

2019-07-29

Toward a Theory of Ethical Mindedness: Moral Psychology and the Evolution of Human Nature

Jordan B. Ainsley

University of Miami, jordan.ainsley@gmail.com

Follow this and additional works at: https://scholarlyrepository.miami.edu/oa_dissertations

Recommended Citation

Ainsley, Jordan B., "Toward a Theory of Ethical Mindedness: Moral Psychology and the Evolution of Human Nature" (2019). *Open Access Dissertations*. 2343.

https://scholarlyrepository.miami.edu/oa_dissertations/2343

This Open access is brought to you for free and open access by the Electronic Theses and Dissertations at Scholarly Repository. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of Scholarly Repository. For more information, please contact repository.library@miami.edu.

UNIVERSITY OF MIAMI

TOWARD A THEORY OF ETHICAL MINDEDNESS:
MORAL PSYCHOLOGY AND THE EVOLUTION OF HUMAN NATURE

By

Jordan B. Ainsley

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

August 2019

©2019
Jordan B. Ainsley
All Rights Reserved

UNIVERSITY OF MIAMI

A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

TOWARD A THEORY OF ETHICAL MINDEDNESS:
MORAL PSYCHOLOGY AND THE EVOLUTION OF HUMAN NATURE

Jordan B. Ainsley

Approved:

Blaine Fowers, Ph.D.
Professor of Educational and
Psychological Studies

Soyeon Ahn, Ph.D.
Professor of Educational and
Psychological Studies

Isaac Prilleltensky, Ph.D.
Vice Provost for Institutional Culture
Professor of Educational and
Psychological Studies

Guillermo Prado, Ph.D..
Dean of the Graduate School

Brad Cokelet, Ph.D.
Assistant Professor of Philosophy
University of Kansas

AINSLEY, JORDAN B.

(Ph.D., Counseling Psychology)

Toward a Theory of Ethical Mindedness:
Moral Psychology and the Evolution of Human Nature.

(August 2019)

Abstract of a dissertation at the University of Miami.

Dissertation supervised by Professor Blaine Fowers.

No. of pages in text. (151)

The current dissertation is an examination of research and theoretical frameworks for explaining ethical mindedness, or the human propensity to be concerned with morality and ethics. Despite a pervasive belief that morality is central to what it means to be human, the idea of ethical mindedness as a basic feature of human nature is often assumed but not fully explained by the extant research and theory. Given evidence that such a shared quality exists, it is concerning that we appear to lack of a cogent, systematic explanation for a basic feature of human nature. The research questions guiding this dissertation take a twofold approach: First, is ethical mindedness a demonstrable characteristic of human nature? Second, what might an adequate theory of ethical mindedness look like?

I approach these ontological and theoretical questions through the lens of evolutionary science. In the first chapter, I will review the evolutionary criteria for identifying adaptations and thus evolved features of human nature. In the second chapter, I will review how the current theoretical camps in moral psychology approach the question of ethical mindedness. This chapter will demonstrate that despite making significant contributions to larger theories of moral functioning, moral psychologists

appear to lack a full understanding of ethical mindedness and how it is that humans came to be this way.

I will begin building that understanding in the third chapter by applying the evolutionary criteria to cooperation as a paradigm account of their use and how a theory of ethical mindedness can be built. In the fourth chapter, I will continue to utilize evolutionary reasoning to trace how cooperation may have spread throughout the species and the subsequent emergence of ethical mindedness. Through this examination, I argue that the ultrasociality of human beings suggests a communal nature to cooperation that leads, through cheater detection, indirect reciprocity, and third party punishment, to the emergence of a fairness instinct, or an attunement to the fairness of the group. This basic communal nature and concern for the fairness of our groups is what I suggest leads to ethical mindedness as a basic feature of human nature.

DEDICATIONS

To Blaine Fowers, for his endless patience, mentorship, and wisdom.

To Tyler Lefevor and Austen Anderson for the conceptual and organizational contributions that brought this project into existence.

To Brady Yocom, Charlene Leak, Marina Donovan, and Shelley Payne. I have enjoyed my upsets and wrong turns the most with you.

To my parents, Ann and Steve, and my sister Caitlin. Supporting me to this point has been an act of heroism that will not likely be repaid by any effort to make that easier in the future. You are remarkable and I am fortunate beyond words.

And to my beautiful wife Lisa and our growing family. You have given my life meaning and purpose.

