology that I have been discussing in this thesis make us capable of entertaining shared intentional normative thoughts. These normative thoughts can acquire the characteristic features of prototypical classes of moral judgments through learning and socialization. Acquiring norms resembles many aspects of language acquisition since moral norms are meaningful rules rather than abstract, contentless syntactic principles. Certainly, the above model of norm acquisition is independent of any particular usage-based model of language development, but these models help us to understand how the cultural transmission of social norms is possible through the motivational and cognitive infrastructure of our shared intentional psychology. In order to understand how this form of norm psychology gives rise to these prototypical classes of moral judgments I have provided in this thesis a lineage explanation aimed to explain the gradual emergence of this form of psychology and its normative ramifications.

5.6. Conclusion

Nativist approaches are popular in the contemporary debates on human cognition due to the prominence of linguistic nativism. Moral grammarians, in particular, have argued that humans possess a universal moral grammar similar to the one posited by Chomskyan linguistics (Dwyer, 2007; Hauser, 2006b; Hauser et al., 2008a, 2008b; Mikhail, 2007). Drawing a parallel from usage-based models of language acquisition, I have argued in this chapter instead that norm acquisition is the result of domain-general mechanisms of intention-reading, pattern-recognition, and affective processing. This parallel is no accident since usage-based theories of language are developmental approaches to intentional communication which arguably rely on a similar set of cognitive and motivational skills that are distinctive of our cooperative lifestyle. The central idea of the model is that norms are acquired and represented by the agent as shared intentional states. These states are characterized by a gradient of generalization and abstraction, intrinsic motivation, and punitive attitudes. Although shared intentional states are hybrid mental states characterized by a gradient of generalization and intrinsic motivation, only the representation of shared normative mental states are linked to a motivation to protest, complain, and punish those (including ourselves) who do not meet these social expectations. By tuning the affective profile of these mental states, this framework helps us to identify a cluster of prototypical moral judgments such as those reported in social domain tradition (Nucci & Turiel, 1978; Smetana & Braeges, 1990; Turiel, 1983) and different varieties of moral demands which are of central interests for some philosophers (Joyce, 2006; Mackie, 1977; Mameli, 2013). These clusters are ramifications of the cognitive lineage of our distinctive capacity for shared intentionality and shared intentional normative judgments, but norms that generate those judgments are acquired through social learning.

Chapter 6. Conclusion

The goal of this thesis was to explain and defend a naturalistic theory of normative cognition—in particular, human social norm psychology. More specifically, I aimed to provide a lineage explanation of this form of normative thinking, i.e., an explanation that specifies a sequence of changes that takes us from agents with an ape-like baseline capacity for social cognition to agents with human-like social norm psychology. The explanation I offer in this dissertation relied on a model of great ape and human cooperation that has been built over the last two decades of psychological research within the theoretical framework of shared intentionality. I argued that human social norm psychology is closely linked to our shared intentional psychology and that this capacity is in turn connected to prototypical moral judgments such as judgments with Turiel's signature moral pattern and judgments about inescapable and authority independent moral demands.

6.1. Summary

As explained in chapter 1, the goal of this thesis was to explain and defend a naturalistic theory of normative cognition—in particular, a special form of social norm psychology whose evolutionary lineage can be traced back to the origins of our capacity for sharing intentional mental states. For this purpose, I addressed issues about the motivational and cognitive machinery of normative thinking as well as questions related to their ontogenetic and evolutionary trajectory. More precisely, the thesis aimed to provide a lineage explanation of this central form of normative thinking. In other words, it aimed to specify a sequence of changes that takes us from agents with an ape-like baseline capacity for social cognition to agents with some capacity for shared normative guidance. First, I tried to show how the emergence of our distinctive capacity to follow norms and make normative judgments is genealogically linked to the evolutionary lineage of our capacity to form shared intentional states. Second, I tried to show how this capacity is related to a diverse cluster of prototypical moral judgments such as judgments with Turiel's signature moral pattern and judgments about inescapable and authority independent moral demands. Moreover, I argued that in explaining the evolution of normative cognition we also require an understanding of the developmental trajectory of this cognitive capacity since normative thinking is to a significant extent the result of developmental tuning of our shared intentional capacities.

Having outlined the general approach of the thesis, I argued in chapter 2 that moral judgments define a quite heterogeneous class of mental states. The best theory that we have to assume that moral judgments are a unified natural kind in psychology, i.e., Turiel's social domain theory of moral development (Nucci & Turiel, 1978; Nucci et al., 1983; Turiel, 1983, 1998), fails in its attempt. There seems to be no property (or set of properties) that unifies moral thinking as an explanatory target for psychology. Instead, we should understand particular accounts of moral judgments as targeting particular classes of judgments that are prototypically moral. For example, judgments that are inescapable and authority independent pick out a class of normative judgments that is important for independent philosophical reasons (Joyce, 2001; Mackie, 1977; Mameli, 2013), although not all moral judgments are judgments about moral demands in the sense of being inescapable and authority independent (Joyce, 2006, 2014). Similarly, not only moral demands are inescapable and authority independent, e.g., what counts as the correct solution for a mathematical equation can also be judged as inescapable or authority independent as prototypical moral demands. I argued that judgments that evoke Turiel's signature moral pattern or judgments about inescapability and authority-independent demands are, at best, prototypically moral. This does not imply the collapse of the psychological investigation of moral judgments, but rather that moral judgments cannot be investigated as if they were natural kinds or a single unified psychological domain. We could, instead, explain in what sense a certain class of normative judgments is prototypically moral as well as why and how they have the family resemblance they share by, for instance, linking those judgments (which may be culturally relative) to the lineage of more robust and developmentally entrenched psychological mechanisms. I ended this chapter by proposing that we could link those normative judgments to the lineage explanation of the motivational and cognitive infrastructure of our shared intentional psychology. Shared intentionality can help us to define a special kind of social norm psychology that is developmentally robust but which ramifies into different branches of prototypical moral judgments due to the diversification of our cultural practices and how we learn them.

Following the above proposal, I focused in chapter 3 on the reconstruction of our ape-like baseline capacity for social cognition, which will serve as the starting point of my hypothesized lineage explanation. I began explaining how debates about the reconstruction of early hominins can affect philosophical debates about normative and moral psychology as well as what the perils of reconstructing the social behaviors of these ancestors are. Then, I have argued that we are more justified in using an alternative model of the *Pan*/*Homo* LCA based on a mosaic hypothesis of human origins. According to this hypothesis, this ancestor would have been a large mosaic of traits seen in many primate species, including chimpanzees and bonobos. In order to support this model, I focused especially on comparative studies between humans and both *Pan* species, and discussed whether these similarities are traits inherited from a common ancestor, traits evolved independently due to similar selective pressures, or (as I think is more likely) a combination of both—i.e., with some traits inherited from the common ancestor while others being the result of parallel evolution. In particular, I argued that this alternative model does not fit well with some prominent views of human evolution such as demonic male view (Wrangham & Peterson, 1996) or the killer ape hypothesis (Dart, 1953), which have been very influential in the philosophical literature. Instead, they support a model of early hominins characterized by a level of social tolerance and prosocial skills that goes beyond these views. Taking as an example Kitcher's (1998, 2006, 2011) account of the evolution of our capacity for normative guidance, I suggested that a better explanation of the shift toward a top-down capacity to understand and respond to commands could be framed in terms of shared intentionality (Tomasello, 2015; Tomasello & Carpenter, 2007; Tomasello et al., 2005).

I developed further this idea in chapter 4. I argued that the above model of the *Pan*/*Homo* LCA is still compatible with the idea that early hominins hunted collaboratively in groups (see also Pickering, 2013). This model not only provided a cognitive baseline for my lineage explanation of normative guidance but also it helps us to explain the gradual expansion of early hominin cooperative capacities since collaborative foraging would be implausible in a very aggressive and dominant lineage. The evolution of collaborative foraging is closely linked to the evolution of our shared intentional capacities. According to the interdependence hypothesis of the origins of shared intentionality, for instance, collaborative foraging, especially in the form of collective hunting, was the main driver in the evolution of shared intentionality (Tomasello et al., 2012). However, this explanation leaves open the question of why these capacities emerge so early in ontogeny. For this reason, I argued that the initial driving force in the evolution of our share intentional capacities was cooperative breeding (see Hawkes, 2012; Hawkes, 2014) while the distinctive normative dimension of shared intentionality emerged as a consequence of the selective pressures on collaborative hunting. In this view, obligate collective hunting requires shared intentional states that regulate collective activities in a distinctive normative way since these mental states are characterized by a gradient of generalizability, intrinsic motivation, which is common to all shared intentional states, but also a distinctive cluster of punitive and corrective attitudes (see Sripada & Stich, 2007). At the end of this chapter, I proposed an evolutionary model of shared intentionality that takes into account the tradeoffs that emerge during the transition from childhood to adulthood in order to integrate insights from both the interdependence hypothesis and the cooperative breeding hypothesis (Tomasello & Gonzalez-Cabrera, 2017).

Expanding on the above model, I argue in chapter 5 that while the basic cognitive and motivational infrastructure underlying shared intentionality is highly entrenched in development, we acquire social norms from our cultural environments through this infrastructure. In particular, I argued that norm acquisition is the result of domain-general mechanisms of intention-reading, pattern-recognition, and affective processing. This model of norm acquisition builds upon a parallel with usage-based models of language acquisition. This parallel is no accident since these models of language acquisition are developmental approaches to intentional communication, which arguably rely on the same motivational and cognitive infrastructure that supports the distinctive cooperative lifestyle of our species. On this model, the agent learns and represents norms as shared intentional states. As other intentional states of this kind, shared normative mental states are hybrid mental states that possess a gradient of generalization and intrinsic motivation, but which are distinctively linked to a set of punitive attitudes toward those (including ourselves) who do not meet these social expectations. I argued that differences in the affective profile of these mental states help us to account for a cluster of prototypical moral judgments such as those reported in the social domain tradition (Nucci & Turiel, 1978; Smetana & Braeges, 1990; Turiel, 1983) and a prototypical form of moral demands which has been of central importance for some philosophers (Joyce, 2006; Mackie, 1977; Mameli, 2013).

6.2. Limitations and future research

Admittedly, as a direct consequence of its methodology, this dissertation encounters a number of limitations that are unknown to other forms of philosophical understanding and which need to be considered here in some length. The most obvious ones are the evidential issues that come with historical approaches. Since the connection between evidence and distant events erodes with time, reconstructing the past is particularly difficult. Evidence is scarce, and it tends to disappear as history unfolds. Reconstructing the social-cognitive lineage of our normative capacities just takes these issues a step further since they only can be inferred indirectly from this physical evidence. As a result, a reconstruction of the deep hominin ancestry of our social-cognitive capacities requires serious evidential bootstrapping, for no single strand of evidence can tell us the complete story about our hominin past.

In this thesis, I have compared closely related species such as chimpanzees and bonobos in order to provide information about the baseline cognitive capacities of our early ancestors. Even unrelated taxa can offer valuable information about our evolutionary past through parallel, convergent evolution if enough similarity between the respective selective environments is guaranteed. This type of parallel is indeed more difficult to assess given the obvious differences between, say, dogs, macaques, and humans. But although risky, these inferences are not hopeless. For example, I have argued that the retention of juvenile traits such as playfulness and social tolerance could be explained by a parallel process of self-domestication, while cooperative breeding makes relevant the comparison between humans and callitrichid monkeys. Well-crafted comparative analyses within different lineages can yield meaningful conclusions, even if they do not solve all the problems of scarcity and low resolution of evidence that are common to evolutionary reconstructions. Future research will benefit from the inclusion of broader comparative analyses.

The paleoanthropological record is also typically scattered. Although techniques to read the past from ancient cut marks in animal bones can be very informative about the origins of butchery, inferences about how hominins procured meat still need to be taken with a grain of salt. It is difficult to assess, for instance, to what extent there was division of labor or role interchangeability when the first evidence of butchering emerge in the fossil record. It is also difficult to evaluate with the available data whether meat was procured through scavenging or such-and-such form of hunting. Yet the fossil record is still crucial to determine the significance of the similarities and differences we find between closely related species. As pointed out in the introduction of this thesis, although all the three genera predating the Lucy specimen, i.e., *Sahelanthropus*, *Orrorin*, and *Ardipithecus*, have been found in the past 20 years, they have crucially changed the way we think about the origin of many distinctive human traits such as bipedalism, diet, and sexual dimorphism. Thus, future investigation should focus on more detailed evolutionary scenarios of hunting such as those linked to the ability to throw projectiles and endurance running, which could significantly benefit from these comparative analyses. For example, while great apes occasionally throw objects with a fair degree of accuracy, only humans can routinely do it with both speed and accuracy (Roach, Venkadesan, Rainbow, & Lieberman, 2013). Certainly, some of the anatomical features required for this are already present in australopithecines and other early hominins, but the entire suite of traits only seem to have appeared about 2 mya in *Homo erectus*, around the same time that the features linked to endurance running emerge in the hominin fossil record.

All this evidence was framed in the context of a particular view of the evolution of human sociality and cooperation that capitalized upon the idea of shared intentionality. Although there are many unresolved issues regarding the cognitive architecture and implementation of this capacity (e.g., to what extent shared intentional states are reducible to recursive mindreading capacities), shared intentionality has proved to be a functionally relevant distinction in contemporary developmental psychology. Part of its success relies on its capacity to explain and predict differences between great ape and human cooperation. But another is its peculiar and well-document developmental pattern, which branches into various skills for sharing different kinds of mental states. I have argued in this thesis, for instance, that part of this developmental process is the emergence of a distinctive capacity for shared intentional normative cognition and that this capacity emerged as a consequence of shifts in hominin life history and selective tradeoffs between different developmental stages. Future work should look at these issues of cognitive architecture and implementation, as well as how they support or undermine the tradeoffs proposed by the evo-devo account of shared intentionality proposed in chapter 4.

It remains controversial, however, when and to what extent our hominin ancestors engaged in shared intentional interactions. In this respect, the argument proposed in this dissertation relies on several assumptions about hominin cognitive competence and the likelihood that the social complex behaviors such as cooperative breeding and collaborative foraging would have emerged without the scaffolding of a capacity for sharing intentional mental states—in particular, the capacity for sharing social expectations about what others should or should not do in a particular situation. Further research should not only pay attention to the advancement of fields such as paleoanthropology and comparative genetics but also to the empirical study of the cognitive demands that these social behaviors presuppose as well as the developmental trajectory that these skills have across different ecological and cultural environments.

Finally, the literature in normative psychology has been traditionally focused on first-order normative cognition, while folk metaethical thinking has been less studied. We not only judge things to be right or wrong, but we also think about these judgments themselves. Suppose that two individuals are discussing whether the Earth is flat. They may think that there is only one correct answer to this question since it is a fact of the matter whether the Earth is flat or not. But would they think the same about moral issues? A number of moral philosophers have endorsed the claim that ordinary people have objectivist intuitions regarding these issues. Yet recent experimental evidence suggests that people's judgments about the objectivity or relativity of moral issues depend on factors such as subjects' age, personality traits, and social proximity. Moreover, it has been argued that children tend to be moral objectivists, while later in development people develop more relativistic intuitions (Knobe et al., 2012). However, this developmental trajectory has been inferred without applying the same experimental paradigms applied to adult subjects. Future research should look carefully at folk metaethical thinking in adults and children, incorporating these findings into a more comprehensive framework of moral psychology.

References

Aktipis, C. A. (2004). Know when to walk away: contingent movement and the evolution of cooperation. *Journal of Theoretical Biology, 231*(2), 249-260. doi:10.1016/j.jtbi.2004.06.020

Alberts, J. R., & Gubernick, D. J. (1984). Early learning as ontogenetic adaptation for ingestion by rats. *Learning and Motivation, 15*(4), 334-359. doi:10.1016/0023-9690(84)90002-X

Alexander, R. D. (1987). *The biology of moral systems*. Hawthorne, N.Y.: A. de Gruyter.

Altmann, S. A., & Altmann, J. (1970). *Baboon Ecology: African Field Research*. Chicago, IL: University of Chicago Press.

Aquinas, T. (13th c./1975). Summa Theologica II-II, Q. 64, “Murder” (Vol. 3). New York, NY: Catholic Way Publishing.

Ariew, A. (1996). Innateness and Canalization. *Philosophy of Science, 63*, S19-S27.

Ariew, A. (1999). Innateness is canalization: In defense of a developmental account of innateness. In V. G. Hardcastle (Ed.), *Where biology meets psychology: philosophical essays* (pp. 117-138). Cambridge, MA: MIT Press.

Ariew, A. (2007). Innateness. In M. Matthen & C. Stevens (Eds.), *Philosophy of Biology* (pp. 567-584). Amsterdam: Elsevier.

Arifeen, S., Black, R. E., Antelman, G., Baqui, A., Caulfield, L., & Becker, S. (2001). Exclusive breastfeeding reduces acute respiratory infection and diarrhea deaths among infants in Dhaka slums. *Pediatrics, 108*(4), e67-e67.

Arsenio, W. F. (1988). Children’s Conceptions of the Situational Affective Consequences of Sociomoral Events. *Child Development, 59*(6), 1611-1622. doi:10.1111/j.1467-8624.1988.tb03688.x

Arsenio, W. F., & Ford, M. E. (1985). The Role of Affective Information in Social-Cognitive Development: Children’s Differentiation of Moral and Conventional Events. *Merrill-Palmer Quarterly, 31*(1), 1-17.

Arsenio, W. F., & Kramer, R. (1992). Victimizers and their victims: Children’s conceptions of the mixed emotional consequences of moral transgressions. *Child Development, 63*(4), 915-927.

Bakeman, R., & Adamson, L. B. (1984). Coordinating Attention to People and Objects in Mother-Infant and Peer-Infant Interaction. *Child Development, 55*(4), 1278-1289. doi:10.2307/1129997

Baker, M. C. (2001). *The atoms of language* (1st ed.). New York, NY: Basic Books.

Banerjee, R. (2002). Children's Understanding of Self-Presentational Behavior: Links With Mental-State Reasoning and the Attribution of Embarrassment. *Merrill-Palmer Quarterly, 48*(4), 378-404. doi:10.1353/mpq.2002.0015

Bar-Yosef Mayer, D. E., Vandermeersch, B., & Bar-Yosef, O. (2009). Shells and ochre in Middle Paleolithic Qafzeh Cave, Israel: indications for modern behavior. *Journal of Human Evolution*, *56*(3), 307-314. doi:10.1016/j.jhevol.2008.10.005

Barclay, P. (2006). Reputational benefits for altruistic punishment. *Evolution and Human Behavior, 27*(5), 325-344. doi:10.1016/j.evolhumbehav.2006.01.003

Bard, K. A. (2012). Emotional engagement: how chimpanzee minds develop. In F. B. M. de Waal & P. F. Ferrari (Eds.), *The primate mind: built to engage with other minds* (pp. 224-245). Cambridge, MA: Harvard University Press.