TABLE OF CONTENTS

Chapter	Page
INTRODUCTION	1
1 EVOLUTIONARY PSYCHOLOGY	9
Understanding Evolutionary Psychology	9
Evolutionary Psychology and Understanding Human Nature	13
Identifying Adaptations	18
Criteria for identifying adaptations	19
Fitness benefits.....	19
Lock-and-key fit between adaptive problem and adaptation.....	21
Cannot be explained by any other adaptation	23
Presence/absence in phylogenetic relatives	24
Develops naturally and without explicit learning	26
Ubiquitous across cultures	29
Neural circuits.....	31
Concluding Points.....	32
2 MORAL PSYCHOLOGY	33
Social Cognitive Domain.....	34
Summary	43
Moral Identity Perspective.....	45
Summary	54
Moral Foundations Theory	55
Summary	65
Concluding Points.....	66
3 A PARADIGM ACCOUNT OF COOPERATION AS AN ELEMENT OF ETHICAL MINDEDNESS	70
Cooperation.....	72
Fitness benefits.....	73
Lock-and-key fit between adaptive problem and adaptation.....	80
Cannot be explained by any other adaptation	82
Presence/absence in phylogenetic relatives	84
Develops naturally and without explicit learning	87
Ubiquitous across cultures	93
Neural circuits.....	96
Concluding Points.....	99
4 THE CONNECTION TO ETHICAL MINDEDNESS AND A FINAL NOTE ON MORAL PSYCHOLOGY	102
Cooperation, Trust, and the Connection to Ethical Mindedness	102
Cheater detection	104
Indirect reciprocity.....	110

Punishment and third party punishment	112
Communal fairness and the connection to ethical mindedness	118
Communal Fairness and Moral Psychology	123
Social Cognitive Domain Theory	123
Moral Identity	125
Moral Foundations Theory	127
A final note on what is missing.....	129
CONCLUSION AND DIRECTIONS FOR FUTURE RESEARCH.....	133
REFERENCES	140

INTRODUCTION

The current dissertation is an examination of research and theoretical frameworks for explaining ethical mindedness, or the human propensity to be concerned with morality and ethics. To quote Darwin (1871), “I fully subscribe to the judgment that all of the differences between man and the lower animals, the moral sense of conscience is by far the most important” (p. 40). This belief that moral sense separates humans from other animals is reiterated by Lapsley (1996) in introducing his text on moral psychology with “Perhaps the study of moral conduct, ethical thinking, and values is so preeminent in the social sciences just because moral qualities appear to be what are so distinctive about the human species” (p. xiii). While there is ample evidence to suggest that some versions of moral functioning exists in other species, there is general agreement that the full range of morality and ethics among humans is a specifically and uniquely human feature. For example, although de Waal (1996, 2007) has argued that chimpanzees are capable of empathy and contends that other primates demonstrate “proto-moral” features; neither de Waal, nor any other scholar to my knowledge, argues that these features represent the specific type of moral and ethical functioning in human beings.

It is curious then that despite this pervasive belief that such morality is central to what it means to be human, the idea of ethical mindedness as a basic feature of human nature is, as we see in Lapsley’s statement, assumed but not fully explained by the extant subdiscipline of moral psychology. Because this failure to account for the origin of what Darwin suggested is a species-typical characteristic suggests a limitation in our conceptualization of human nature, it has the potential to impact research and theory across the social sciences, especially the field of moral psychology. Without an

understanding of ethical mindedness as a core feature of human nature, and thus moral functioning as emergent from that basic ethical nature, it would be easy to assume that human nature is fundamentally selfish. This unquestioned assumption is an easy one to make in the absence of a more widely accepted view of our basic ethical mindedness and moral nature as a context for even seemingly conflicting behaviors and history.

Moral functioning has been studied by developmental, social, evolutionary, cognitive and other psychologists. This diversity of perspectives has led to an increasingly decentralized literature (Lapsley, 1996). Nevertheless, it seems reasonable to look to moral psychology as a subdiscipline that should be able to explain why ethical mindedness appears to be a feature of human nature. It does not appear that this field has yet produced a fully systematic, cogent, and compelling explanation, however. Because a feature such as ethical mindedness is so potentially important to the field of moral psychology, my primary purposes in this dissertation are to identify the necessary criteria for producing such an explanation, analyze how ethical mindedness is addressed and explained in the contemporary moral psychological literature, and begin the process of documenting ethical mindedness as a species characteristic feature according to the identified criteria.

So what, exactly, do I mean by *ethical mindedness*, and why does it matter? At its core, I use the term *ethical mindedness* to refer to the “strong, rapid, and automatic inclination to see one’s own and others’ actions as right or wrong, good or bad” (Fowers, 2015, p. 3). It is the seemingly species-typical investment in the basic questions of right and wrong, good and bad, and in the morality of our own and others’ actions and choices. Ethical mindedness is the sense of being compelled to evaluate ethics and morality in the

first place, but it does not prescribe a particular content or conclusion for those evaluations. Ethical mindedness, similar to other human instincts, captures the set of instincts *to* learn an ethical viewpoint, but not what is taught or ultimately socialized upon those instincts (Cosmides & Tooby, 1997). This is an important distinction because in order to focus on human nature, we must recognize the richness of difference and diversity of ethical understandings that has emerged among human groups (Fowers, 2015). While this richness and diversity is itself a valid and compelling area of study, the current focus is framed as ethical mindedness to highlight primarily what is shared across the human species, not what differs within it.

Given evidence that such a shared quality exists, it is the lack of a cogent, systematic explanation for a basic feature of human nature that is so immediately concerning. Without the inclusion of ethical mindedness in our existing theoretical frameworks and a robust explanation for it, we cannot guide research in a direction that is consistent with the nature of our species. If it is true that ethical mindedness is a part of our basic nature, then our research and theory should capture and respond to that reality. Beyond this is perhaps the even more critical and practical issue of moral education. That is, without an adequate understanding of our own inherited ethical mindedness, how can we develop effective moral education?

Given this importance, the research question guiding this dissertation takes a twofold approach: First, is ethical mindedness a demonstrable characteristic of human nature? Second, what might an adequate theory of ethical mindedness look like?

I will approach these questions first by presenting a chapter that briefly discusses the value of considering ethical mindedness through the lens of evolutionary science,

followed by a detailed review of what constitutes an adequate evolutionary argument for evolved psychological species characteristics (as laid out by evolutionary psychologists). This is intended to provide both a framework for demonstrating the limitations in how ethical mindedness is accounted for in the existing theories of moral psychology as well as a governing structure for the development of an adequate case for ethical mindedness as a species-typical characteristic of human nature.

Once oriented to both the rationale and requirements for a theory grounded in evolutionary psychology, the second chapter will present existing research and theoretical developments in the field of moral psychology. For the most part, moral psychology is divided into three primary subgroups (Gibbs, 2019). These groups are born primarily from the influence of Kohlberg and Piaget and the theoretical frameworks can be identified as Moral Foundations Theory, Social Cognitive Domain Theory, and the moral identity viewpoint (Gibbs, 2019). While this distinction between theoretical perspectives is a commonly held understanding, it is not one that is frequently discussed. I will use it in the current dissertation primarily because it is useful in providing a framework to systematically review the relevant literature. There have also been important recent developments. For example, Narvaez (2014) has expanded significantly on the theoretical frameworks in Moral Identity with her Triune Ethics Theory.

It is arguably within moral psychology that explanations for ethical mindedness as a basic feature of human nature should be the richest and most comprehensive, but researchers from each subgroup appear to presume ethical mindedness while offering only thin explanations for its ubiquity in human nature. Although the evolutionary explanations that moral psychologists have offered for ethical mindedness provide useful

starting points and are, to some degree, consistent with my position in this dissertation, they remain incomplete relative to the prevailing standards required for an adequate evolutionary argument. This second chapter will thus begin with a brief overview of the field, followed by a detailed analysis of how the existing theoretical perspectives currently conceptualize and account for ethical mindedness. I will focus specifically on the limits of current theories of moral psychology in terms of both inadequate evolutionary arguments and a related failure to account for ethical mindedness as a species characteristic feature of human nature.

The third chapter will begin the process of developing an adequate case for ethical mindedness by reviewing evidence and research across disciplines. This chapter will focus on cooperation as a prime example of an adaptation involved in ethical mindedness. West, El Mouden, and Gardner (2011) define cooperation as “a behavior which provides benefit to another individual (recipient) and which is selected for because of its beneficial effect on the recipient” (p. 235). Although research in anthropology, evolutionary psychology, developmental psychology, social psychology, and neuroscience indicates that ethical mindedness is species characteristic for humans, it appears that researchers across disciplines have not yet come to grips with the centrality of ethical mindedness for human beings. Accounting for ethical mindedness as a central feature of human nature may be made difficult by a general discomfort with addressing ethics and values directly in psychology. Many have dealt with this discomfort by imposing a fact-value dichotomy, believing that they can insist only on facts and retain an arm’s length distance from any personal or social ideals and values. In other words, while the evidence suggests that ethical mindedness is species characteristic, the producers of that evidence have

failed to expand their focus from what they take to be value-free observable facts to an interpretation that suggests a deeper, more fundamental nature. They worry that acknowledging the values inherent in concepts like ethical mindedness could open the way to “contaminating” science with values.