Bard, K. A., & Vauclair, J. (1984). The communicative context of object manipulation in ape and human adult-infant pairs. *Journal of Human Evolution, 13*(2), 181-190. doi:10.1016/S0047-2484(84)80062-7

Baron-Cohen, S. (1995). *Mindblindness: an essay on autism and theory of mind*. Cambridge, MA: MIT Press.

Basile, B. M., Hampton, R. R., Suomi, S. J., & Murray, E. A. (2009). An assessment of memory awareness in tufted capuchin monkeys (*Cebus apella*). *Animal Cognition, 12*(1), 169-180. doi:10.1007/s10071-008-0180-1

Bates, E., & MacWhinney, B. (1979). The functionalist approach to the acquisition of grammar. In E. Ochs & B. B. Schieffelin (Eds.), *Developmental pragmatics* (pp. 167-221). New York, NY: Academic Press.

Bates, E., & MacWhinney, B. (1989). Functionalism and the Competition Model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 3-73). Cambridge, UK: Cambridge University Press.

Baumard, N. (2016). *The origins of fairness: how evolution explains our moral nature*. Oxford: Oxford University Press.

Baumard, N., André, J. B., & Sperber, D. (2013). A mutualistic approach to morality: The evolution of fairness by partner choice. *Behavioral and Brain Sciences, 36*(1), 59-78. doi:10.1017/S0140525x11002202

Beran, M. J., Smith, J. D., Coutinho, M. V., Couchman, J. J., & Boomer, J. (2009). The psychological organization of "uncertainty" responses and "middle" responses: a dissociation in capuchin monkeys (*Cebus apella*). *Journal of Experimental Psychology Animal Behavior Processes, 35*(3), 371-381. doi:10.1037/a0014626

Berge, C., & Penin, X. (2004). Ontogenetic allometry, heterochrony, and interspecific differences in the skull of African apes, using tridimensional Procrustes analysis. *American Journal of Physical Anthropology, 124*(2), 124-138. doi:10.1002/Ajpa.10333

Binford, L. R. (1981). *Bones: ancient men and modern myths*. Orlando: Academic Press.

Binford, L. R. (1987). Searching for camps and missing the evidence? Another look at the Lower Paleolithic. In O. Soffer (Ed.), *The Pleistocene old world: regional perspectives* (pp. 17-31). New York, NY: Plenum.

Blair, R. J. R. (1995). A cognitive developmental approach to morality: investigating the psychopath. *Cognition, 57*(1), 1-29. doi:10.1016/0010-0277(95)00676-P

Blair, R. J. R. (1996). Brief report: Morality in the autistic child. *Journal of Autism and Developmental Disorders, 26*(5), 571-579. doi:10.1007/BF02172277

Blair, R. J. R. (1997). Moral reasoning and the child with psychopathic tendencies. *Personality and Individual Differences, 22*(5), 731-739. doi:10.1016/S0191-8869(96)00249-8

Blair, R. J. R. (2007). The amygdala and ventromedial prefrontal cortex in morality and psychopathy. *Trends in Cognitive Science, 11*(9), 387-392. doi:10.1016/j.tics.2007.07.003

Blair, R. J. R. (2008). The amygdala and ventromedial prefrontal cortex: functional contributions and dysfunction in psychopathy. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*(1503), 2557-2565. doi:10.1098/rstb.2008.0027

Blair, R. J. R., & Cipolotti, L. (2000). Impaired social response reversal: A case of ‘acquired sociopathy’. *Brain, 123*, 1122-1141.

Blair, R. J. R., Monson, J., & Frederickson, N. (2001). Moral reasoning and conduct problems in children with emotional and behavioural difficulties. *Personality and Individual Differences, 31*(5), 799-811. doi:10.1016/S0191-8869(00)00181-1

Blumenschine, R. J. (1986). *Early hominid scavenging opportunities: implications of carcass availability in the Serengeti and Ngorongoro ecosystems*. Oxford, England: B.A.R.

Blumenschine, R. J., Cavallo, J. A., & Capaldo, S. D. (1994). Competition for Carcasses and Early Hominid Behavioral Ecology: a Case-Study and Conceptual-

Framework. *Journal of Human Evolution, 27*(1-3), 197-213. doi:10.1006/jhev.1994.1042

Blurton-Jones, N. G. (1987). Tolerated Theft, Suggestions About the Ecology and Evolution of Sharing, Hoarding and Scrounging. *Social Science Information Sur Les Sciences Sociales, 26*(1), 31-54.

Blurton-Jones, N. G., Hawkes, K., & O’Connell, J. F. (1999). Some current ideas about the evolution of the human life history. In P. C. Lee (Ed.), *Comparative primate socioecology* (pp. 140-166). Cambridge: Cambridge University Press.

Bobe, R., & Behrensmeyer, A. K. (2004). The expansion of grassland ecosystems in Africa in relation to mammalian evolution and the origin of the genus *Homo*. *Palaeogeography, Palaeoclimatology, Palaeoecology, 207*(3-4), 399-420. doi:10.1016/j.palaeo.2003.09.033

Boehm, C. (1999). *Hierarchy in the forest: the evolution of egalitarian behavior*. Cambridge, MA: Harvard University Press.

Boehm, C. (2007). *Conscience origins, sanctioning selection, and the evolution of altruism in Homo Sapiens*. Department of Anthropology. University of Southern California.

Boehm, C. (2008). Purposive Social Selection and the Evolution of Human Altruism. *Cross-Cultural Research, 42*(4), 319-352. doi:10.1177/1069397108320422

Boehm, C. (2012). *Moral origins: the evolution of virtue, altruism, and shame*. New York, NY: Basic Books.

Boesch, C. (1994). Cooperative hunting in wild chimpanzees. *Animal Behaviour, 48*(3), 653-667. doi:10.1006/anbe.1994.1285

Boesch, C. (2002). Cooperative hunting roles among Taï chimpanzees. *Human Nature, 13*(1), 27-46. doi:10.1007/s12110-002-1013-6

Boesch, C., & Boesch-Achermann, H. (2000). *The chimpanzees of the Taï Forest: behavioural ecology and evolution*. Oxford: Oxford University Press.

Boesch, C., & Boesch, H. (1989). Hunting behavior of wild chimpanzees in the Taï National Park. *Am. J. Phys. Anthropol., 78*, 547.

Boesch, C., Hohmann, G., & Marchant, L. F. (2002). *Behavioural diversity in chimpanzees and bonobos*. Cambridge: Cambridge University Press.

Botterill, G., & Carruthers, P. (1999). *The Philosophy of Psychology*. Oxford: Cambridge University Press.

Bowles, S. (2008). Being human: Conflict: Altruism’s midwife. *Nature, 456*(7220), 326-327. doi:10.1038/456326a

Bowles, S. (2009). Did Warfare Among Ancestral Hunter-Gatherers Affect the Evolution of Human Social Behaviors? *Science, 324*(5932), 1293-1298. doi:10.1126/science.1168112

Bowles, S., & Gintis, H. (2011). *A Cooperative Species: Human Reciprocity and Its Evolution*. Princeton, NJ: Princeton University Press.

Boyd, R. (1991). Realism, Anti-Foundationalism and the Enthusiasm for Natural Kinds. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition, 61*(1/2), 127-148. doi:10.2307/4320174

Boyd, R. (1999). Homeostasis, species, and higher taxa. In R. A. Wilson (Ed.), *Species: New Interdisciplinary Essays* (pp. 141-185): MIT Press.

Boyd, R., & Richerson, P. J. (1992). Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups. *Ethology and Sociobiology, 13*(3), 171-195. doi:10.1016/0162-3095(92)90032-Y

Brace, C. L. (1972). Sexual dimorphism in human evolution. *Yearbook of Physical Anthropology, 16*, 31-49.

Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology, 45*(4), 602-607. doi:10.1111/j.1469-8986.2008.00654.x

Bramble, D. M., & Lieberman, D. E. (2004). Endurance running and the evolution of *Homo*. *Nature, 432*(7015), 345-352. doi:10.1038/Nature03052

Bratman, M. (1992). Shared Cooperative Activity. *Philosophical Review, 101*(2), 327-340. doi:10.2307/2185537

Breedlove, S. M., Watson, N. V., & Rosenzweig, M. R. (2010). *Biological psychology: an introduction to behavioral, cognitive, and clinical neuroscience* (6th ed.). Sunderland, MA: Sinauer Associates, Inc. Publishers.

Breuer, T., Ndoundou-Hockemba, M., & Fishlock, V. (2005). First observation of tool use in wild gorillas. *PLoS Biology, 3*(11), 2041-2043. doi:10.1371/journal.pbio.0030380

Brosnan, S. F. (2006). Nonhuman Species’ Reactions to Inequity and their Implications for Fairness. *Social Justice Research, 19*(2), 153-185. doi:10.1007/pl00022136

Brown, D. E. (1991). *Human universals*. Philadelphia: Temple University Press.

Brownell, C. A., Svetlova, M., & Nichols, S. (2009). To share or not to share: When do toddlers respond to another’s needs? *Infancy, 14*(1), 117-130. doi:10.1080/15250000802569868

Bruner, J. P. (n.d.). *Evolution, emotion and negative reciprocity*. University of California, Irvine, USA. Unpublished manuscript.

Brunet, M., Guy, F., Pilbeam, D., Mackaye, H. T., Likius, A., Ahounta, D., . . . Zollikofer, C. (2002). A new hominid from the Upper Miocene of Chad, Central Africa. *Nature, 418*(6894), 145-151. doi:10.1038/nature00879

Bullinger, A. F., Melis, A. P., & Tomasello, M. (2011). Chimpanzees, *Pan troglodytes*, prefer individual over collaborative strategies towards goals. *Animal Behaviour, 82*(5), 1135-1141. doi:10.1016/j.anbehav.2011.08.008

Bunn, H. T. (2007). Meat made us human. In P. S. Ungar (Ed.), *Evolution of the human diet: the known, the unknown, and the unknowable* (pp. 191-211). Oxford: Oxford University Press.

Bunn, H. T., & Blumenschine, R. J. (1987). On "Theoretical Framework and Tests" of Early Hominid Meat and Marrow Acquisition: A Reply to Shipman. *American Anthropologist, 89*(2), 444-448. doi:10.1525/aa.1987.89.2.02a00170

Bunn, H. T., Harris, J. W. K., Isaac, G., Kaufulu, Z., Kroll, E., Schick, K., . . . Behrensmeyer, A. K. (1980). Fxjj50: an Early Pleistocene Site in Northern Kenya. *World Archaeology, 12*(2), 109-136.

Bunn, H. T., & Pickering, T. R. (2010). Bovid mortality profiles in paleoecological context falsify hypotheses of endurance running-hunting and passive scavenging by early Pleistocene hominins. *Quaternary Research, 74*(3), 395-404. doi:10.1016/j.yqres.2010.07.012

Burkart, J. M., Allon, O., Amici, F., Fichtel, C., Finkenwirth, C., Heschl, A., . . . van Schaik, C. P. (2014). The evolutionary origin of human hyper-cooperation. *Nature Communications, 5*. doi:10.1038/ncomms5747

Burkart, J. M., Hrdy, S. B., & van Schaik, C. P. (2009). Cooperative Breeding and Human Cognitive Evolution. *Evolutionary Anthropology, 18*(5), 175-186. doi:10.1002/Evan.20222

Burkart, J. M., & van Schaik, C. P. (2010). Cognitive consequences of cooperative breeding in primates? *Animal Cognition, 13*(1), 1-19. doi:10.1007/s10071-009-0263-7

Bybee, J. L. (1985). *Morphology: A study on the relation between meaning and form*. Amsterdam: John Benjamins.

Bybee, J. L. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press.

Calcott, B. (2009). Lineage Explanations: Explaining How Biological Mechanisms Change. *The British Journal for the Philosophy of Science, 60*(1), 51-78. doi:10.1093/bjps/axn047

Call, J. (2009). Contrasting the Social Cognition of Humans and Nonhuman Apes: The Shared Intentionality Hypothesis. *Topics in Cognitive Science, 1*(2), 368-379.

Call, J. (2010). Do apes know that they could be wrong? *Animal Cognition, 13*(5), 689-700. doi:10.1007/s10071-010-0317-x

Call, J., & Carpenter, M. (2001). Do apes and children know what they have seen? *Animal Cognition, 3*(4), 207-220. doi:10.1007/s100710100078

Call, J., & Carpenter, M. (2002). Three Sources of Information in Social Learning. In K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in Animals and Artifacts* (pp. 211-228). Cambridge, MA: MIT Press.

Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences, 12*(5), 187-192. doi:10.1016/j.tics.2008.02.010

Callaghan, T., Moll, H., Rakoczy, H., Warneken, F., Liszkowski, U., Behne, T., & Tomasello, M. (2011). Early social cognition in three cultural contexts. *Monogr Soc Res Child Dev, 76*(2), vii-viii, 1-142. doi:10.1111/j.1540-5834.2011.00603.x

Camp, E., & Shupe, E. (2017). Instrumental reasoning in nonhuman animals. In K. Andrews & J. Beck (Eds.), *The Routledge handbook of philosophy of animal minds* (1 edition . ed., pp. 100-108). London: Routledge.

Cariani, F. (2016). Deontic Modals and Probabilities: One Theory to Rule Them All? In N. Charlow & M. Chrisman (Eds.), *Deontic Modality* (pp. 11-46). Oxford: Oxford University Press.

Carpenter, M., & Call, J. (2013). How joint is the joint attention of apes and human infants? In J. Metcalfe & H. S. Terrace (Eds.), *Agency and joint attention* (pp. 49-61). Oxford: Oxford University Press.

Carpenter, M., Tomasello, M., & Striano, T. (2005). Role Reversal Imitation and Language in Typically Developing Infants and Children With Autism. *Infancy, 8*(3), 253-278. doi:10.1207/s15327078in0803_4

Carter, D. B., & Patterson, C. J. (1982). Sex roles as social conventions: The development of children’s conceptions of sex-role stereotypes. *Developmental Psychology, 18*(6), 812-824. doi:10.1037/0012-1649.18.6.812

Casler, K., Terziyan, T., & Greene, K. (2009). Toddlers view artifact function normatively. *Cognitive Development, 24*(3), 240-247. doi:10.1016/j.cogdev.2009.03.005

Chang, L. J., Smith, A., Dufwenberg, M., & Sanfey, A. G. (2011). Triangulating the Neural, Psychological, and Economic Bases of Guilt Aversion. *Neuron, 70*(3), 560-572. doi:10.1016/j.neuron.2011.02.056

Charlow, N. (2016). Decision Theory: Yes! Truth Conditions: No! *Deontic Modality* (pp. 47-81). Oxford: Oxford University Press.

Cheney, D. L., & Seyfarth, R. M. (1990). *How monkeys see the world: inside the mind of another species*. Chicago: University of Chicago Press.

Chevalier-Skolnikoff, S. (1990). Tool Use by Wild Cebus Monkeys at Santa-Rosa-National-Park, Costa-Rica. *Primates, 31*(3), 375-383. doi:10.1007/Bf02381108

Chiang, M. (1967). Use of Tools by Wild Macaque Monkeys in Singapore. *Nature, 214*(5094), 1258-1259. doi:10.1038/2141258a0

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Chomsky, N. (1968). *Language and mind*. New York, NY: Harcourt.

Chomsky, N. (1980). Rules and Representations. *Behavioral and Brain Sciences, 3*(1), 1-15.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, Holland/Cinnaminson, NJ: Foris Publications.

Chomsky, N. (1982). *Some concepts and consequences of the theory of government and binding*. Cambridge, MA: MIT Press.

Chomsky, N. (1986). *Knowledge of language: its nature, origin, and use*. New York, NY: Praeger.

Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale & S. J. Keyser (Eds.), *The View From Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*: MIT Press.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: The MIT Press.

Chomsky, N. (2001). Beyond Explanatory Adequacy. *MIT Occasional Papers in Linguistics, 20*(1), 1-28.

Ciochon, R. L. (1983). Hominoid cladistics and the ancestry of modern apes and humans. In R. L. Ciochon & R. S. Corruccini (Eds.), *New interpretations of ape and human ancestry* (pp. 783-837). New York, NY: Plenum Press.

Clark, C. (1977). A preliminary report on weaning among chimpanzees of the Gombe National Park, Tanzania. In S. Chevalier-Skolnikoff & F. Poirier (Eds.), *Primate Bio-Social Development* (pp. 235-260). New York, NY: Garland Press.

Clark, G., & Henneberg, M. (2015). The life history of Ardipithecus ramidus: a heterochronic model of sexual and social maturation. *Anthropological Review, 78*(2), 109-132. doi:10.1515/anre-2015-0009

Clemens, J. D., Sack, D. A., Harris, J. R., Khan, M. R., Chakraborty, J., Chowdhury, S., . . . Holmgren, J. (1990). Breast-Feeding and the Risk of Severe Cholera in Rural Bangladeshi Children. *American Journal of Epidemiology, 131*(3), 400-411.

Clutton-Brock, T. H., Harvey, P. H., & Rudder, B. (1977). Sexual Dimorphism, Socionomic Sex-Ratio and Body-Weight in Primates. *Nature, 269*(5631), 797-800. doi:10.1038/269797a0

Colby, A., & Kohlberg, L. (1987). *The measurement of moral judgment* (Vol. 1). Cambridge: Cambridge University Press.

Conway, C. M., & Christiansen, M. H. (2001). Sequential learning in non-human primates. *Trends in Cognitive Sciences, 5*(12), 539-546. doi:10.1016/S1364-6613(00)01800-3

Corkum, V., & Moore, C. (1995). Development of joint visual attention in infants. In C. Moore & P. J. Dunham (Eds.), *Joint attention: its origins and role in development* (pp. 61-83). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Couchman, J. J., Coutinho, M. V., Beran, M. J., & Smith, J. D. (2010). Beyond stimulus cues and reinforcement signals: a new approach to animal metacognition. *Journal of Comparative Psychology, 124*(4), 356-368. doi:10.1037/a0020129

Cowie, F. (1999). *What's Within? Nativism Reconsidered*. Oxford: Oxford University Press.

Crofoot, M. C., & Wrangham, R. W. (2010). Intergroup Aggression in Primates and Humans: The Case for a Unified Theory. In P. M. Kappeler & J. B. Silk (Eds.), *Mind the gap: tracing the origins of human universals* (pp. 171-195). Berlin: Springer.

Croft, W. (2000). *Explaining language change: an evolutionary approach*. New York, NY: Longman.