Many have questioned whether such a fact-value dichotomy can be maintained in the sciences (Brinkmann, 2009; Bruner, 1990; Cushman, 1990; Danziger, 1990; Dewey, 1926; Fowers, 2010, 2015; Richardson, Fowers, & Guignon, 1999). By demonstrating the connections from cooperation (an observable and demonstrable adaptation) to trust, trust to the instinct for fairness, and the fairness instinct to ethical mindedness, I will ultimately argue that there is a seamless evolutionary connection to ethical mindedness as an evolved characteristic. One important note here is to understand that when I discuss fairness, I am focusing on the instinct or interest in fairness and specifically on interpersonal fairness, not institutional or abstract conceptions of fairness. This is primarily because the basic focus of this dissertation is on understanding the fundamental and evolved features of human nature, or the basic instincts or propensities. While I believe that all humans possess an instinct or interest in fairness, what we come to understand as fair is vastly different and contextualized across cultures and time. To emphasize this distinction, I will often refer to a “fairness instinct” instead of simply making reference to the idea of fairness.

The core purpose of Chapter 3, however, will be to begin the larger argument for ethical mindedness by integrating evidence from across disciplines and demonstrating a paradigm account of cooperation. Although I will leave most of the work of applying the criteria for evolutionary argument to others, and ultimately a full range, of moral domains

to future publications, it is my belief that once an argument for ethical mindedness has been made for one domain or characteristic, it will not be difficult to establish that argument for other characteristics.

In the fourth chapter of this dissertation, the evolutionary argument for the centrality of ethical mindedness will be constructed following the demonstration of cooperation as an evolved feature of human nature. I will continue to assume an evolutionary perspective and track the adaptation of cooperation as it evolved through cheater detection, indirect reciprocation, and third party punishment. These adaptations provide insight into the role of trust and instinct for fairness, which produces a view of cooperation that is fundamentally communal and based in the ultrasocial nature of human beings. By following this line of reasoning, I argue that there is a seamless connection from cooperation as an observable, demonstrable adaptation and the emergence of ethical mindedness. This is intended to serve as one demonstration of how evolutionary reasoning leads to a theoretical understanding of human nature as constituted by a basic ethical mindedness.

Having demonstrated the connection between cooperation and ethical mindedness and the fundamental communality of fairness and moral functioning, I will then revisit the existing work of moral psychologists. In this discussion, I will focus primarily on illustrating the consequences of having neglected to explain moral functioning in terms of its evolutionary history. I will offer brief examples of what moral psychologists have said about fairness and justice and whether those claims line up with the reality of how moral functioning emerges based on evolutionary reasoning.

I will conclude this dissertation by discussing the overarching narrative, reiterating its importance. This conclusion is intended to have three lasting effects on both the specific field of moral psychology and the broader discipline as a whole. At the most immediate level, this argument is (firstly) intended to strengthen and refine the theoretical frameworks in moral psychology. By doing so, and by illustrating a core feature of human nature that moral psychologists are uniquely positioned to continue studying, this argument is (secondly) intended to encourage greater attention to the potential influence of an otherwise small subfield of psychology. This established, it is hoped that (thirdly) moral psychologists in particular, and psychologists in general, can begin to research and theorize from perspectives that account for the centrality of ethical mindedness to human nature.

CHAPTER 1: EVOLUTIONARY PSYCHOLOGY

Perhaps the first question that ought to be acknowledged before delving into the rationale and nuance of evolutionary argument is: Why evolutionary theory? The answer, at least in terms of the current dissertation, is relatively simple. The first and most straightforward reason for focusing on evolutionary theory is that the current work began with a study of moral psychology and a search for how the existing theoretical frameworks accounted for and explained the presence and ubiquity of human ethical mindedness. Moral psychologists, however inadequately, tend ultimately to reference evolution as their case for ethical mindedness as a species-typical characteristic. I therefore follow that perspective while deepening it toward greater theoretical adequacy. The second reason for this is well-stated by Fowers (2015) in the following:

The theory of evolution has been controversial from the outset, but it provides a systematic and scientific account of the origin and characteristics of life on earth that is entirely without peer... Evolutionary theory provides a uniquely detailed and fruitful way to understand human nature, and evolutionary psychology has developed to capitalize on the resources of evolutionary theory to provide a wide-ranging account of the human mind. (p. 43)

My purpose is not to make a case for evolutionary theory. That case has been well-documented elsewhere (e.g., Buss, 2005). My purpose in focusing on evolutionary psychology is to utilize the reigning scientific theory for understanding life on earth as an approach to deepening the theoretical frameworks of moral psychology and explaining ethical mindedness as species-typical feature of human nature.