Crompton, R. H., Vereecke, E. E., & Thorpe, S. K. S. (2008). Locomotion and posture from the common hominoid ancestor to fully modern hominins, with special reference to the last common panin/hominin ancestor. *Journal of Anatomy, 212*(4), 501-543. doi:10.1111/j.1469-7580.2008.00870.x

Crook, J. H. (1972). Sexual Selection, Dimorphism, and Social Organization in the Primates. In B. G. Campbell (Ed.), *Sexual Selection and the Descent of Man* (pp. 231-281). Chicago, IL: Aldine.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*(4), 148-153. doi:10.1016/j.tics.2009.01.005

Cushman, F., & Young, L. (2011). Patterns of Moral Judgment Derive From Nonmoral Psychological Representations. *Cognitive Science, 35*(6), 1052-1075. doi:10.1111/j.1551-6709.2010.01167.x

Dąbrowska, E. (n.d.). *What exactly is Universal Grammar, and has anyone seen it?* Department of Humanities. Northumbria University. Unpublished manuscript. Retrieved from https://www.northumbria.ac.uk/static/5007/sasspdf/whatexactlyisUG.pdf

Damasio, A. R. (1994). *Descartes' error: emotion, reason and the human brain*. London: Picador.

Damasio, A. R. (1999). *The feeling of what happens: body and emotion in the making of consciousness* (1st ed.). New York, NY: Harcourt Brace.

Dart, R. (1953). The Predatory Transition from Ape to Man. *International Anthropological and Linguistic Review, 1*, 201-217.

Darwall, S. L. (2006). *The second-person standpoint: morality, respect, and accountability*. Cambridge, MA: Harvard University Press.

Davidson, R. J., Putnam, K. M., & Larson, C. L. (2000). Dysfunction in the neural circuitry of emotion regulation: A possible prelude to violence. *Science, 289*(5479), 591-594. doi:10.1126/science.289.5479.591

Davila-Ross, M., Allcock, B., Thomas, C., & Bard, K. A. (2011). Aping expressions? Chimpanzees produce distinct laugh types when responding to laughter of others. *Emotion, 11*(5), 1013-1020. doi:10.1037/a0022594

de Hooge, I. E., Zeelenberg, M., & Breugelmans, S. M. (2007). Moral sentiments and cooperation: Differential influences of shame and guilt. *Cognition and Emotion, 21*(5), 1025-1042. doi:10.1080/02699930600980874

de Villiers, P. A., & de Villiers, J. G. (2012). Deception dissociates from false belief reasoning in deaf children: implications for the implicit versus explicit theory of mind distinction. *British Journal of Developmental Psychology, 30*(1), 188-209. doi:10.1111/j.2044-835X.2011.02072.x

de Waal, F. B. M. (1989). Food sharing and reciprocal obligations among chimpanzees. *J. Hum. Evol., 18*, 433.

de Waal, F. B. M. (1995). Bonobo sex and society: the behavior of a close relative challenges assumptions about male supremacy in human evolution. *Scientific American, 272*(3), 82-88.

de Waal, F. B. M. (2001). Apes from Venus: Bonobos and Human Social Evolution. In F. B. M. d. Waal (Ed.), *Tree of origin: what primate behavior can tell us about human social evolution* (pp. 39-68). Cambridge, MA: Harvard University Press.

de Waal, F. B. M., & Lanting, F. (1997). *Bonobo: the forgotten ape*. Berkeley, CA: University of California Press.

Dennett, D. C. (1987). *The intentional stance*. Cambridge, MA: MIT Press.

DeVore, I., & Washburn, S. L. (1963). Baboon ecology and human evolution. In F. C. Howell & F. o. Bourlière (Eds.), *African ecology and human evolution* (pp. 335-367). Chicago: Aldine.

Dindo, M., Thierry, B., & Whiten, A. (2008). Social diffusion of novel foraging methods in brown capuchin monkeys. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1631), 187-193. doi:10.1098/rspb.2007.1318

Dubreuil, B. (2010a). *Human evolution and the origins of hierarchies: the state of nature*. Cambridge: Cambridge University Press.

Dubreuil, B. (2010b). Paleolithic public goods games: why human culture and cooperation did not evolve in one step. *Biology & Philosophy, 25*(1), 53-73. doi:10.1007/s10539-009-9177-7

Duda, P., & Zrzavý, J. (2013). Evolution of life history and behavior in Hominidae: Towards phylogenetic reconstruction of the chimpanzee-human last common ancestor. *Journal of Human Evolution, 65*(4), 424-446. doi:10.1016/j.jhevol.2013.07.009

Duguid, S., Wyman, E., Bullinger, A. F., Herfurth-Majstorovic, K., & Tomasello, M. (2014). Coordination strategies of chimpanzees and human children in a Stag Hunt game. *Proceedings of the Royal Society of London, Series B, 281*(1796). doi:10.1098/rspb.2014.1973

Dunham, Y., Baron, A. S., & Banaji, M. R. (2008). The development of implicit intergroup cognition. *Trends in Cognitive Sciences, 12*(7), 248-253. doi:10.1016/j.tics.2008.04.006

Dupoux, E., & Jacob, P. (2007). Universal moral grammar: a critical appraisal. *Trends in Cognitive Sciences, 11*(9), 373-378. doi:10.1016/j.tics.2007.07.001

Durrleman, S., Pennec, X., Trouve, A., Ayache, N., & Braga, J. (2012). Comparison of the endocranial ontogenies between chimpanzees and bonobos via temporal regression and spatiotemporal registration. *Journal of Human Evolution, 62*(1), 74-88. doi:10.1016/j.jhevol.2011.10.004

Dwyer, S. (1999). Moral competence. In K. Murasugi & R. Stainton (Eds.), *Philosophy and linguistics* (pp. 169-190). Boulder: Westview Press.

Dwyer, S. (2007). How good is the linguistic analogy? In P. Carruthers, S. Laurence, & S. P. Stich (Eds.), *The Innate Mind: Culture and Cognition* (Vol. 2, pp. 237-255). Oxford: Oxford University Press.

Dwyer, S., & Hauser, M. D. (2008). Dupoux and Jacob's moral instincts: throwing out the baby, the bathwater and the bathtub. *Trends in Cognitive Sciences, 12*(1), 1-2. doi:10.1016/j.tics.2007.10.006

Dwyer, S., Huebner, B., & Hauser, M. D. (2010). The Linguistic Analogy: Motivations, Results, and Speculations. *Topics in Cognitive Science, 2*(3), 486-510. doi:10.1111/j.1756-8765.2009.01064.x

Eisenberg, N., & Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychol. Bull., 101*, 91.

Eisenberg, T., Garvey, S. P., & Wells, M. T. But was he sorry? The role of remorse in capital sentencing. *Cornell Law Review, 83*(6), 1599.

Ekman, P. (1999). Basic Emotions. In T. Dalgleish & M. J. Power (Eds.), *Handbook of cognition and emotion* (pp. 45-60). New York, NY: Wiley.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *J Pers Soc Psychol, 17*(2), 124-129.

Ellis, B. D. (2001). *Scientific essentialism*. Cambridge: Cambridge University Press.

Ellis, B. D. (2002). *The philosophy of nature: a guide to the new essentialism*. Montreal/Ithaca N.Y.: McGill-Queen's University Press.

Engelmann, J. M., Herrmann, E., & Tomasello, M. (2012). Five-Year Olds, but Not Chimpanzees, Attempt to Manage Their Reputations. *PLoS ONE, 7*(10). doi:10.1371/journal.pone.0048433

Evans, J. S. B. T., & Frankish, K. (2009). *In two minds: dual processes and beyond*. Oxford: Oxford University Press.

Fairbairn, D. J. (1997). Allometry for sexual size dimorphism: Pattern and process in the coevolution of body size in males and females. *Annual Review of Ecology and Systematics, 28*, 659-687. doi:10.1146/annurev.ecolsys.28.1.659

Feldman, F. (1986). *Doing the best we can: an essay in informal deontic logic*. Dordrecht: D. Reidel Pub. Co.

Fernandes, M. E. B. (1991). Tool Use and Predation of Oysters (Crassostrea-Rhizophorae) by the Tufted Capuchin, Cebus-Apella-Apella, in Brackish Water Mangrove Swamp. *Primates, 32*(4), 529-531. doi:10.1007/Bf02381944

Fernandez-Duque, E., Valeggia, C. R., & Mendoza, S. P. (2009). The Biology of Paternal Care in Human and Nonhuman Primates. *Annual Review of Anthropology, 38,* 115-130. doi:10.1146/annurev-anthro-091908-164334

Fessler, D. M. T., Barrett, H. C., Kanovsky, M., Stich, S., Holbrook, C., Henrich, J. P., . . . Laurence, S. (2015). Moral parochialism and contextual contingency across seven societies. *Proceedings of the Royal Society B: Biological Sciences, 282*(1813), 20150907. doi:10.1098/rspb.2015.0907

Fischbacher, U., & Utikal, V. (2013). On the acceptance of apologies. *Games and Economic Behavior, 82*, 592-608. doi:10.1016/j.geb.2013.09.003

Flanagan, O. J. (1991). *Varieties of moral personality: ethics and psychological realism*. Cambridge, MA: Harvard University Press.

Flannery, K. V., & Marcus, J. (2012). *The creation of inequality: how our prehistoric ancestors set the stage for monarchy, slavery, and empire*. Cambridge: Harvard University Press.

Fletcher, G. E., Warneken, F., & Tomasello, M. (2012). Differences in cognitive processes underlying the collaborative activities of children and chimpanzees. *Cognitive Development, 27*(2), 136-153. doi:10.1016/j.cogdev.2012.02.003

Fodor, J. A. (1983). *The modularity of mind: an essay on faculty psychology*. Cambridge, MA.: MIT Press.

Fodor, J. D. (2001). Setting syntactic parameters. In M. R. Baltin & C. Collins (Eds.), *The handbook of contemporary syntactic theory* (pp. 730-767). Malden, Mass.: Blackwell Publishers.

Fodor, J. D., & Sakas, W. G. (2004). Evaluating models of parameter setting. In A. Brugos, L. Micciulla, & C. E. Smith (Eds.), *Proceedings of the 28th Annual Boston University Conference on Language Development* (Vol. 1, pp. 1-27).

Foot, P. (1972). Morality as a System of Hypothetical Imperatives. *Philosophical Review, 81*(3), 305-316. doi:10.2307/2184328

Fox, M., Sear, R., Beise, J., Ragsdale, G., Voland, E., & Knapp, L. A. (2009). Grandma plays favourites: X-chromosome relatedness and sex-specific childhood mortality. *Proceedings of the Royal Society B: Biological Sciences, 277*(1681), 567-573. doi:10.1098/rspb.2009.1660

Frankish, K. (2010). Dual-Process and Dual-System Theories of Reasoning. *Philosophy Compass, 5*(10), 914-926. doi:10.1111/j.1747-9991.2010.00330.x

Frankish, K., & Evans, J. S. B. T. (2009). The duality of mind: An historical perspective. In J. S. B. T. Evans & K. Frankish (Eds.), *In two minds: dual processes and beyond* (pp. 1-29). Oxford: Oxford University Press.

Fraser, B. J. (2010). *Sexual Selection and the Evolution of Morality*. (PhD Thesis), Australian National University, Canberra, ACT.

Fraser, B. J. (2012). The nature of moral judgements and the extent of the moral domain. *Philosophical Explorations, 15*(1), 1-16. doi:10.1080/13869795.2012.647356

Frith, C. D., & Amodio, D. M. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience, 7*(4), 268-277. doi:10.1038/nrn1884

Fujita, K. (2009). Metamemory in tufted capuchin monkeys (*Cebus apella*). *Animal Cognition, 12*(4), 575-585. doi:10.1007/s10071-009-0217-0

Galdikas, B. M., & Wood, J. W. (1990). Birth spacing patterns in humans and apes. *American Journal of Physical Anthropology, 83*(2), 185-191. doi:10.1002/ajpa.1330830207

Galef, B. G. (1996). Food selection: Problems in understanding how we choose foods to eat. *Neuroscience and Biobehavioral Reviews, 20*(1), 67-73. doi:10.1016/0149-7634(95)00041-C

Galef, B. G. (2009). Culture in animals? In K. N. Laland & B. G. Galef (Eds.), *The question of animal culture* (pp. 222-246). Cambridge, MA: Harvard University Press.

Galef, B. G., & Wigmore, S. W. (1983). Transfer of Information Concerning Distant Foods: a Laboratory Investigation of the 'Information-Center' Hypothesis. *Animal Behaviour, 31*(Aug), 748-758. doi:10.1016/S0003-3472(83)80232-2

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of "theory of mind". *Trends in Cognitive Sciences, 7*(2), 77-83.

Gamble, C., Gowlett, J., & Dunbar, R. I. (2014). *Thinking big: how the evolution of social life shaped the human mind*. London: Thames & Hudson.

Garcia, J., Hankins, W. G., & Rusiniak, K. W. (1974). Behavioral regulation of the milieu interne in man and rat. *Science, 185*(4154), 824-831.

Garner, R. (2007). Abolishing morality. *Ethical Theory and Moral Practice, 10*(5), 499-513.

Geary, D. C., Bailey, D., & Oxford, J. (2011). Reflections on the human family. In C. Salmon & T. K. Shackelford (Eds.), *The Oxford handbook of evolutionary family psychology* (pp. 365-385). Oxford: Oxford University Press.

Geary, D. C., & Flinn, M. V. (2001). Evolution of Human Parental Behavior and the Human Family. *Parenting: Science and Practice, 1*(1-2), 5-61. doi:10.1207/S15327922par011&2_2

Geertz, C. (1975). Thick Description: Toward an Interpretative Theory of Culture *The interpretation of cultures: selected essays* (pp. 3-30). New York, NY: Basic Books.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52*(1), 45-56. doi:10.1037/0003-066x.52.1.45

Geraci, A., & Surian, L. (2011). The developmental roots of fairness: infants' reactions to equal and unequal distributions of resources. *Developmental Science, 14*(5), 1012-1020. doi:10.1111/j.1467-7687.2011.01048.x

Gibbard, A. (1990). *Wise choices, apt feelings: a theory of normative judgment*. Cambridge, MA: Harvard University Press.

Gilbert, M. (1989). *On social facts*. London: Routledge.

Gilby, I. C. (2006). Meat sharing among the Gombe chimpanzees: harassment and reciprocal exchange. *Animal Behaviour, 71*, 953-963. doi:10.1016/j.anbehav.2005.09.009

Givón, T. (1979). *On understanding grammar*. New York, NY: Academic Press.

Göckeritz, S., Schmidt, M. F. H., & Tomasello, M. (2014). Young children's creation and transmission of social norms. *Cognitive Development, 30*(0), 81-95. doi:10.1016/j.cogdev.2014.01.003

Godfrey-Smith, P. (2014). *Philosophy of biology*. Princeton: Princeton University Press.

Gold, G. J., & Weiner, B. (2000). Remorse, confession, group identity, and expectancies about repeating a transgression. *Basic and Applied Social Psychology, 22*(4), 291-300. doi:10.1207/15324830051035992

Goldberg, A. E. (2006). *Constructions at work: the nature of generalization in language*. Oxford: Oxford University Press.

Goldberg, T. L., & Wrangham, R. W. (1997). Genetic correlates of social behaviour in wild chimpanzees: evidence from mitochondrial DNA. *Animal Behaviour, 54*, 559-570. doi:10.1006/anbe.1996.0450

Goldman, A. I. (1993). Ethics and Cognitive Science. *Ethics, 103*(2), 337-360. doi:10.2307/2381527

Gómez, J. C., Sarriá, E., & Tamarit, J. (1993). The comparative study of early communication and theories of mind: Ontogeny, phylogeny, and pathology. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding Other Minds: Perspectives from Autism* (pp. 397-426). Oxford: Oxford University Press.

Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*(2), 109-135. doi:10.1016/S0010-0277(99)00003-7

Gonzalez-Cabrera, I. (forthcoming). On social tolerance and the evolution of human normative guidance. *British Journal for the Philosophy of Science*.

Goodall, J. (1964). Tool-Using and Aimed Throwing in a Community of Free-Living Chimpanzees. *Nature, 201*(4926), 1264-1266.

Goodall, J. (1986). *The chimpanzees of Gombe: patterns of behavior*. Cambridge, MA: Belknap Press of Harvard University Press.

Gowlett, J. A. (2016). The discovery of fire by humans: a long and convoluted process. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*(1696). doi:10.1098/rstb.2015.0164

Gräfenhain, M., Behne, T., Carpenter, M., & Tomasello, M. (2009). Young children's understanding of joint commitments. *Developmental Psychology, 45*(5), 1430-1443. doi:10.1037/a0016122

Gräfenhain, M., Carpenter, M., & Tomasello, M. (2013). Three-year-olds' understanding of the consequences of joint commitments. *PLoS ONE*, *8*(9), e73039.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral Foundations Theory: The Pragmatic Validity of Moral Pluralism. *Advances in Experimental Social Psychology, 47*, 55-130. doi:10.1016/B978-0-12-407236-7.00002-4

Gredeback, G., & Melinder, A. (2010). Infants' understanding of everyday social interactions: a dual process account. *Cognition*, *114*(2), 197-206. doi:10.1016/j.cognition.2009.09.004

Greenberg, J. R., Hamann, K., Warneken, F., & Tomasello, M. (2010). Chimpanzee helping in collaborative and noncollaborative contexts. *Animal Behaviour, 80*(5), 873-880. doi:10.1016/j.anbehav.2010.08.008

Greene, J. D. (2013). *Moral tribes: emotion, reason, and the gap between us and them*. London: The Penguin Press.

Greenfield, L. O. (1992). Origin of the human canine: A new solution to an old enigma. *American Journal of Physical Anthropology, 35*(S15), 153-185. doi:10.1002/ajpa.1330350607

Grehan, J. R., & Schwartz, J. H. (2009). Evolution of the second orangutan: phylogeny and biogeography of hominid origins. *Journal of Biogeography, 36*(10), 1823-1844. doi:10.1111/j.1365-2699.2009.02141.x

Griffiths, P. E., & Machery, E. (2008). Innateness, Canalization, and 'Biologicizing the Mind'. *Philosophical Psychology, 21*(3), 397-414. doi:10.1080/09515080802201146

Gross, J. J. (1998a). Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology, 74*(1), 224-237.

Gross, J. J. (1998b). The emerging field of emotion regulation: An integrative review. *Review of General Psychology, 2*(3), 271-299.

Gross, J. J. (1999). Emotion regulation: Past, present, and future. *Cognition & Emotion, 13*(5), 551-573.

Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology, 39*(3), 281-291.

Gruber, T., Clay, Z., & Zuberbühler, K. (2010). A comparison of bonobo and chimpanzee tool use: evidence for a female bias in the *Pan* lineage. *Animal Behaviour, 80*(6), 1023-1033. doi:10.1016/j.anbehav.2010.09.005

Gurven, M. (2004). To give and to give not: The behavioral ecology of human food transfers. *Behavioral and Brain Sciences, 27*(4), 543-560.

Haidt, J. (2003). The moral emotions. In R. J. Davidson, H. H. Goldsmith, & K. R. Scherer (Eds.), *Handbook of affective sciences* (pp. 852-870). Oxford: Oxford University Press.

Haidt, J. (2007). The new synthesis in moral psychology. *Science, 316*(5827), 998-1002. doi:10.1126/science.1137651

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. doi:10.1007/s11211-007-0034-z

Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, Culture, and Morality, or Is It Wrong to Eat Your Dog? *Journal of Personality and Social Psychology, 65*(4), 613-628. doi:10.1037//0022-3514.65.4.613

Haile-Selassie, Y. (2001). Late Miocene hominids from the Middle Awash, Ethiopia. *Nature, 412*(6843), 178-181. doi:10.1038/35084063

Haile-Selassie, Y., Suwa, G., & White, T. D. (2004). Late Miocene Teeth from Middle Awash, Ethiopia, and Early Hominid Dental Evolution. *Science, 303*(5663), 1503-1505. doi:10.1126/science.1092978

Haile-Selassie, Y., Suwa, G., & White, T. D. (2009). Hominidae. In Y. Haile-Selassie & G. WoldeGabriel (Eds.), *Ardipithecus kadabba: Late Miocene evidence from the Middle Awash, Ethiopia* (pp. 159-236). Berkeley, CA: University of California Press.

Haile-Selassie, Y., & WoldeGabriel, G. (2009). *Ardipithecus kadabba: late miocene evidence from the Middle Awash, Ethiopia*. Berkeley, CA: University of California Press.

Hall, K. R. L. (1962). Numerical data, maintenance activities and locomotion of the wild chacma baboon, *Papio ursinus*. *Proceedings of the Zoological Society of London, 139*(2), 181-220. doi:10.1111/j.1469-7998.1962.tb01827.x

Hamann, K., Warneken, F., Greenberg, J. R., & Tomasello, M. (2011). Collaboration encourages equal sharing in children but not in chimpanzees. *Nature, 476*(7360), 328-331. doi:10.1038/nature10278

Hamann, K., Warneken, F., & Tomasello, M. (2012). Children’s Developing Commitments to Joint Goals. *Child Development, 83*(1), 137-145. doi:10.1111/j.1467-8624.2011.01695.x

Hamann, S., Herman, R. A., Nolan, C. L., & Wallen, K. (2004). Men and women differ in amygdala response to visual sexual stimuli. *Nat Neurosci, 7*(4), 411-416. doi:10.1038/nn1208

Hamilton, W. J., Buskirk, R. E., & Buskirk, W. H. (1978). Omnivory and Utilization of Food Resources by Chacma Baboons, *Papio ursinus*. *American Naturalist, 112*(987), 911-924.

Hampton, R. R. (2001). Rhesus monkeys know when they remember. *Proceedings of the National Academy of Sciences, 98*(9), 5359-5362. doi:10.1073/pnas.071600998

Hare, B., & Kwetuenda, S. (2010). Bonobos voluntarily share their own food with others. *Current Biology, 20*(5), R230-R231.

Hare, B., Melis, A. P., Woods, V., Hastings, S., & Wrangham, R. W. (2007). Tolerance Allows Bonobos to Outperform Chimpanzees on a Cooperative Task. *Current Biology, 17*(7), 619-623. doi:10.1016/j.cub.2007.02.040

Hare, B., Plyusnina, I., Ignacio, N., Schepina, O., Stepika, A., Wrangham, R. W., & Trut, L. (2005). Social cognitive evolution in captive foxes is a correlated by-product of experimental domestication. *Current Biology, 15*(3), 226-230. doi:10.1016/j.cub.2005.01.040

Hare, B., & Tomasello, M. (2005). Human-like social skills in dogs? *Trends in Cognitive Sciences, 9*(9), 439-444. doi:10.1016/j.tics.2005.07.003

Hare, B., Wobber, V., & Wrangham, R. W. (2012). The self-domestication hypothesis: evolution of bonobo psychology is due to selection against aggression. *Animal Behaviour, 83*(3), 573-585. doi:10.1016/j.anbehav.2011.12.007

Hart, D., & Sussman, R. W. (2005). *Man the hunted: primates, predators, and human evolution*. New York, NY: Westview Press.

Haspelmath, M. (2007). Pre-established categories don't exist: Consequences for language description and typology *Linguistic Typology* (Vol. 11, pp. 119-132).

Hauser, M. D. (2006a). The liver and the moral organ. *Social Cognitive and Affective Neuroscience, 1*(3), 214-220. doi:10.1093/Scan/Nsl026

Hauser, M. D. (2006b). *Moral minds: how nature designed our universal sense of right and wrong*. New York, NY: Ecco Press/HarperCollins.

Hauser, M. D., Young, L., & Cushman, F. (2008a). On misreading the linguistic analogy: response to Jesse Prinz and Ron Mallon. In W. Sinnott-Armstrong (Ed.), *Moral Psychology: The Cognitive Science of Morality: Intuition and Diversity* (Vol. 2, pp. 171-180). Cambridge, MA: MIT Press.

Hauser, M. D., Young, L., & Cushman, F. (2008b). Reviving Rawl's Linguistic Analogy. In W. Sinnott-Armstrong (Ed.), *Moral Psychology: The Cognitive Science of Morality: Intuition and Diversity* (Vol. 2, pp. 107-144). Cambridge, MA: MIT Press.

Hawkes, K. (2003). Grandmothers and the evolution of human longevity. *American Journal of Human Biology, 15*(3), 380-400. doi:10.1002/Ajhb.10156

Hawkes, K. (2012). Stag Hunts or Rearing Environments? *Current Anthropology, 53*(6), 687-688.

Hawkes, K. (2014). Primate Sociality to Human Cooperation: Why Us and Not Them? *Human Nature, 25*(1), 28-48. doi:10.1007/s12110-013-9184-x

Hawkes, K., & Coxworth, J. E. (2013). Grandmothers and the Evolution of Human Longevity: A Review of Findings and Future Directions. *Evolutionary Anthropology, 22*(6), 294-302. doi:10.1002/Evan.21382

Hawkes, K., O’Connell, J. F., & Jones, N. G. B. (2001). Hunting and nuclear families: Some lessons from the Hadza about men’s work. *Current Anthropology, 42*(5), 681-709. doi:10.1086/322559

Hawkes, K., O’Connell, J. F., Jones, N. G. B., Alvarez, H., & Charnov, E. L. (1998). Grandmothering, menopause, and the evolution of human life histories. *Proceedings of the National Academy of Sciences of the United States of America, 95*(3), 1336-1339.

Helwig, C. C., Tisak, M. S., & Turiel, E. (1990). Children’s Social Reasoning in Context: Reply to Gabennesch. *Child Development, 61*(6), 2068-2078. doi:10.1111/j.1467-8624.1990.tb03588.x

Henrich, J. P. (2016). *The secret of our success: how culture is driving human evolution, domesticating our species, and making us smarter*. Princeton: Princeton University Press.

Henrich, J. P., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behav Brain Sci, 33*(2-3), 61-83; discussion 83-135. doi:10.1017/S0140525X0999152X

Henrich, J. P., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., . . . Ziker, J. (2006). Costly punishment across human societies. *Science, 312*(5781), 1767-1770. doi:10.1126/science.1127333

Henshilwood, C., & Dubreuil, B. (2011). The Still Bay and Howiesons Poort, 77-59 ka: Symbolic Material Culture and the Evolution of the Mind during the African Middle Stone Age. *Current Anthropology, 52*(3), 361-400. doi:10.1086/660022

Henshilwood, C. S., & d’Errico, F. (2011). *Homo symbolicus: the dawn of language, imagination and spirituality*. Amsterdam: John Benjamins.

Henshilwood, C. S., d’Errico, F., & Watts, I. (2009). Engraved ochres from the Middle Stone Age levels at Blombos Cave, South Africa. *Journal of Human Evolution, 57*(1), 27-47. doi:10.1016/j.jhevol.2009.01.005

Hepach, R., Vaish, A., & Tomasello, M. (2012). Young Children Are Intrinsically Motivated to See Others Helped. *Psychological Science, 23*(9), 967-972. doi:10.1177/0956797612440571

Hepach, R., Vaish, A., & Tomasello, M. (2013). Young Children Sympathize Less in Response to Unjustified Emotional Distress. *Developmental Psychology, 49*(6), 1132-1138. doi:10.1037/A0029501

Hernandez-Aguilar, R. A., Moore, J., & Travis, R. P. (2007). Savanna chimpanzees use tools to harvest the underground storage organs of plants. *Proceedings of the*

National Academy of Sciences of the United States of America, 104(49), 19210-19213. doi:10.1073/pnas.0707929104

Herrmann, E., Hare, B., Call, J., & Tomasello, M. (2010). Differences in the Cognitive Skills of Bonobos and Chimpanzees. *PLoS ONE, 5*(8), e12438. doi:10.1371/journal.pone.0012438

Herrmann, E., Hare, B., Cissewski, J., & Tomasello, M. (2011). A comparison of temperament in nonhuman apes and human infants. *Dev Sci, 14*(6), 1393-1405. doi:10.1111/j.1467-7687.2011.01082.x

Heyes, C., & Frith, C. D. (2014). The cultural evolution of mind reading. *Science, 344*(6190). doi:10.1126/science.1243091

Hill, K. (2002). Altruistic cooperation during foraging by the Ache, and the evolved human predisposition to cooperate. *Human Nature, 13*(1), 105.

Hill, K. (2012). *Social structure and inter-group interactions in hunter-gatherers: how humans achieved cumulative culture*. Paper presented at the Cultural evolution: patterns of cultural change and diversification, Stockholm.

Hirata, S., & Fuwa, K. (2007). Chimpanzees (*Pan troglodytes*) learn to act with other individuals in a cooperative task. *Primates, 48*(1), 13-21. doi:10.1007/s10329-006-0022-1

Hobbes, T. (1668/1994). Leviathan: with selected variants from the Latin edition of 1668. In E. M. Curley (Ed.), (pp. lxxvii, 584 p.). Indianapolis: Hackett.

Hobson, J. A., & Hobson, R. P. (2007). Identification: The missing link between joint attention and imitation? *Development and Psychopathology*, *19*(2), 411-431. doi:10.1017/S0954579407070204

Hohmann, G., & Fruth, B. (2003). Intra- and inter-sexual aggression by bonobos in the context of mating. *Behaviour, 140*, 1389-1413. doi:10.1163/156853903771980648

Hohmann, G., Gerloff, U., Tautz, D., & Fruth, B. (1999). Social bonds and genetic ties: Kinship association and affiliation in a community of bonobos (*Pan paniscus*). *Behaviour, 136*, 1219-1235. doi:10.1163/156853999501739

Hohmann, G., Mundry, R., & Deschner, T. (2009). The Relationship Between Socio-Sexual Behavior and Salivary Cortisol in Bonobos: Tests of the Tension Regulation Hypothesis. *American Journal of Primatology, 71*(3), 223-232. doi:10.1002/Ajp.20640

Hollos, M., Leis, P. E., & Turiel, E. (1986). Social Reasoning in Ijo Children and Adolescents in Nigerian Communities. *Journal of Cross-Cultural Psychology, 17*(3), 352-374. doi:10.1177/0022002186017003007

Hopper, P. (1987). Emergent grammar. In J. Aske, N. Beery, L. Michaelis, & H. Filip (Eds.), *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society* (pp. 139-157). Berkeley, CA: Berkeley Linguistics Society.

Hrdy, S. B. (1999). *Mother nature: a history of mothers, infants, and natural selection* (1st ed.). New York, NY: Pantheon Books.

Hrdy, S. B. (2005). Comes the child before the man: how cooperative breeding and prolonged postweaning dependence shaped human potentials. In B. S. Hewlett & M. E. Lamb (Eds.), *Hunter-gatherer childhoods: evolutionary, developmental, & cultural perspectives* (pp. 65-91). New Brunswick, N.J.: Aldine Transaction.

Hrdy, S. B. (2006). Evolutionary context of human development: The cooperative breeding model. In C. S. Carter, L. Ahnert, Grossmann, K. E., S. B. Hrdy, M. E. Lamb, S. W. Porges, & N. Sachser (Eds.), *Attachment and bonding: a new synthesis* (pp. 9-32). Cambridge, MA: MIT Press.

Hrdy, S. B. (2009a). Meet the Alloparents: shared child care may be the secret of human evolutionary success. *Natural History, 118*(3), 24-19.

Hrdy, S. B. (2009b). *Mothers and others: the evolutionary origins of mutual understanding*. Cambridge, MA: Belknap Press of Harvard University Press.

Hrdy, S. B. (2014). Development + Social Selection in the Emergence of “Emotionally Modern” Humans. In J. Decety & Y. Christen (Eds.), *New Frontiers in Social Neuroscience* (Vol. 21, pp. 57-91): Springer International Publishing.

Hrdy, S. B. (2016). Development plus social selection in the emergence of “emotionally modern humans”. In L. Meehan & A. Crittenden (Eds.), *Childhood: Origins, evolution, and implications*. Albuquerque: University of New Mexico Press and School for Advanced Research Press.

Hughes, C., Jaffee, S. R., Happe, F., Taylor, A., Caspi, A., & Moffitt, T. E. (2005). Origins of individual differences in theory of mind: from nature to nurture? *Child Development, 76*(2), 356-370. doi:10.1111/j.1467-8624.2005.00850.x

Hume, D. (1740/1978). A treatise of human nature. In L. A. Selby-Bigge & P. H. Nidditch (Eds.), (2d ed., pp. xix, 743 p.). Oxford: Oxford University Press.

Inbar, Y., Pizarro, D. A., Knobe, J., & Bloom, P. (2009a). Disgust Sensitivity Predicts Intuitive Disapproval of Gays. *Emotion, 9*(3), 435-443.

Inbar, Y., Pizarro, D. A., Knobe, J., & Bloom, P. (2009b). Disgust sensitivity predicts intuitive disapproval of gays. *Emotion, 9*(3), 435-439. doi:10.1037/a0015960

Ingmanson, E. J. (1996). Tool-using behavior in wild *Pan paniscus*: Social and ecological considerations. In A. E. Russon, K. A. Bard, & S. T. Parker (Eds.), *Reaching into thought: the minds of the great apes* (pp. 190-210). Cambridge: Cambridge University Press.

Jablonski, N. G., & Leakey, M. G. (Eds.). (2008). *Koobi Fora research project: The fossil monkeys* (Vol. 6). San Francisco, CA: Academy of Sciences.

Jackson, I., & Sirois, S. (2009). Infant cognition: going full factorial with pupil dilation. *Developmental Science, 12*(4), 670-679. doi:10.1111/j.1467-7687.2008.00805.x

Jamison, C. S., Cornell, L. L., Jamison, P. L., & Nakazato, H. (2002). Are all grandmothers equal? A review and a preliminary test of the “grandmother hypothesis” in Tokugawa Japan. *American Journal of Physical Anthropology, 119*(1), 67-76. doi:10.1002/ajpa.10070

Johnson, M. (1993). *Moral imagination: implications of cognitive science for ethics*. Chicago: University of Chicago Press.

Jolly, C. J. (1970). The Seed-Eaters: A New Model of Hominid Differentiation Based on a Baboon Analogy. *Man, 5*(1), 5-26. doi:10.2307/2798801

Jolly, C. J. (2001). A proper study for mankind: Analogies from the papionin monkeys and their implications for human evolution. *Yearbook of Physical Anthropology, 44*, 177-204. doi:10.1002/Ajpa.10021

Joyce, R. (2001). *The myth of morality*. Cambridge: Cambridge University Press.

Joyce, R. (2006). *The evolution of morality*. Cambridge, MA: MIT Press.

Joyce, R. (2014). The origins of moral judgment. *Behaviour, 151*(2-3), 261-278. doi:10.1163/1568539X-00003140

Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. *American Psychologist, 58*(9), 697.

Kanō, T. (1992). *The last ape: pygmy chimpanzee behavior and ecology*. Stanford: Stanford University Press.

Kant, I. (1785/1998). *Groundwork of the metaphysics of morals* (M. J. Gregor, Trans. M. J. Gregor Ed.). Cambridge: Cambridge University Press.

Kaplan, H., Hill, K., Lancaster, J. B., & Hurtado, A. M. (2000). A theory of human life history evolution: Diet, intelligence, and longevity. *Evolutionary Anthropology, 9*(4), 156-185.

Kappeler, P. M., & van Schaik, C. P. (2004). *Sexual selection in primates: new and comparative perspectives*. Cambridge: Cambridge University Press.

Kappelman, J. (1996). The evolution of body mass and relative brain size in fossil hominids. *Journal of Human Evolution, 30*(3), 243-276. doi:10.1006/jhev.1996.0021

Karama, S., Lecours, A. R., Leroux, J., Bourgouin, P., Beaudoin, G., Joubert, S., & Beauregard, M. (2002). Areas of brain activation in males and females. During viewing of erotic film excerpts. *Human Brain Mapping, 16*(1), 1-13. doi:10.1002/Hbm.10014

Karg, K., Schmelz, M., Call, J., & Tomasello, M. (2015). The goggles experiment: can chimpanzees use self-experience to infer what a competitor can see? *Animal Behaviour, 105*, 211-221. doi:10.1016/j.anbehav.2015.04.028

Katz, L. D. (2000). *Evolutionary origins of morality: cross-disciplinary perspectives*. Thorverton, Devon: Imprint Academic.

Kayne, R. S. (2005). Some notes on comparative syntax, with particular reference to English and French. In G. Cinque & R. S. Kayne (Eds.), *The Oxford handbook of comparative syntax* (pp. 3-69). Oxford: Oxford University Press.

Kelemen, D. (1999). The scope of teleological thinking in preschool children. *Cognition, 70*(3), 241-272.

Kelly, D. (2011). *Yuck!: the nature and moral significance of disgust*. Cambridge, MA: MIT Press.

Kelly, D., & Stich, S. (2008). Two Theories About the Cognitive Architecture Underlying Morality. In P. Carruthers, S. Laurence, & S. Stich (Eds.), The Innate Mind: Foundations and the Future (Vol. 3, pp. 348-366). Oxford: Oxford University Press.

Kelly, D., Stich, S., Haley, K. J., Eng, S. J., & Fessler, D. M. T. (2007). Harm, affect, and the moral/conventional distinction. *Mind & Language, 22*(2), 117-131.

Kelly, R. C. (2005). The evolution of lethal intergroup violence. *Proceedings of the National Academy of Sciences of the United States of America, 102*(43), 15294-15298. doi:10.1073/pnas.0505955102

Kennedy, G. E. (2005). From the ape’s dilemma to the weanling’s dilemma: early weaning and its evolutionary context. *J Hum Evol, 48*(2), 123-145. doi:10.1016/j.jhevol.2004.09.005

Ketelaar, T., & Au, W. T. (2003). The effects of feelings of guilt on the behavior of uncooperative individuals in repeated social bargaining games: An affect-as-information interpretation of the role of emotion in social interaction. *Cognition and Emotion, 17*, 429-453.

Killen, M. (1991). Social and moral development in early childhood. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Handbook of Moral Behavior and Development* (Vol. 2, pp. 115-138). Hillsdale, NJ: L. Erlbaum.