Understanding Evolutionary Psychology

In Cosmides and Tooby's (1997) seminal primer on evolutionary psychology, the authors identify psychology as the "branch of biology that studies (1) brains, (2) how brains process information, and (3) how the brain's information-processing programs

generate behavior.” Within evolutionary psychology, there is a common distinction between brains and minds (Cosmides & Tooby, 1997). This distinction refers to the brain as a physical organ while the mind can be thought of as an emergent property (Bunge, 1977). In this sense, the mind is a capacity that emerges from the brain; it cannot happen without a brain, but a brain in and of itself is not enough to make it happen (Mainzer, 2008).

In other words, if we think of the brain as an information processing machine, in some ways similar to a computer, then the brain is the physical machine that one can touch and see. The emergent capacity of that machine, or the operation of information processing that it performs, is the mind. And it “learns” to perform that operation through programming and input. That programming, in this analogy, is the inherited characteristics and the social interaction and experiential input a brain receives are inputs that modify the programming. The social world interacts with the physical brain, and the mind emerges as the information-processing operation that occurs as a result (Cosmides & Tooby, 1992; Fowers, 2015).

It is important to understand this distinction because brain-mind discussions must be organized by distinct levels of analysis for both the physical organ and its emergent capacity. These separate levels are related but cannot be reduced to one or the other. This is critical to the current dissertation (and to understanding the perspective of evolutionary psychology) because one must be clear that the proposition is not reductive. Evolutionary psychologists see attempts to reduce analysis of the mind to an analysis of the brain at an organic, neurochemical, or cellular level as a serious category mistake.

Not only is this not a perspective of reductionism, but by identifying different levels of analysis, it becomes clear that each level must be treated with different tools, and perhaps even different theoretical frameworks if we are to begin seeing the truth of it. When we talk about brains, we can talk about fMRI machines, but an fMRI offers us no insight into the process, capacity, function, or experience of the human mind. To discuss the mind, we must focus on abstract qualities or capacities like identity or prospective thinking; properties possessed as universal features of human nature, but that cannot be seen through any observation of the brain alone. We must, in this sense, focus on the brain, the emergent mind, and the social context in which that mind emerges.

I clarify all of this because biology is essential, but ethical mindedness is a capacity of the mind that can't be reduced to biology; and the mind cannot be understood without understanding its emergence in a social context. We have brains that do a lot of the processing of information and formulation of responses, but all of that occurs in a social context. Ethical mindedness can't happen without a social context of deeply interdependent beings that goes beyond the biology and far beyond the brain.

While Cosmides and Tooby's (1997) description of psychology as a branch of biology may read as a dry or distant description to those who study the human behavior and experiences generated by the brain, this distilling of psychology down to a specific branch of biology creates a perspective from which the "inferential tools developed in biology – its theories, principles, and observations – can be used to understand psychology" (Cosmides & Tooby, 1997). The purpose of this connection of psychology to the brain, and therefore biology, is again not a reductive effort to simplify the field or

devalue existing research. On the contrary, it is an effort to create connection between empirical findings and theoretical frameworks, biology and psychology.

Cosmides and Tooby (1997) further clarify the importance of this connection to biology when they describe evolutionary psychology as “not an area of study, like vision, reasoning, or social behavior. It is a *way of thinking* about psychology that can be applied to any topic within it” (emphasis in original). In this sense, evolutionary psychology is a way of thinking that connects psychological research and theory to the study of the brain, how the brain functions, the behavior generated by that function, and the evolutionary process through which all of this emerged. One goal of this assimilation of research is to achieve conceptual integration. Cosmides and Tooby (1992) describe this integration as “the principle that various disciplines within the behavioral and social sciences should make themselves mutually consistent, and consistent with what is known in the natural sciences as well” (p. 4). This is critical when we consider it in light of the disparate, decentralized research that characterizes our understanding of moral functioning.