Killen, M., & Smetana, J. G. (2015). Origins and Development of Morality. In M. E. Lamb (Ed.), *Handbook of child psychology and developmental science* (7th ed., Vol. 3, pp. 701-749). New York, NY: Wiley-Blackwell.

Killen, M., & Turiel, E. (1998). Adolescents' and young adults' evaluations of helping and sacrificing for others. *Journal of Research on Adolescence, 8*(3), 355-375. doi:10.1207/s15327795jra0803_4

Kimbel, W. H., Suwa, G., Asfaw, B., Rak, Y., & White, T. D. (2014). *Ardipithecus ramidus* and the evolution of the human cranial base. *Proceedings of the National Academy of Sciences of the United States of America, 111*(3), 948-953. doi:10.1073/pnas.1322639111

King, G. E. (1976). Society and territory in human evolution. *Journal of Human Evolution, 5*(4), 323-331. doi:10.1016/0047-2484(76)90035-X

King, J. A., Blair, R. J. R., Mitchell, D. G., Dolan, R. J., & Burgess, N. (2006). Doing the right thing: a common neural circuit for appropriate violent or compassionate behavior. *NeuroImage, 30*(3), 1069-1076. doi:10.1016/j.neuroimage.2005.10.011

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences of the United States of America, 104*(30), 12577-12580. doi:10.1073/pnas.0705345104

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.

Kitcher, P. (2006). Biology and Ethics. In D. Copp (Ed.), *The Oxford handbook of ethical theory* (pp. 163-185). Oxford: Oxford University Press.

Kitcher, P. (2011). *The ethical project*. Cambridge, MA: Harvard University Press.

Knobe, J., Buckwalter, W., Nichols, S., Robbins, P., Sarkissian, H., & Sommers, T. (2012). Experimental Philosophy. *Annual Review of Psychology, 63*, 81-99. doi:10.1146/annurev-psych-120710-100350

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological Research on Joint Action: Theory and Data. In H. R. Brian (Ed.), *Psychology of Learning and Motivation* (Vol. Volume 54, pp. 59-101): Academic Press.

Knoblich, G., & Jordan, J. S. (2003). Action coordination in groups and individuals: Learning anticipatory control. *Journal of Experimental Psychology-Learning Memory and Cognition, 29*(5), 1006-1016. doi:10.1037/0278-7393.29.5.1006

Koban, L., Corradi-Dell'Acqua, C., & Vuilleumier, P. (2013). Integration of Error Agency and Representation of Others' Pain in the Anterior Insula. *Journal of Cognitive Neuroscience, 25*(2), 258-272.

Kobayashi, H., & Kohshima, S. (2001). Unique morphology of the human eye and its adaptive meaning: comparative studies on external morphology of the primate eye. *Journal of Human Evolution, 40*(5), 419-435. doi:10.1006/jhev.2001.0468

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. A. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 347-480). Chicago: Rand McNally.

Konner, M. (2010). *The evolution of childhood: relationships, emotion, mind*. Cambridge, MA: Belknap Press of Harvard University Press.

Korsgaard, C. M. (2010). Reflections on the evolution of morality. *The Amherst Lecture in Philosophy, 5*, 1-29. Retrieved from Amherst Lecture in Philosophy website: http://www.amherstlecture.org/korsgaard2010/

Kramer, K. L. (2005). Children's Help and the Pace of Reproduction: Cooperative Breeding in Humans. *Evolutionary Anthropology, 14*(6), 224-237. doi:10.1002/evan.20082

Kramer, K. L. (2010). Cooperative Breeding and its Significance to the Demographic Success of Humans. *Annual Review of Anthropology, 39*(1), 417-436. doi:10.1146/annurev.anthro.012809.105054

Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science, 354*(6308), 110-114. doi:10.1126/science.aaf8110

Kübler, S., Owenga, P., Reynolds, S. C., Rucina, S. M., & King, G. C. P. (2015). Animal movements in the Kenya Rift and evidence for the earliest ambush hunting by hominins. *Scientific Reports, 5*, 14011. doi:10.1038/srep14011

http://www.nature.com/articles/srep14011#supplementary-information

Kuhn, T. S. (2000). *The road since Structure*. Chicago: University of Chicago Press.

Lancy, D. F. (1996). *Playing on the mother-ground: cultural routines for children's development*. New York, NY: Guilford press.

Langacker, R. W. (1987). *Foundations of cognitive grammar*. Stanford: Stanford University Press.

Lanjouw, A. (2002). Behavioural adaptations to water scarcity in Tongo chimpanze. In C. Boesch, G. Hohmann, & L. F. Marchant (Eds.), *Behavioural diversity in chimpanzees and bonobos* (pp. 52-60). Cambridge: Cambridge University Press.

Lassiter, D. (2016). Linguistic and Philosophical Considerations on Bayesian Semantics *Deontic Modality* (pp. 82-116). Oxford: Oxford University Press.

Laurence, S., & Margolis, E. (2001). The poverty of the stimulus argument. *British Journal for the Philosophy of Science, 52*(2), 217-276.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Berlin: Springer.

Leavens, D. A., & Hopkins, W. D. (1998). Intentional communication by chimpanzees: a cross-sectional study of the use of referential gestures. *Dev Psychol, 34*(5), 813-822.

LeDoux, J. E. (1993). Emotional networks in the brain. In J. M. Haviland & M. J. Lewis (Eds.), *Handbook of emotions* (pp. 109-118). New York, NY: Guilford.

LeDoux, J. E. (1996). *The emotional brain: the mysterious underpinnings of emotional life*. New York, NY: Simon & Schuster.

Lee, R. B., & DeVore, I. (1969). *Man the hunter*. Chicago: Aldine Pub. Co.

Leigh, S. R., & Shea, B. T. (1995). Ontogeny and the Evolution of Adult Body-Size Dimorphism in Apes. *American Journal of Primatology, 36*(1), 37-60. doi:10.1002/ajp.1350360104

Levenson, R. W. (2003). Blood, sweat, and fears: The autonomic architecture of emotion. *Annals of the New York Academy of Sciences, 1000*, 348-366. doi:10.1196/annals.1280.016

Levine, S., Rottman, J., Davis, T., Stich, S., & Machery, E. (n.d.). *Religion and the scope of the moral domain*.

Levis, N. A., & Pfennig, D. W. (2016). Evaluating 'Plasticity-First' Evolution in Nature: Key Criteria and Empirical Approaches. *Trends in Ecology & Evolution, 31*(7), 563-574. doi:10.1016/j.tree.2016.03.012

Lewis, K. P., & Barton, R. A. (2006). Amygdala size and hypothalamus size predict social play frequency in nonhuman primates: A comparative analysis using independent contrasts. *Journal of Comparative Psychology, 120*(1), 31-37. doi:10.1037/0735-7036.120.1.31

Liebal, K., Behne, T., Carpenter, M., & Tomasello, M. (2009). Infants use shared experience to interpret pointing gestures. *Developmental Science, 12*(2), 264-271. doi:10.1111/j.1467-7687.2008.00758.x

Lieberman, D. E. (2011). *The evolution of the human head*. Cambridge, MA: Belknap Press of Harvard University Press.

Lieberman, D. E., Bramble, D. M., Raichlen, D. A., & Shea, J. J. (2009). Brains, Brawn, and the Evolution of Human Endurance Running Capabilities. In F. E. Grine, J. G. Fleagle, & R. E. Leakey (Eds.), *The First Humans* (pp. 77-92). Dordrecht: Springer Netherlands.

Lieberman, D. E., Carlo, J., de Leon, M. P., & Zollikofer, C. P. E. (2007). A geometric morphometric analysis of heterochrony in the cranium of chimpanzees and bonobos. *Journal of Human Evolution, 52*(6), 647-662. doi:10.1016/j.jhevol.2006.12.005

Liszkowski, U., Carpenter, M., Striano, T., & Tomasello, M. (2006). 12-and 18-month-olds point to provide information for others. *Journal of Cognition and Development, 7*(2), 173-187. doi:10.1207/s15327647jcd0702_2

Locke, J. (1689/1988). Two treatises of government. In P. Laslett (Ed.), (Student ed., pp. 464 p.). Cambridge: Cambridge University Press.

Loewenfeld, I. E., & Lowenstein, O. (1993). *The pupil: anatomy, physiology, and clinical applications* (1st ed.). Detroit, MI: Wayne State University Press.

Lohmann, H., & Tomasello, M. (2003). The role of language in the development of false belief understanding: a training study. *Child Development, 74*(4), 1130-1144.

Lorenz, K. (1940). Durch Domestikation verursachte Störungen arteigenen Verhaltens. *Zeitschrift für angewandte Psychologie und Charakterkunde, 59*(1/2), 2-81.

Lovejoy, C. O. (1981). The Origin of Man. *Science, 211*(4480), 341-350. doi:10.1126/science.211.4480.341

Lovejoy, C. O. (2009). Reexamining Human Origins in Light of *Ardipithecus ramidus*. *Science, 326*(5949), 74, 74e71-74e78. doi:10.1126/science.1175834

Lovejoy, C. O., Latimer, B., Suwa, G., Asfaw, B., & White, T. D. (2009). Combining Prehension and Propulsion: The Foot of *Ardipithecus ramidus*. *Science, 326*(5949), 72, 72e71-72e78. doi:10.1126/science.1175832

Lovejoy, C. O., Simpson, S. W., White, T. D., Asfaw, B., & Suwa, G. (2009). Careful Climbing in the Miocene: The Forelimbs of *Ardipithecus ramidus* and Humans

Are Primitive. *Science, 326*(5949), 70, 70e71-70e78. doi:10.1126/science.1175827

Lowe, E. J. (2006). *The four-category ontology: a metaphysical foundation for natural science*. Oxford: Clarendon Press.

Lowenstein, O., Feinberg, R., & Loewenfeld, I. E. (1963). Pupillary Movements during Acute and Chronic Fatigue: a New Test for the Objective Evaluation of Tiredness. *Investigative Ophthalmology, 2*(2), 138-157.

Machery, E. (2012). Delineating the Moral Domain. *Baltic International Yearbook of Cognition, Logic and Communication, 7*, 1-14. doi:10.4148/biyclc.v7i0.1777

Machery, E., & Mallon, R. (2010). Evolution of morality. In J. M. Doris & M. P. R. Group. (Eds.), *The moral psychology handbook* (pp. xii, 493 p.). Oxford: Oxford University Press.

Mackie, J. L. (1977). *Ethics: inventing right and wrong*. Harmondsworth: Penguin.

MacLean, E., & Hare, B. (2013). Spontaneous Triadic Engagement in Bonobos (*Pan paniscus*) and Chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology, 127*(3), 245-255. doi:10.1037/A0030935

Mah, V. K., & Ford-Jones, E. L. (2012). Spotlight on middle childhood: Rejuvenating the ‘forgotten years’. *Paediatr Child Health, 17*(2), 81-83.

Malti, T., & Krettenauer, T. (2013). The relation of moral emotion attributions to prosocial and antisocial behavior: a meta-analysis. *Child Dev, 84*(2), 397-412. doi:10.1111/j.1467-8624.2012.01851.x

Mameli, M. (2013). Meat made us moral: a hypothesis on the nature and evolution of moral judgment. *Biology & Philosophy, 28*(6), 903-931. doi:10.1007/s10539-013-9401-3

Mameli, M., & Bateson, P. (2011). An evaluation of the concept of innateness. *Philosophical Transactions of the Royal Society B: Biological Sciences, 366*(1563), 436-443. doi:10.1098/rstb.2010.0174

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science, 283*(5398), 77-80. doi:10.1126/science.283.5398.77

Marlowe, F. W. (2005). Hunter-gatherers and human evolution. *Evolutionary Anthropology, 14*(2), 54-67. doi:10.1002/Evan.20046

Marsh, K. L., Richardson, M. J., & Schmidt, R. C. (2009). Social Connection Through Joint Action and Interpersonal Coordination. *Topics in Cognitive Science, 1*(2), 320-339. doi:10.1111/j.1756-8765.2009.01022.x

McClure, S. M., Botvinick, M. M., Yeung, N., Greene, J. D., & Cohen, J. D. (2007). Conflict monitoring in cognition-emotion competition. In J. J. Gross (Ed.), *Handbook of emotion regulation* (pp. 204-226). New York, NY: Guilford Press.

McGhee, G. R. (2011). *Convergent evolution: limited forms most beautiful*. Cambridge, MA: MIT Press.

McGrew, W. C. (1981). The female chimpanzee as a human evolutionary prototype. In F. Dahlberg (Ed.), *Woman the gatherer* (pp. 35-73). New Haven: Yale University Press.

McGrew, W. C. (2010). In search of the last common ancestor: new findings on wild chimpanzees. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365*(1556), 3267-3276. doi:10.1098/rstb.2010.0067

McHenry, H. M. (1991). Sexual Dimorphism in *Australopithecus Afarensis*. *Journal of Human Evolution, 20*(1), 21-32. doi:10.1016/0047-2484(91)90043-U

McPherron, S. P., Alemseged, Z., Marean, C. W., Wynn, J. G., Reed, D., Geraads, D., . . . Bearat, H. A. (2010). Evidence for stone-tool-assisted consumption of animal tissues before 3.39 million years ago at Dikika, Ethiopia. *Nature, 466*(7308), 857-860. doi:10.1038/nature09248

Mead, G. H. (2015). *Mind, self, and society* (C. W. Morris Ed. The definitive edition. ed.). Chicago: University of Chicago Press.

Melis, A. P., Altrichter, K., & Tomasello, M. (2013). Allocation of resources to collaborators and free-riders in 3-year-olds. *Journal of Experimental Child Psychology, 114*(2), 364-370. doi:10.1016/j.jecp.2012.08.006

Melis, A. P., Hare, B., & Tomasello, M. (2006a). Engineering cooperation in chimpanzees: tolerance constraints on cooperation. *Animal Behaviour, 72*, 275-286. doi:DOI 10.1016/j.anbehav.2005.09.018

Melis, A. P., Hare, B., & Tomasello, M. (2006b). Engineering cooperation in chimpanzees: tolerance constraints on cooperation. *Anim. Behav., 72*, 275.

Melis, A. P., Hare, B., & Tomasello, M. (2008). Do chimpanzees reciprocate received favors? *Anim. Behav., 76*, 951.

Melis, A. P., Hare, B., & Tomasello, M. (2009). Chimpanzees coordinate in a negotiation game. *Evolution and Human Behavior, 30*(6), 381-392. doi:10.1016/j.evolhumbehav.2009.05.003

Melis, A. P., Schneider, A.-C., & Tomasello, M. (2011). Chimpanzees, Pan troglodytes, share food in the same way after collaborative and individual food acquisition. *Animal Behaviour, 82*(3), 485-493. doi:10.1016/j.anbehav.2011.05.024

Meristo, M., Hjelmquist, E., & Morgan, G. (2012). How access to language affects theory of mind in deaf children. In M. Siegal & L. Surian (Eds.), *Access to language and cognitive development* (pp. 44-61). Oxford: Oxford University Press.

Meyer-Lindenberg, A., Buckholtz, J. W., Kolachana, B., Hariri, A. R., Pezawas, L., Blasi, G., . . . Weinberger, D. R. (2006). Neural mechanisms of genetic risk for impulsivity and violence in humans. *Proceedings of the National Academy of Sciences of the United States of America, 103*(16), 6269-6274. doi:10.1073/pnas.0511311103

Mikhail, J. (2007). Universal moral grammar: theory, evidence and the future. *Trends in Cognitive Sciences, 11*(4), 143-152. doi:10.1016/j.tics.2006.12.007

Mikhail, J. (2008). Moral cognition and computational theory. In W. Sinnott-Armstrong (Ed.), *Moral Psychology: The Neuroscience of Morality: Emotion, Brain Disorders, and Development* (Vol. 3, pp. 81-92). Cambridge, MA: The MIT Press.

Mikhail, J. (2009). Moral grammar and intuitive jurisprudence: A formal model of unconscious moral and legal knowledge. In B. H. Ross, D. M. Bartels, C. W. Bauman, L. J. Skitka, & D. L. Medin (Eds.), *Psychology of Learning and Motivation, Vol. 50: Moral Judgment and Decision Making*: Academic Press.

Mikhail, J. (2011). *Elements of moral cognition: Rawls' linguistic analogy and the cognitive science of moral and legal judgment*. Cambridge: Cambridge University Press.

Mikhail, J., Sorrentino, C. M., & Spelke, E. S. (1998). Toward a universal moral grammar. In M. A. Gernsbacher & S. J. Derry (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 1250). Mahwah, NJ: Lawrence Erlbaum.

Miller, G. F. (2000). *The mating mind: how sexual choice shaped the evolution of human nature* (1st ed.). New York, NY: Doubleday.

Miller, G. F. (2007). Sexual selection for moral virtues. *Quarterly Review of Biology, 82*(2), 97-125. doi:10.1086/517857

Millikan, R. G. (1999). Historical Kinds and the "Special Sciences". *Philosophical Studies, 95*(1-2), 45-65. doi:10.1023/A:1004532016219

Misch, A., Over, H., & Carpenter, M. (2016). I won't tell: Young children show loyalty to their group by keeping group secrets. *Journal of Experimental Child Psychology, 142*, 96-106. doi:10.1016/j.jecp.2015.09.016

Mitteroecker, P., Gunz, P., Bernhard, M., Schaefer, K., & Bookstein, F. L. (2004). Comparison of cranial ontogenetic trajectories among great apes and humans. *Journal of Human Evolution, 46*(6), 679-697. doi:10.1016/j.jhevol.2004.03.006

Mitteroecker, P., Gunz, P., & Bookstein, F. L. (2005). Heterochrony and geometric morphometrics: a comparison of cranial growth in *Pan paniscus* versus *Pan troglodytes*. *Evolution & Development, 7*(3), 244-258. doi:10.1111/j.1525-142X.2005.05027.x

Moll, J., Zahn, R., de Oliveira-Souza, R., Krueger, F., & Grafman, J. (2005). The neural basis of human moral cognition. *Nature Reviews Neuroscience, 6*(10), 799-809. doi:10.1038/Nrn1768

Moore, G. E. (1903). *Principia ethica*. Cambridge: Cambridge University press.

Moore, J. (1996). Savanna chimpanzees, referential models and the last common ancestor. In W. C. McGrew, L. F. Marchant, & T. Nishida (Eds.), *Great ape societies* (pp. 275-292). Cambridge: Cambridge University Press.

Moore, R., Mueller, B., Kaminski, J., & Tomasello, M. (2015). Two-year-old children but not domestic dogs understand communicative intentions without language, gestures, or gaze. *Dev Sci, 18*(2), 232-242. doi:10.1111/desc.12206

Moran, N. A. (1992). The Evolutionary Maintenance of Alternative Phenotypes. *The American Naturalist, 139*(5), 971-989.

Mundy, P., & Jarrold, W. (2010). Infant joint attention, neural networks and social cognition. *Neural Networks, 23*(8-9), 985-997. doi:10.1016/j.neunet.2010.08.009

Nado, J., Kelly, D., & Stich, S. (2009). Moral judgment. In J. Symons & P. Calvo (Eds.), *The Routledge companion to philosophy of psychology* (pp. 621-633). London: Routledge.

Nagel, T. (1986). *The view from nowhere*. Oxford: Oxford University Press.

Nelissen, R. M. A., Dijker, A. J. M., & deVries, N. K. (2007). How to turn a hawk into a dove and vice versa: Interactions between emotions and goals in a give-some dilemma game. *Journal of Experimental Social Psychology, 43*(2), 280-286. doi:10.1016/j.jesp.2006.01.009

Nelissen, R. M. A., & Zeelenberg, M. (2009). When Guilt Evokes Self-Punishment: Evidence for the Existence of a Dobby Effect. *Emotion, 9*(1), 118-122. doi:10.1037/a0014540

Nesse, R. M. (2007). Runaway Social Selection for Displays of Partner Value and Altruism. *Biological Theory, 2*(2), 143-155.

New, C. (1974). Saints, Heroes and Utilitarians. *Philosophy, 49*(188), 179-189. doi:10.1017/S0031819100048051

Newman, S. W. (1999). The medial extended amygdala in male reproductive behavior: A node in the mammalian social behavior network. *Advancing from the Ventral Striatum to the Extended Amygdala, 877*, 242-257. doi:10.1111/j.1749-6632.1999.tb09271.x

Nichols, S. (2001). Mindreading and the cognitive architecture underlying altruistic motivation. *Mind & Language, 16*(4), 425-455. doi:10.1111/1468-0017.00178

Nichols, S. (2002). Norms with feeling: towards a psychological account of moral judgment. *Cognition, 84*(2), 221-236. doi:10.1016/S0010-0277(02)00048-3

Nichols, S. (2004). *Sentimental Rules: On the Natural Foundations of Moral Judgment*: Oxford University Press.

Nichols, S. (2005). Innateness and moral psychology. In P. Carruthers, S. Laurence, & S. P. Stich (Eds.), *The Innate Mind: Structure and Contents* (pp. 353-369). Oxford: Oxford University Press.

Nichols, S., & Folds-Bennett, T. (2003). Are children moral objectivists? Children’s judgments about moral and response-dependent properties. *Cognition, 90*(2), B23-32.

Niedenthal, P. M., & Cantor, N. (1984). Making use of social prototypes: From fuzzy concepts to firm decisions. *Fuzzy Sets and Systems, 14*, 5-28.

Nietzsche, F. W. (1887/1967). On the genealogy of morals. In W. A. Kaufmann (Ed.), (pp. viii, 367 p.). New York, NY: Vintage Books.

Nisan, M. (1987). Moral Norms and Social Conventions: A Cross-Cultural Comparison. *Developmental Psychology, 23*(5), 719-725. doi:10.1037/0012-1649.23.5.719

Noë, R., & Hammerstein, P. (1994). Biological markets: supply and demand determine the effect of partner choice in cooperation, mutualism and mating. *Behavioral Ecology and Sociobiology, 35*(1), 1-11. doi:10.1007/bf00167053

Nucci, L. P. (1982). Conceptual Development in the Moral and Conventional Domains: Implications for Values Education. *Review of Educational Research, 52*(1), 93-122. doi:10.2307/1170274

Nucci, L. P. (2001). *Education in the moral domain*. Cambridge: Cambridge University Press.

Nucci, L. P., & Herman, S. (1982). Behavioral disordered children’s conceptions of moral, conventional, and personal issues. *Journal of Abnormal Child Psychology, 10*(3), 411-425. doi:10.1007/BF00912330

Nucci, L. P., & Lee, J. (1993). Morality and personal autonomy. In G. Noam & T. E. Wren (Eds.), *The moral self* (pp. 123-148). Cambridge, MA: MIT Press.

Nucci, L. P., & Nucci, M. S. (1982). Children's Social Interactions in the Context of Moral and Conventional Transgressions. *Child Development, 53*(2), 403-412. doi:10.2307/1128983

Nucci, L. P., & Turiel, E. (1978). Social Interactions and the Development of Social Concepts in Preschool Children. *Child Development, 49*(2), 400-407. doi:10.2307/1128704

Nucci, L. P., & Turiel, E. (1993). God's Word, Religious Rules, and Their Relation to Christian and Jewish Children's Concepts of Morality. *Child Development, 64*(5), 1475-1491. doi:10.2307/1131547

Nucci, L. P., Turiel, E., & Encarnacion-Gawrych, G. (1983). Children's Social Interactions and Social Concepts: Analyses of Morality and Convention in the Virgin Islands. *Journal of Cross-Cultural Psychology, 14*(4), 469-487. doi:10.1177/0022002183014004006

Nunnally, J. C., Knott, P. D., Duchnowski, A., & Parker, R. (1967). Pupillary Response as a General Measure of Activation. *Perception & Psychophysics, 2*(4), 149-155. doi:10.3758/Bf03210310

Nunner-Winkler, G., & Sodian, B. (1988). Children's understanding of moral emotions. *Child Dev, 59*(5), 1323-1338.

O'Connell, J. F., Hawkes, K., & Blurton Jones, N. G. (1999). Grandmothering and the evolution of *Homo erectus*. *Journal of Human Evolution, 36*(5), 461-485. doi:10.1006/jhev.1998.0285

O'Connor, C. (n.d.). *The Evolution of Guilt: a Model-Based Approach*. University of California, Irvine, USA. Unpublished manuscript.

Ohtsubo, Y., & Watanabe, E. (2009). Do sincere apologies need to be costly? Test of a costly signaling model of apology. *Evolution and Human Behavior, 30*(2), 114-123. doi:10.1016/j.evolhumbehav.2008.09.004

Okamoto, S., Tanaka, M., & Tomonaga, M. (2004). Looking back: The "representational mechanism" of joint attention in an infant chimpanzee (*Pan troglodytes*). *Japanese Psychological Research, 46*(3), 236-245. doi:10.1111/j.1468-5584.2004.00255.x

Okamoto, S., Tomonaga, M., Ishii, K., Kawai, N., Tanaka, M., & Matsuzawa, T. (2002). An infant chimpanzee (*Pan troglodytes*) follows human gaze. *Animal Cognition, 5*(3), 191-191. doi:10.1007/s10071-002-0138-7

Olson, J. (2010). In Defense of Moral Error Theory. In M. Brady (Ed.), *New Waves in Metaethics* (pp. 62-84). Basingstoke: Palgrave Macmillan.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-Month-Old Infants Understand False Beliefs? *Science, 308*(5719), 255-258. doi:10.1126/science.1107621

Oyen, O. J. (1979). Tool-use in free-ranging baboons of Nairobi National Park. *Primates, 20*(4), 595-597. doi:10.1007/BF02373442

Pääbo, S. (2003). The mosaic that is our genome. *Nature, 421*(6921), 409-412.

Pacherie, E., & Dokic, J. (2006). From mirror neurons to joint actions. *Cognitive Systems Research, 7*(2-3), 101-112. doi:10.1016/j.cogsys.2005.11.012

Palagi, E. (2006). Social play in bonobos (*Pan paniscus*) and chimpanzees (*Pan troglodytes*): Implications for natural social systems and interindividual relationships. *American Journal of Physical Anthropology, 129*(3), 418-426. doi:10.1002/ajpa.20289

Parish, A. R. (1996). Female relationships in bonobos (*Pan paniscus*): Evidence for bonding, cooperation, and female dominance in a male-philopatric species. *Human Nature, 7*(1), 61-96.

Park, J. J. (2013). Prototypes, Exemplars, and Theoretical & Applied Ethics. *Neuroethics, 6*(2), 237-247. doi:10.1007/s12152-011-9106-8

Parkinson, C., Sinnott-Armstrong, W., Koralus, P. E., Mendelovici, A., McGeer, V., & Wheatley, T. (2011). Is Morality Unified? Evidence that Distinct Neural Systems Underlie Moral Judgments of Harm, Dishonesty, and Disgust. *Journal of Cognitive Neuroscience, 23*(10), 3162-3180.

Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies, 59*(1-2), 185-198. doi:10.1016/S1071-5819(03)00017-X

Paukner, A., Anderson, J. R., & Fujita, K. (2006). Redundant food searches by capuchin monkeys (*Cebus apella*): a failure of metacognition? *Anim Cogn, 9*(2), 110-117. doi:10.1007/s10071-005-0007-2

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical Learning in a Natural Language by 8-Month-Old Infants. *Child Development, 80*(3), 674-685. doi:10.1111/j.1467-8624.2009.01290.x

Penin, X., Berge, C., & Baylac, M. (2002). Ontogenetic study of the skull in modern humans and the common chimpanzees: Neotenic hypothesis reconsidered with a tridimensional procrustes analysis. *American Journal of Physical Anthropology, 118*(1), 50-62. doi:10.1002/Ajpa.10044

Perry, G. H., Dominy, N. J., Claw, K. G., Lee, A. S., Fiegler, H., Redon, R., . . . Stone, A. C. (2007). Diet and the evolution of human amylase gene copy number variation. *Nature Genetics, 39*(10), 1256-1260. doi:10.1038/ng2123

Perry, S. (1997). Male-female social relationships in wild white-faced capuchins (Cebus capucinus). *Behaviour, 134*, 477-510. doi:10.1163/156853997x00494

Perry, S., Baker, M., Fedigan, L., Gros-Louis, J., Jack, K., MacKinnon, K. C., . . . Rose, L. (2003). Social conventions in wild white-faced capuchin monkeys: Evidence for traditions in a neotropical primate. *Current Anthropology, 44*(2), 241-268. doi:10.1086/345825

Pettit, P. (1990). *Virtus Normativa*: Rational Choice Perspectives. *Ethics, 100*(4), 725-755. doi:10.2307/2381776

Pezawas, L., Meyer-Lindenberg, A., Drabant, E. M., Verchinski, B. A., Munoz, K. E., Kolachana, B. S., . . . Weinberger, D. R. (2005). 5-HTTLPR polymorphism impacts human cingulate-amygdala interactions: a genetic susceptibility mechanism for depression. *Nature Neuroscience, 8*(6), 828-834. doi:10.1038/Nn1463

Phillips, K. A. (1998). Tool use in wild capuchin monkeys (*Cebus albifrons trinitatis*). *American Journal of Primatology, 46*(3), 259-261. doi:10.1002/(Sici)1098-2345(1998)46:3<259::Aid-Ajp6>3.0.Co;2-R

Piaget, J. (1932). *The moral judgment of the child* (M. Gabain, Trans.). London: K. Paul, Trench, Trubner & Co. ltd.

Piaget, J. (1952). *The origins of intelligence in children*. New York, NY: International Universities Press.

Pickering, T. R. (2013). Rough and tumble: aggression, hunting, and human evolution. Berkeley: University of California Press.

Pickford, M., & Senut, B. (2001). The geological and faunal context of Late Miocene hominid remains from Lukeino, Kenya. *Comptes Rendus de l'Académie des Sciences, 332*(2), 145-152. doi:10.1016/S1251-8050(01)01528-2

Pika, S., & Zuberbühler, K. (2008). Social games between bonobos and humans: Evidence for shared intentionality? *American Journal of Primatology, 70*(3), 207-210. doi:10.1002/Ajp.20469

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Pinker, S. (1994). *The language instinct* (1st ed.). New York, NY: W. Morrow and Co.

Pinker, S. (2011). *The better angels of our nature: why violence has declined*. New York, NY: Viking.

Pizarro, D. (2000). Nothing more than feelings? The role of emotions in moral judgment. *Journal for the Theory of Social Behaviour, 30*(4), 355-375.

Pizarro, D. (2006). The Creativity of Everyday Moral Reasoning: Empathy, Disgust and Moral Persuasion. In J. C. Kaufman & J. Baer (Eds.), *Creativity and Reason in Cognitive Development* (pp. 81-98). Cambridge: Cambridge University Press.

Plavcan, J. M., Lockwood, C. A., Kimbel, W. H., Lague, M. R., & Harmon, E. H. (2005). Sexual dimorphism in *Australopithecus afarensis* revisited: How strong is the case for a human-like pattern of dimorphism? *Journal of Human Evolution*, *48*(3), 313-320. doi:10.1016/j.jhevol.2004.09.006

Plavcan, J. M., & van Schaik, C. P. (1997). Interpreting hominid behavior on the basis of sexual dimorphism. *Journal of Human Evolution, 32*(4), 345-374. doi:10.1006/jhev.1996.0096

Plooij, F. X. (1984). *The behavioral development of free-living chimpanzee babies and infants*. Norwood, NJ: Ablex Pub. Corp.

Ponce de Leon, M. S., & Zollikofer, C. P. E. (2001). Neanderthal cranial ontogeny and its implications for late hominid diversity. *Nature, 412*(6846), 534-538.

Post, D. G. (1982). Feeding behavior of yellow baboons (*Papio cynocephalusin*) the Amboseli National Park, Kenya. *International Journal of Primatology, 3*(4), 403-430. doi:10.1007/BF02693741

Potts, R. (1984). Home Bases and Early Hominids. *American Scientist*, *72*(4), 338-347.

Prinz, J. J. (2007). *The Emotional Construction of Morals*. Oxford: Oxford University Press.

Prinz, J. J. (2009). Against moral nativism. In D. Murphy & M. A. Bishop (Eds.), *Stich and his critics* (pp. 167-189). Malden, MA: Wiley-Blackwell.

Prüfer, K., Munch, K., Hellmann, I., Akagi, K., Miller, J. R., Walenz, B., . . . Paabo, S. (2012). The bonobo genome compared with the chimpanzee and human genomes. *Nature, 486*(7404), 527-531. doi:10.1038/Nature11128

Pybus, E. M. (1982). Saints and Heroes. *Philosophy, 57*(220), 193-199. doi:10.1017/S0031819100050762

Pyers, J. E., & Senghas, A. (2009). Language promotes false-belief understanding: evidence from learners of a new sign language. *Psychological Science, 20*(7), 805-812. doi:10.1111/j.1467-9280.2009.02377.x

Quine, W. V. (1969). Natural kinds *Ontological relativity and other essays* (pp. 114-138). New York, NY: Columbia University Press.

Quintelier, K. J. P., & Fessler, D. M. T. (2015). Confounds in moral/conventional studies. *Philosophical Explorations, 18*(1), 58-67. doi:10.1080/13869795.2013.874496

Quintelier, K. J. P., Fessler, D. M. T., & De Smet, D. (2012). The case of the drunken sailor: On the generalisable wrongness of harmful transgressions. *Thinking & Reasoning, 18*(2), 183-195. doi:10.1080/13546783.2012.669738

Rakison, D. H., & Oakes, L. M. (2003). *Early category and concept development: making sense of the blooming, buzzing confusion*. Oxford: Oxford University Press.

Rakoczy, H. (2008). Taking fiction seriously: Young children understand the normative structure of joint pretence games. *Developmental Psychology, 44*(4), 1195-1201. doi:10.1037/0012-1649.44.4.1195

Rakoczy, H., Brosche, N., Warneken, F., & Tomasello, M. (2009). Young children’s understanding of the context-relativity of normative rules in conventional games. *British Journal of Developmental Psychology, 27*, 445-456. doi:10.1348/026151008x337752

Rakoczy, H., Gräfenhain, M., Cluver, A., Dalhoff, A. C., & Sternkopf, A. (2014). Young children’s agent-neutral representations of action roles. *J Exp Child Psychol, 128*, 201-209. doi:10.1016/j.jecp.2014.06.004

Rakoczy, H., Hamann, K., Warneken, F., & Tomasello, M. (2010). Bigger knows better: Young children selectively learn rule games from adults rather than from peers. *British Journal of Developmental Psychology, 28*(4), 785-798. doi:10.1348/026151009x479178

Rakoczy, H., & Schmidt, M. F. H. (2013). The Early Ontogeny of Social Norms. *Child Development Perspectives, 7*(1), 17-21. doi:10.1111/Cdep.12010

Rakoczy, H., Wameken, F., & Tomasello, M. (2008). The sources of normativity: Young children’s awareness of the normative structure of games. *Developmental Psychology, 44*(3), 875-881. doi:10.1037/0012-1649.44.3.875

Rakoczy, H., Warneken, F., & Tomasello, M. (2009). Young children’s selective learning of rule games from reliable and unreliable models. *Cognitive Development, 24*(1), 61-69. doi:10.1016/j.cogdev.2008.07.004

Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science, 288*(5464), 349-351. doi:10.1126/science.288.5464.349

Rauss, K., & Pourtois, G. (2013). What is Bottom-Up and What is Top-Down in Predictive Coding? *Frontiers in Psychology, 4,* 276. doi:10.3389/fpsyg.2013.00276

Reddy, V. (2015). Joining intentions in infancy. *Journal of Consciousness Studies, 22*(1-2), 24-44.

Regan, J. W. (1971). Guilt, perceived injustice, and altruistic behavior. *J Pers Soc Psychol, 18*(1), 124-132.

Rekers, Y., Haun, D. B. M., & Tomasello, M. (2011). Children, but not chimpanzees, prefer to collaborate. *Current Biology, 21*(20), 1756-1758.

Reno, P. L., McCollum, M. A., Meindl, R. S., & Lovejoy, C. O. (2010). An enlarged postcranial sample confirms *Australopithecus afarensis* dimorphism was similar to modern humans. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365*(1556), 3355-3363. doi:10.1098/rstb.2010.0086

Reno, P. L., Meindl, R. S., McCollum, M. A., & Lovejoy, C. O. (2003). Sexual dimorphism in *Australopithecus afarensis* was similar to that of modern humans. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(16), 9404-9409. doi:10.1073/pnas.1133180100

Rilling, J. K., Scholz, J., Preuss, T. M., Glasser, M. F., Errangi, B. K., & Behrens, T. E. (2012). Differences between chimpanzees and bonobos in neural systems supporting social cognition. *Social Cognitive and Affective Neuroscience, 7*(4), 369-379. doi:10.1093/scan/nsr017

Roach, N. T., Venkadesan, M., Rainbow, M. J., & Lieberman, D. E. (2013). Elastic energy storage in the shoulder and the evolution of high-speed throwing in *Homo*. *Nature, 498*(7455), 483-486. doi:10.1038/nature12267

http://www.nature.com/nature/journal/v498/n7455/abs/nature12267.html#supplementary-information

Roberts, I., & Holmberg, A. (2005). On the role of parameters in universal grammar: a reply to Newmeyer. In H. Broekhuis, N. Corver, R. Huybregts, U. Kleinhenz, & J. Koster (Eds.), *Organizing grammar: studies in honor of Henk van Riemsdijk* (pp. 538-553). Berlin: Mouton de Gruyter.

Roberts, S. (1979). *Order and dispute: an introduction to legal anthropology*. Oxford: Martin Robertson.

Rochat, P. (2009). *Others in mind: social origins of self-consciousness*. Cambridge, MA: Cambridge University Press.

Rosati, A. G., & Hare, B. (2012). Chimpanzees and bonobos exhibit divergent spatial memory development. *Developmental Science, 15*(6), 840-853. doi:10.1111/j.1467-7687.2012.01182.x

Rossano, F., Rakoczy, H., & Tomasello, M. (2011). Young children's understanding of violations of property rights. *Cognition, 121*(2), 219-227. doi:10.1016/j.cognition.2011.06.007

Rousseau, J.-J. (1755/1992). Discourse on the origin of inequality. In D. A. Cress (Ed.), (pp. xx, 92 p.). Indianapolis: Hackett Pub. Co.

Rousseau, J.-J. (1762/1987). On the social contract. In D. A. Cress (Ed.), (pp. vi, 103 p.). Indianapolis: Hackett.

Rowland, M. G. M., Barrell, R. A. E., & Whitehead, R. G. (1978). The weanling's dilemma: Bacterial contamination in traditional Gambian weaning foods. *The Lancet, 311*(8056), 136-138. doi:10.1016/S0140-6736(78)90432-4

Ruse, M. (1986). Evolutionary Ethics: a Phoenix Arisen. *Zygon, 21*(1), 95-112. doi:10.1111/j.1467-9744.1986.tb00736.x

Ruse, M. (1998). *Taking Darwin seriously: a naturalistic approach to philosophy*. Amherst, NY: Prometheus Books.

Ruse, M., & Wilson, E. O. (1986). Moral Philosophy as Applied Science. *Philosophy, 61*(236), 173-192.

Ryle, G. (1971). The Thinking of Thoughts: What is 'Le Penseur' Doing? *Collected Papers* (Vol. 2, pp. 480-496). London: Hutchinson.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926-1928. doi:10.1126/science.274.5294.1926

Salomo, D., & Liszkowski, U. (2013). Sociocultural settings influence the emergence of prelinguistic deictic gestures. *Child Development, 84*(4), 1296-1307. doi:10.1111/cdev.12026

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind". *NeuroImage, 19*(4), 1835-1842. doi:10.1016/S1053-8119(03)00230-1

Saxe, R., Xiao, D. K., Kovacs, G., Perrett, D. I., & Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions.

Neuropsychologia, 42(11), 1435-1446. doi:10.1016/j.neuropsychologia.2004.04.015

Sayers, K., Raghanti, M. A., & Lovejoy, C. O. (2012). Human Evolution and the Chimpanzee Referential Doctrine. *Annual Review of Anthropology, 41*, 119-138. doi:10.1146/annurev-anthro-092611-145815

Sayre-McCord, G. (2007). Metaethics. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2014 Edition ed.).

Scally, A., Dutheil, J. Y., Hillier, L. W., Jordan, G. E., Goodhead, I., Herrero, J., . . . Durbin, R. (2012). Insights into hominid evolution from the gorilla genome sequence. *Nature, 483*(7388), 169-175. doi:http://www.nature.com/nature/journal/v483/n7388/abs/nature10842.html#supplementary-information

Scanlon, T. (1998). *What we owe to each other*. Cambridge, MA: Belknap Press of Harvard University Press.

Schmidt, M. F. H., Butler, L. P., Heinz, J., & Tomasello, M. (2016). Young Children See a Single Action and Infer a Social Norm: Promiscuous Normativity in 3-Year-Olds. *Psychological Science, 27*(10), 1360-1370. doi:10.1177/0956797616661182

Schmidt, M. F. H., Gonzalez-Cabrera, I., & Tomasello, M. (2017). Children's developing metaethical judgments. *Journal of Experimental Child Psychology, 164*, 163-177. doi:https://doi.org/10.1016/j.jecp.2017.07.008

Schmidt, M. F. H., Rakoczy, H., & Tomasello, M. (2011a). Young children attribute normativity to novel actions without pedagogy or normative language. *Developmental Science, 14*(3), 530-539. doi:10.1111/j.1467-7687.2010.01000.x

Schmidt, M. F. H., Rakoczy, H., & Tomasello, M. (2011b). Young children attribute normativity to novel actions without pedagogy or normative language. *Dev Sci, 14*(3), 530-539. doi:10.1111/j.1467-7687.2010.01000.x

Schmidt, M. F. H., Rakoczy, H., & Tomasello, M. (2012). Young children enforce social norms selectively depending on the violator's group affiliation. *Cognition, 124*(3), 325-333. doi:10.1016/j.cognition.2012.06.004

Schmidt, M. F. H., & Tomasello, M. (2012). Young Children Enforce Social Norms. *Current Directions in Psychological Science, 21*(4), 232-236. doi:10.1177/0963721412448659

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin, 34*(8), 1096-1109. doi:10.1177/0146167208317771

Schneider, W. (1999). Automaticity. In R. A. Wilson & F. C. Keil (Eds.), *The MIT encyclopedia of the cognitive sciences* (pp. 63-64). Cambridge, MA: MIT Press.

Schroeder, M. A. (2008). *Being for: evaluating the semantic program of expressivism*. Oxford: Oxford University Press.

Schroeder, M. A. (2012). Philosophy of language for metaethics. In G. Russell & D. Graff Fara (Eds.), *The Routledge Companion to the Philosophy of Language* (pp. 705-715). London: Routledge.

Schroepfer-Walker, K., Wobber, V., & Hare, B. (2015). Experimental evidence that grooming and play are social currency in bonobos and chimpanzees. *Behaviour, 152*(3-4), 545-562. doi:10.1163/1568539x-00003258

Schultz, A. H. (1941). The relative size of the cranial capacity in primates. *American Journal of Physical Anthropology, 28*(3), 273-287. doi:10.1002/ajpa.1330280303

Schwarcz, H. P., Grün, R., Vandermeersch, B., Bar-Yosef, O., Valladas, H., & Tchernov, E. (1988). ESR dates for the hominid burial site of Qafzeh in Israel. *Journal of Human Evolution, 17*(8), 733-737. doi:10.1016/0047-2484(88)90063-2

Schwartz, J. H. (1987). *The red ape: orang-utans and human origins*. Boston: Houghton Mifflin.

Schwartz, J. H. (2004). Trying to make chimpanzees into humans. *History and Philosophy of the Life Sciences, 26*(2), 271-277.

Seabright, P. (2004). *The Company of Strangers: A Natural History of Economic Life*. Princeton, NJ: Princeton University Press.

Sear, R., Mace, R., & McGregor, I. A. (2000). Maternal grandmothers improve nutritional status and survival of children in rural Gambia. *Proceedings of the Royal Society B: Biological Sciences, 267*(1453), 1641-1647. doi:10.1098/rspb.2000.1190

Searle, J. R. (1990). Collective Intentions and Actions. In P. R. Cohen, J. L. Morgan, & M. E. Pollack (Eds.), *Intentions in communication* (pp. 401-415). Cambridge, MA: MIT Press.

Searle, J. R. (1995). *The construction of social reality*. London: Allen Lane.

Searle, J. R. (1998). *Mind, Language and Society: Philosophy in the Real World*. New York, NY: Basic Books.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences, 10*(2), 70-76. doi:10.1016/j.tics.2005.12.009

Seeger, A. (1981). *Nature and Society in Central Brazil: The Suyá Indians of Mato Grosso*. Cambridge, MA: Harvard University Press.

Sellars, W. (1963). Imperatives, Intentions, and the Logic of 'Ought'. In H.-N. Castañeda & G. Nakhnikian (Eds.), *Morality and the language of conduct* (pp. 159-214). Detroit: Wayne State University Press.

Shea, B. T. (1983). Paedomorphosis and neoteny in the pygmy chimpanzee. *Science, 222*(4623), 521-522.

Shea, B. T. (1989). Heterochrony in Human-Evolution: the Case for Neoteny Reconsidered. *Yearbook of Physical Anthropology, 32*, 69-101.

Shipman, P. (1986). Scavenging or Hunting in Early Hominids: Theoretical Framework and Tests. *American Anthropologist, 88*(1), 27-43. doi:10.1525/aa.1986.88.1.02a00020

Shweder, R. A., Mahapatra, M., & Miller, J. (1987). Culture and moral development. In J. Kagan & S. Lamb (Eds.), *The emergence of morality in young children* (pp. 1-83). Chicago: University of Chicago Press.

Sierksma, J., Thijs, J., & Verkuyten, M. (2014). Ethnic Helping and Group Identity: A Study among Majority Group Children. *Social Development, 23*(4), 803-819. doi:10.1111/sode.12077

Silfver, M. (2007). Coping with guilt and shame: a narrative approach. *Journal of Moral Education, 36*(2), 169-183. doi:10.1080/03057240701325274

Sinnott-Armstrong, W. (2008). Is moral phenomenology unified? *Phenomenology and the Cognitive Sciences, 7*(1), 85-97. doi:10.1007/s11097-007-9065-z

Sinnott-Armstrong, W., & Wheatley, T. (2012). The disunity of morality and why it matters to philosophy. *The Monist, 95*(3), 355-377.

Sinnott-Armstrong, W., & Wheatley, T. (2013). Are moral judgments unified? *Philosophical Psychology, 27*(4), 451-474. doi:10.1080/09515089.2012.736075

Skinner, B. F. (1957). *Verbal Behavior*. New York, NY: Appleton-Century-Crofts.

Skybreak, A. (1984). *Of primeval steps and future leaps: an essay on the emergence of human beings, the source of women's oppression, and the road to emancipation*. Chicago: Banner Press.

Smetana, J. G. (1984). Toddlers' social interactions regarding moral and conventional transgressions. *Child Development, 55*(5), 1767-1776.

Smetana, J. G. (1985). Preschool children's conceptions of transgressions: Effects of varying moral and conventional domain-related attributes. *Developmental Psychology, 21*(1), 18-29. doi:10.1037/0012-1649.21.1.18

Smetana, J. G. (1993). Understanding of social rules. In M. Bennett (Ed.), *The Development of social cognition: the child as psychologist* (pp. 111-114). New York/London: Guilford Press.

Smetana, J. G. (1995). Morality in context: Abstractions, ambiguities, and applications. In R. Vasta (Ed.), *Annals of Child Development* (Vol. 10, pp. 83-139). London: Jessica Kingsley.

Smetana, J. G. (1997). Parenting and the development of social knowledge reconceptualized: A social domain analysis. In E. Grusec & L. Kuczynski (Eds.), *Parenting and the internalization of values* (pp. 162-192). New York, NY: Wiley.

Smetana, J. G., & Braeges, J. L. (1990). The Development of Toddlers Moral and Conventional Judgments. *Merrill-Palmer Quarterly-Journal of Developmental Psychology, 36*(3), 329-346.

Smetana, J. G., Kelly, M., & Twentyman, C. T. (1984). Abused, Neglected, and Nonmaltreated Children's Conceptions of Moral and Social-Conventional Transgressions. *Child Development, 55*(1), 277-287. doi:10.2307/1129852

Smetana, J. G., Killen, M., & Turiel, E. (2000). Children's reasoning about moral conflicts. In W. Craig (Ed.), *Childhood social development: The essential readings* (pp. 275-302). Oxford: Blackwell.

Smetana, J. G., Toth, S. L., Cicchetti, D., Bruce, J., Kane, P., & Daddis, C. (1999). Maltreated and nonmaltreated preschoolers' conceptions of hypothetical and actual moral transgressions. *Developmental Psychology, 35*(1), 269-281. doi:10.1037/0012-1649.35.1.269

Smith, A. (1759/2002). The theory of moral sentiments. In K. Haakonssen (Ed.), (pp. 551 p.). Cambridge: Cambridge University Press.

Smith, M. (1994). *The moral problem*. Oxford: Blackwell.

Smith, T. M., Tafforeau, P., Le Cabec, A., Bonnin, A., Houssaye, A., Pouech, J., . . . Menter, C. G. (2015). Dental Ontogeny in Pliocene and Early Pleistocene Hominins. *PLoS ONE, 10*(2), e0118118. doi:10.1371/journal.pone.0118118

Smith, T. M., Tafforeau, P., Reid, D. J., Grün, R., Eggins, S., Boutakiout, M., & Hublin, J.-J. (2007). Earliest evidence of modern human life history in North African early Homo sapiens. *Proceedings of the National Academy of Sciences, 104*(15), 6128-6133. doi:10.1073/pnas.0700747104

Sober, E., & Wilson, D. S. (1998). *Unto others: the evolution and psychology of unselfish behavior*. Cambridge, MA: Harvard University Press.

Solso, R. L., & McCarthy, J. E. (1981). Prototype formation of faces: A case of pseudo-memory. *British Journal of Psychology, 72*(4), 499-503. doi:10.1111/j.2044-8295.1981.tb01779.x

Spivey, M. (2007). *The continuity of mind*. Oxford: Oxford University Press.

Sripada, C. S., & Stich, S. (2007). A Framework for the Psychology of Norms. In P. Carruthers, S. Laurence, & S. P. Stich (Eds.), *The Innate Mind: Culture and Cognition* (Vol. 2, pp. 280-301). Oxford: Oxford University Press.

Stanford, C. B. (1998a). *Chimpanzee and red colobus: the ecology of predator and prey*. Cambridge, MA: Harvard University Press.

Stanford, C. B. (1998b). The social behavior of chimpanzees and bonobos: empirical evidence and shifting assumptions. *Current Anthropology, 39*(4), 399-420. doi:10.1086/204757

Stanford, C. B., & Allen, J. S. (1991). On Strategic Storytelling: Current Models of Human Behavioral Evolution. *Current Anthropology, 32*(1), 58-61. doi:10.1086/203914

Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Stanovich, K. E., & Toplak, M. E. (2012). Defining features versus incidental correlates of Type 1 and Type 2 processing. *Mind and Society, 11*(1), 3-13. doi:10.1007/s11299-011-0093-6

Sterelny, K. (2003). *Thought in a hostile world: the evolution of human cognition*. Malden, MA: Blackwell.

Sterelny, K. (2010). Moral Nativism: A Sceptical Response. *Mind & Language, 25*(3), 279-297.

Sterelny, K. (2012a). *The evolved apprentice: how evolution made humans unique*. Cambridge, MA: The MIT Press.

Sterelny, K. (2012b). Morality’s Dark Past. *Analyse & Kritik, 34*(1), 95-115.

Sterelny, K. (2013). Life in Interesting Times: Cooperation and Collective Action in the Holocene. In K. Sterelny, R. Joyce, B. Calcott, & B. Fraser (Eds.), *Cooperation and Its Evolution* (pp. 89-108). Cambridge, MA: MIT Press.

Sterelny, K. (2014). A Paleolithic Reciprocation Crisis: Symbols, Signals, and Norms. *Biological Theory, 9*(1), 65-77. doi:10.1007/s13752-013-0143-x

Sterelny, K. (2016). Cooperation, Culture, and Conflict. *The British Journal for the Philosophy of Science, 67*(1), 31-58. doi:10.1093/bjps/axu024

Stich, S. (1993). Moral philosophy and mental representation. In M. Hechter, L. Nadel, & R. E. Michod (Eds.), *The Origin of values* (pp. 215-228). New York, NY: Aldine de Gruyter.

Stich, S. (2006). Is Morality an Elegant Machine or a Kludge? *Journal of Cognition and Culture, 6*(1-2), 181-189. doi:10.1163/156853706776931349

Stich, S. (2016). The moral domain. In K. Grey & J. Graham (Eds.), *The Atlas of Moral Psychology: Mapping Good and Evil in the Mind*. New York, NY: Guilford.

Street, S. (2006). A Darwinian Dilemma for Realist Theories of Value. *Philosophical Studies, 127*(1), 109-166. doi:10.1007/s11098-005-1726-6

Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers Can Attribute Second-Order Beliefs. *Developmental Psychology, 30*(3), 395-402. doi:10.1037/0012-1649.30.3.395

Surbeck, M., Deschner, T., Schubert, G., Weltring, A., & Hohmann, G. (2012). Mate competition, testosterone and intersexual relationships in bonobos, Pan paniscus. *Animal Behaviour, 83*(3), 659-669. doi:10.1016/j.anbehav.2011.12.010

Surbeck, M., Fowler, A., Deimel, C., & Hohmann, G. (2009). Evidence for the Consumption of Arboreal, Diurnal Primates by Bonobos (*Pan paniscus*). *American Journal of Primatology, 71*(2), 171-174. doi:10.1002/Ajp.20634

Surbeck, M., & Hohmann, G. (2008). Primate hunting by bonobos at LuiKotale, Salonga National Park. *Current Biology, 18*(19), R906-R907. doi:10.1016/j.cub.2008.08.040

Sussman, R. W. (1999). The Myth of Man the Hunter, Man the Killer and the Evolution of Human Morality. *Zygon, 34*(3), 453-471. doi:10.1111/0591-2385.00226

Suwa, G., Kono, R. T., Simpson, S. W., Asfaw, B., Lovejoy, C. O., & White, T. D. (2009). Paleobiological Implications of the *Ardipithecus ramidus* Dentition. *Science, 326*(5949), 69, 94-99. doi:10.1126/science.1175824

Svensson, R., Weerman, F. M., Pauwels, L. J. R., Bruinsma, G. J. N., & Bernasco, W. (2013). Moral emotions and offending: Do feelings of anticipated shame and guilt mediate the effect of socialization on offending? *European Journal of Criminology, 10*(1), 22-39. doi:10.1177/1477370812454393

Svetlova, M., Nichols, S. R., & Brownell, C. A. (2010). Toddlers' prosocial behavior: from instrumental to empathic to altruistic helping. *Child Development, 81*(6), 1814-1827. doi:10.1111/j.1467-8624.2010.01512.x

Takemoto, H., Kawamoto, Y., & Furuichi, T. (2015). How did bonobos come to range south of the congo river? Reconsideration of the divergence of *Pan paniscus* from other *Pan* populations. *Evolutionary Anthropology, 24*(5), 170-184. doi:10.1002/evan.21456

Tan, J. Z., & Hare, B. (2013). Bonobos Share with Strangers. *PLoS ONE, 8*(1), e51922. doi:10.1371/journal.pone.0051922

Taylor, P. W. (1978). On Taking the Moral Point of View. *Midwest Studies in Philosophy, 3*(1), 35-61. doi:10.1111/j.1475-4975.1978.tb00347.x

Tennie, C., Call, J., & Tomasello, M. (2006). Push or Pull: Imitation vs. Emulation in Great Apes and Human Children. *Ethology, 12*(112), 1159-1169.

Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: on the evolution of cumulative culture. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1528), 2405-2415. doi:10.1098/rstb.2009.0052

Tennie, C., & Hedwig, D. (2010). How Latent Solution Experiments Can Help to Study Differences Between Human Culture and Primate Traditions. In E. Potocki & J. Krasiński (Eds.), *Primatology: Theories, Methods and Research* (pp. 95-112). New York, NY: Nova Publishers.

Tisak, M. S. (1995). Domains of social reasoning and beyond. *Annals of Child Development, 11*, 95-130.

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention: its origins and role in development* (pp. 103-130). Hillsdale, NJ: Lawrence Erlbaum Associates.

Tomasello, M. (1996). Do apes ape? In B. G. J. Galef (Ed.), *Social Learning in Animals: The Roots of Culture* (pp. 319-346). San Diego: Academic Press.

Tomasello, M. (1998). Reference: Intending that others jointly attend. *Pragmatics & Cognition, 6*(1), 229-243. doi:10.1075/pc.6.1-2.12tom

Tomasello, M. (2003). *Constructing a language: a usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Tomasello, M. (2005). Beyond formalities: The case of language acquisition. *Linguistic Review, 22*(2-4), 183-197. doi:10.1515/tlir.2005.22.2-4.183

Tomasello, M. (2008). *Origins of human communication*. Cambridge, MA: MIT Press.

Tomasello, M. (2009a). The question of chimpanzee culture, plus postscript (chimpanzee culture, 2009). In K. N. Laland & B. G. Galef (Eds.), *The Question of Animal Culture* (pp. 198-221). Cambridge, MA: Harvard University Press.

Tomasello, M. (2009b). *Why we cooperate*. Cambridge, MA: MIT Press.

Tomasello, M. (2014). *A natural history of human thinking*. Cambridge, MA: Harvard University Press.

Tomasello, M. (2015). *A natural history of human morality*. Cambridge, MA: Harvard University Press.

Tomasello, M., & Call, J. (1997). *Primate cognition*. Oxford: Oxford University Press.

Tomasello, M., Call, J., & Hare, B. (2003). Chimpanzees understand psychological states—the question is which ones and to what extent. *Trends in Cognitive Sciences, 7*(4), 153-156. doi:10.1016/S1364-6613(03)00035-4

Tomasello, M., & Carpenter, M. (2005). The emergence of social cognition in three young chimpanzees. *Monographs of the Society for Research in Child Development, 70*(1), vii-132. doi:10.1111/j.1540-5834.2005.00324.x

Tomasello, M., & Carpenter, M. (2007). Shared intentionality. *Developmental Science, 10*(1), 121-125. doi:10.1111/j.1467-7687.2007.00573.x

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28*(5), 675-691.

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Dev, 78*(3), 705-722. doi:10.1111/j.1467-8624.2007.01025.x

Tomasello, M., & Gonzalez-Cabrera, I. (2017). The Role of Ontogeny in the Evolution of Human Cooperation. *Human Nature*, 1-15. doi:10.1007/s12110-017-9291-1

Tomasello, M., & Gonzalez-Cabrera, I. (forthcoming). The role of ontogeny in the evolution of human cooperation. *Human Nature.*

Tomasello, M., & Hamann, K. (2012). Collaboration in young children. *Q J Exp Psychol (Hove), 65*(1), 1-12. doi:10.1080/17470218.2011.608853

Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis. *Journal of Human Evolution, 52*(3), 314-320. doi:10.1016/j.jhevo1.2006.10.001

Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural Learning. *Behavioral and Brain Sciences, 16*(3), 495-511.

Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., & Herrmann, E. (2012). Two Key Steps in the Evolution of Human Cooperation: The Interdependence Hypothesis. *Current Anthropology, 53*(6), 673-692. doi:10.1086/668207

Tomasello, M., & Vaish, A. (2013). Origins of Human Cooperation and Morality. *Annual Review of Psychology, 64,* 231-255. doi:10.1146/annurev-psych-113011-143812

Tomonaga, M., Tanaka, M., Matsuzawa, T., Myowa-yamakoshi, M., Kosugi, D., Mizuno, Y., . . . Bard, K. A. (2004). Development of social cognition in infant chimpanzees (*Pan troglodytes*): Face recognition, smiling, gaze, and the lack of triadic interactions. *Japanese Psychological Research, Volume 46*(Issue 3), 227-235. doi:10.1111/j.1468-5584.2004.00254.x

Tooby, J., & DeVore, I. (1987). The reconstruction of hominid behavioral evolution through strategic modeling. In W. G. Kinzey (Ed.), *The Evolution of human behavior: primate models* (pp. 183-238). Albany: State University of New York Press.

Trevarthen, C. (1979). Instincts for human understanding and for cultural cooperation: Their development in infancy. In M. von Cranach, K. Foppa, W. Lepenies, & D. Ploog (Eds.), *Human Ethology: Claims and Limits of a New Discipline* (pp. 530-571). Cambridge: Cambridge University Press.

Tronick, E. Z., Als, H., Adamson, L., Wise, S., & Brazelton, T. B. (1978). The Infant’s Response to Entrapment between Contradictory Messages in Face-to-Face Interaction. *Journal of the American Academy of Child Psychiatry, 17*(1), 1-13. doi:10.1016/S0002-7138(09)62273-1

Tuomela, R. (2003). The We-Mode and the I-Mode. In F. F. Schmitt (Ed.), *Socializing Metaphysics: The Nature of Social Reality* (pp. 93-128). Lanham, Md: Rowman and Littlefield.

Turiel, E. (1978). Distinct conceptual and developmental domains: Social convention and morality. In H. E. Howe & C. B. Keasey (Eds.), *Nebraska Symposium on Motivation, 1977: Social Cognitive Development* (Vol. 25, pp. 77-116). Lincon, NE: University of Nebraska Press.

Turiel, E. (1983). *The development of social knowledge: morality and convention*. Cambridge: Cambridge University Press.

Turiel, E. (1998). The development of morality. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (5th ed. ed., Vol. 3, pp. 863-932). New York, NY: Wiley.

Ueno, A., & Matsuzawa, T. (2004). Food transfer between chimpanzee mothers and their infants. *Primates, 45*(4), 231-239. doi:10.1007/s10329-004-0085-9

Ungar, P. S., Grine, F. E., Teaford, M. F., & El Zaatari, S. (2006). Dental microwear and diets of African early *Homo*. *Journal of Human Evolution, 50*(1), 78-95. doi:10.1016/j.jhevol.2005.08.007

Ungar, P. S., Krueger, K. L., Blumenschine, R. J., Njau, J., & Scott, R. S. (2012). Dental microwear texture analysis of hominins recovered by the Olduvai Landscape Paleoanthropology Project, 1995-2007. *Journal of Human Evolution, 63*(2), 429-437. doi:10.1016/j.jhevol.2011.04.006

Ungar, P. S., Scott, R. S., Grine, F. E., & Teaford, M. F. (2010). Molar microwear textures and the diets of *Australopithecus anamensis* and *Australopithecus afarensis*. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365*(1556), 3345-3354. doi:10.1098/rstb.2010.0033

Vaish, A., Carpenter, M., & Tomasello, M. (2009). Sympathy Through Affective Perspective Taking and Its Relation to Prosocial Behavior in Toddlers. *Developmental Psychology, 45*(2), 534-543. doi:10.1037/A0014322

Vaish, A., Carpenter, M., & Tomasello, M. (2010). Young children selectively avoid helping people with harmful intentions. *Child Dev., 81*, 1661.

Vaish, A., Carpenter, M., & Tomasello, M. (2011). Young children’s responses to guilt displays. *Dev Psychol, 47*(5), 1248-1262. doi:10.1037/a0024462

Vaish, A., Carpenter, M., & Tomasello, M. (2016). The Early Emergence of Guilt-Motivated Prosocial Behavior. *Child Development, 87*(6), 1772-1782. doi:10.1111/cdev.12628

Vaish, A., Missana, M., & Tomasello, M. (2011). Three-year-old children intervene in third-party moral transgressions. *Br. J. Dev. Psychol., 29*, 124.

Vaish, A., & Tomasello, M. (2013). The Early Ontogeny of Human Cooperation and Morality. In M. Killen & J. G. Smetana (Eds.), *Handbook of Moral Development* (Second Edition ed.): Taylor & Francis.

van Schaik, C. P., Fox, E. A., & Sitompul, A. F. (1996). Manufacture and use of tools in wild Sumatran orangutans. *Naturwissenschaften, 83*(4), 186-188. doi:10.1007/BF01143062

Vervaecke, H., de Vries, H., & van Elsacker, L. (2000). Dominance and its behavioral measures in a captive group of bonobos (*Pan paniscus*). *International Journal of Primatology, 21*(1), 47-68. doi:10.1023/A:1005471512788

Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks, 23*(8-9), 998-1003. doi:10.1016/j.neunet.2010.06.002

Voland, E., & Jan, B. (2002). Opposite Effects of Maternal and Paternal Grandmothers on Infant Survival in Historical Krummhörn. *Behavioral Ecology and Sociobiology, 52*(6), 435-443.

von Rohr, C. R., Burkart, J. M., & van Schaik, C. P. (2011). Evolutionary precursors of social norms in chimpanzees: a new approach. *Philosophy and Biology, 26*(1), 1-30. doi:10.1007/s10539-010-9240-4

Waddington, C. H. (1953). Genetic Assimilation of an Acquired Character. *Evolution, 7*(2), 118-126. doi:10.2307/2405747

Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development, 77*(3), 640-663. doi:10.1111/j.1467-8624.2006.00895.x

Warneken, F., Gräfenhain, M., & Tomasello, M. (2012). Collaborative partner or social tool? New evidence for young children's understanding of joint intentions in collaborative activities. *Developmental Science, 15*(1), 54-61. doi:10.1111/j.1467-7687.2011.01107.x

Warneken, F., Hare, B., Melis, A. P., Hanus, D., & Tomasello, M. (2007). Spontaneous altruism by chimpanzees and young children. *PLoS Biology, 5*(7), e184. doi:07-PLBI-RA-0267 [pii]

10.1371/journal.pbio.0050184

Warneken, F., Lohse, K., Melis, A. P., & Tomasello, M. (2011). Young children share the spoils after collaboration. *Psychol. Sci., 22*, 267.

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science, 311*(5765), 1301-1303. doi:10.1126/science.1121448

Warneken, F., & Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy, 11*(3), 271-294.

Warneken, F., & Tomasello, M. (2013). The emergence of contingent reciprocity in young children. *Journal of Experimental Child Psychology, 116*(2), 338-350. doi:10.1016/j.jecp.2013.06.002

Washburn, D. A., Gulledge, J. P., Beran, M. J., & Smith, J. D. (2010). With his memory magnetically erased, a monkey knows he is uncertain. *Biology Letters, 6*(2), 160-162. doi:10.1098/rsbl.2009.0737

Watts, D. P. (2012). Long-term research on chimpanzee behavioural ecology in Kibale National Park, Uganda. In P. M. Kappeler & D. P. Watts (Eds.), *Long-term field studies of primates* (pp. 313-338). Berlin: Springer.

Wellman, H. M., Lane, J. D., LaBounty, J., & Olson, S. L. (2011). Observant, nonaggressive temperament predicts theory-of-mind development. *Developmental Science, 14*(2), 319-326. doi:10.1111/j.1467-7687.2010.00977.x

West-Eberhard, M. J. (1979). Sexual Selection, Social Competition, and Evolution. *Proceedings of the American Philosophical Society, 123*(4), 222-234.

West-Eberhard, M. J. (2003). *Developmental plasticity and evolution*. Oxford: Oxford University Press.

Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science, 16*(10), 780-784. doi:10.1111/j.1467-9280.2005.01614.x

White, T. D., Ambrose, S. H., Suwa, G., Su, D. F., DeGusta, D., Bernor, R. L., . . . Vrba, E. (2009). Macrovertebrate paleontology and the Pliocene habitat of *Ardipithecus ramidus*. *Science, 326*(5949), 87-93.

White, T. D., Asfaw, B., Beyene, Y., Haile-Selassie, Y., Lovejoy, C. O., Suwa, G., & WoldeGabriel, G. (2009). *Ardipithecus ramidus* and the Paleobiology of Early Hominids. *Science, 326*(5949), 64, 75-86. doi:10.1126/science.1175802

White, T. D., Lovejoy, C. O., Asfaw, B., Carlson, J. P., & Suwa, G. (2015). Neither chimpanzee nor human, *Ardipithecus* reveals the surprising ancestry of both. *Proceedings of the National Academy of Sciences of the United States of America, 112*(16), 4877-4884. doi:10.1073/pnas.1403659111

White, T. D., Suwa, G., & Asfaw, B. (1994). *Australopithecus ramidus*, a new species of early hominid from Aramis, Ethiopia. *Nature, 371*(6495), 306-312. doi:10.1038/371306a0

White, T. D., WoldeGabriel, G., Asfaw, B., Ambrose, S., Beyene, Y., Bernor, R. L., . . . Suwa, G. (2006). Asa Issie, aramis and the origin of *Australopithecus*. *Nature, 440*(7086), 883-889. doi:10.1038/nature04629

Whiten, A., Byrne, R. W., Barton, R. A., Waterman, P. G., & Henzi, S. P. (1991). Dietary and Foraging Strategies of Baboons. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 334*(1270), 187-197. doi:10.1098/rstb.1991.0108

Whiten, A., Byrne, R. W., & Henzi, S. P. (1987). The Behavioral Ecology of Mountain Baboons. *International Journal of Primatology, 8*(4), 367-388. doi:10.1007/Bf02737389

Whiten, A., McGrew, W. C., Aiello, L. C., Boesch, C., Boyd, R., Byrne, R. W., . . . Wrangham, R. W. (2010). Studying Extant Species to Model Our Past. *Science, 327*(5964), 410-410.

Whiten, A., McGuigan, N., Marshall-Pescini, S., & Hopper, L. M. (2009). Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1528), 2417-2428. doi:10.1098/rstb.2009.0069

Wilkins, A. S., Wrangham, R. W., & Fitch, W. T. (2014). The "Domestication Syndrome" in Mammals: A Unified Explanation Based on Neural Crest Cell Behavior and Genetics. *Genetics, 197*(3), 795-808.

Williams, B. (1985). *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press.

Williams, F. L., Godfrey, L. R., & Sutherland, M. R. (2002). Heterochrony and the evolution of Neanderthal and modern human craniofacial form. In N. Minugh-Purvis & K. McNamara (Eds.), *Human evolution through developmental change* (pp. xviii, 508 p.). Baltimore: Johns Hopkins University Press.

Williams, F. L., Godfrey, L. R., & Sutherland, M. R. (2003). Diagnosing heterochronic perturbations in the craniofacial evolution of *Homo* (Neandertals and modern humans) and *Pan* (*P. troglodytes* and *P. paniscus*). In J. L. Thompson, G. E. Krovitz, & A. J. Nelson (Eds.), *Patterns of growth and development in the genus Homo* (pp. 295-319). Cambridge: Cambridge University Press.

Wimsatt, W. C. (1999). Generativity, entrenchment, evolution, and innateness: philosophy, evolutionary biology, and conceptual foundations of science. In V. Harcastle (Ed.), *Where Biology Meets Psychology* (pp. 137-179).

Wobber, V., Hare, B., Maboto, J., Lipson, S., Wrangham, R. W., & Ellison, P. T. (2010). Differential changes in steroid hormones before competition in bonobos and chimpanzees. *Proceedings of the National Academy of Sciences, 107*(28), 12457-12462. doi:10.1073/pnas.1007411107

Wobber, V., Herrmann, E., Hare, B., Wrangham, R. W., & Tomasello, M. (2014). Differences in the early cognitive development of children and great apes. *Dev Psychobiol, 56*(3), 547-573. doi:10.1002/dev.21125

Wolpoff, M. H. (1976). Some Aspects of the Evolution of Early Hominid Sexual Dimorphism. *Current Anthropology, 17*(4), 579-606.

Wrangham, R. W. (1987). The significance of African apes for reconstructing human social evolution. In W. G. Kinzey (Ed.), *The Evolution of human behavior: primate models* (pp. 51-71). Albany: State University of New York Press.

Wrangham, R. W. (1993). The evolution of sexuality in chimpanzees and bonobos. *Human Nature, 4*(1), 47-79. doi:10.1007/BF02734089

Wrangham, R. W. (2009). *Catching fire: how cooking made us human*. London: Profile.

Wrangham, R. W. (2011). Chimpanzees, Bonobos and the Self-Domestication Hypothesis. *American Journal of Primatology, 73*, 33-33.

Wrangham, R. W., & Glowacki, L. (2012). Intergroup Aggression in Chimpanzees and War in Nomadic Hunter-Gatherers. *Human Nature, 23*(1), 5-29. doi:10.1007/s12110-012-9132-1

Wrangham, R. W., & Peterson, D. (1996). *Demonic males: apes and the origins of human violence*. Boston: Houghton Mifflin.

Wrangham, R. W., & Pilbeam, D. (2001). African apes as time machines. In B. Galdikas, N. Briggs, L. Sheeran, G. Shapiro, & J. Goodall (Eds.), *All apes great and small* (Vol. 1, African apes, pp. 5-17). Dordrecht: Kluwer Academic.

Wyman, E., Rakoczy, H., & Tomasello, M. (2009a). Normativity and context in young children’s pretend play. *Cognitive Development, 24*(2), 146-155. doi:10.1016/j.cogdev.2009.01.003

Wyman, E., Rakoczy, H., & Tomasello, M. (2009b). Young children understand multiple pretend identities in their object play. *British Journal of Developmental Psychology, 27*, 385-404. doi:10.1348/026151008x322893

Yamamoto, S. (2015). Non-reciprocal but peaceful fruit sharing in wild bonobos in Wamba. *Behaviour, 152*(3-4), 335-357. doi:10.1163/1568539x-00003257

Yovel, G., & Freiwald, W. A. (2013). Face recognition systems in monkey and human: are they the same thing? *F1000Prime Reports, 5*, 10. doi:10.12703/P5-10

Zahn-Waxler, C., Radkeyarrow, M., Wagner, E., & Chapman, M. (1992). Development of Concern for Others. *Developmental Psychology, 28*(1), 126-136. doi:10.1037/0012-1649.28.1.126

Zelditch, M. L., Sheets, H. D., & Fink, W. L. (2000). Spatiotemporal reorganization of growth rates in the evolution of ontogeny. *Evolution, 54*(4), 1363-1371.

Zihlman, A. L. (1984). Pygmy Chimps, People, and the Pundits. *New Scientist, 104*(1430), 39-40.

Zihlman, A. L., & Bolter, D. R. (2015). Body composition in Pan paniscus compared with Homo sapiens has implications for changes during human evolution.

Proceedings of the National Academy of Sciences. doi:10.1073/pnas.1505071112

Zihlman, A. L., Cronin, J. E., Cramer, D. L., & Sarich, V. M. (1978). Pygmy Chimpanzee as a Possible Prototype for Common Ancestor of Humans, Chimpanzees and Gorillas. *Nature, 275*(5682), 744-746. doi:10.1038/275744a0

Zihlman, A. L., & Lowenstein, J. M. (1983). *Ramapithecus* and *Pan paniscus*: significance for human origins. In R. L. Ciochon & R. S. Corruccini (Eds.), *New interpretations of ape and human ancestry* (pp. 677-694). New York, NY: Plenum Press.