Wilson (1975) asserted that "scientists and humanists should consider together the possibility that the time has come for ethics to be removed temporarily from the hands of the philosophers and biologicized" (p. 562). His view is perhaps extreme, and I will not advocate for this sort of reductionism. Nonetheless Wilson points the current dissertation in a helpful direction. A conceptually integrated theory will be cohesive with the findings across relevant disciplines. Cosmides and Tooby (1992) use as their example of this cohesion that the laws of physics are compatible with the laws of chemistry, which are compatible with those of biology; as well as that a chemist would not propose theories that violate the principles of physics (p. 4). In the social sciences, by contrast, no such

integration seems to exist. Psychological research tends to be conducted in veritable silos. The principles of evolutionary psychology are designed to begin bridging those gaps to develop richer, more meaningful theoretical frameworks and findings. Without this connection to biology, we have no hope of achieving a unified theory of human nature, and we remain at risk of developing theories that conflict with observable biological facts.

I hope to contribute to those efforts both in my use of evolutionary psychology as the governing framework for this dissertation, as well as in my related inclusion of evidence for ethical mindedness that is drawn from a wide range of disciplines.

Evolutionary Psychology and Understanding Human Nature

So how does evolutionary psychology work, and how can we use it to understand human nature? Cosmides and Tooby (1997) describe the brain as an information-processing machine that was designed through natural selection to solve adaptive problems faced by our ancestors. This description includes three terms worthy of clear definitions. I used the analogy of a computer above when clarifying the distinction between mind and brain, and to introduce the influence of social experience on that relationship. That analogy is based on Cosmides and Tooby's (1997) description that the brain functions as an information processing machine. While perhaps more mechanistic than necessary, this term is most apt if we think of computers not as calculators, but as machines that are programmed (i.e. that "learn") to respond to conditional, if-then statements. In this sense, computers are rudimentary "reasoning" machines. A computation is ultimately a conditional statement that processes "if-this-then-that". Brains learn conditional input-output rules in a similar way. If I receive a signal that I am

being excluded or rejected, I feel pain; one that I am working as a team, I feel energized. My brain processes the input of my external environment and generates output. It is, in this sense, an information processing machine.

The second term “natural selection” is used to describe the process through which the human brain was designed. Natural selection refers to “a key process of evolution wherein naturally occurring genetic mutations are selected because they increase the organism’s reproductive success” (Fowers, 2015, p. 43). The key elements include naturally occurring genetic mutations and enhanced likelihood of reproduction that occurs as a result of those otherwise-random mutations. In keeping with one of the examples above (albeit while simplifying an otherwise complex process), suppose our ancestors’ brains were slightly more sensitive to social exclusion. If social exclusion causes them more pain, they would be more apt to avoid it. This avoidance would enhance their likelihood of avoiding group exclusion, thereby increasing their probability of survival, increasing their reproductive success (as well the reproductive success of their offspring should they share that otherwise-random mutation), thus spreading the mutation throughout the population in contrast to one that had no influence on reproduction. In that sense, the mutation was “naturally selected” because it increased reproductive success.

This leads us to the third term worthy of clarification: adaptive problems. Adaptive problems are related to both reproductive success and survival. With regard to reproductive success, we can think of problems such as mate selection or providing for offspring; and with regard to survival, we can think of problems such as finding food and avoiding injury or illness (Fowers, 2015). Longstanding challenges to reproduction and

survival of individuals in a given environment are referred to as adaptive problems because “such longstanding problems are solved through adaptations to the design of the species” (Fowers, 2015, p. 43). In other words, adaptive problems are solved through random mutations that change how an individual is built or functions and are naturally selected and spread throughout the population when the changes result in increased reproductive success.

So when Cosmides and Tooby (1997) describe the brain as an information-processing machine that was designed through natural selection to solve adaptive problems faced by our ancestors, what they refer to is the brain as a machine that has evolved over time as otherwise-random mutations were selected for their unique ability to solve the problems of a given environment and enhance reproductive success. Drawing from the principles of biology, evolutionary psychologists thus attempt to understand the human mind by understanding the features and characteristics of the brain. Darwin’s (1936) theory of evolution provides two evolutionary principles for explaining the characteristics of any animal: (1) common descent, and (2) adaptation driven by natural selection (Cosmides & Tooby, 1997). By common descent, the authors refer to the idea that species who share a common ancestor will share many of the same adaptations, all inherited from that ancestor. Species are considered closely related if they share a more recent common ancestor.

These two terms, common descent, and adaptation driven by natural selection, are critical to understanding how evolutionary psychologists attempt to understand the human brain based on principles of biology. There are two key approaches that emerge from the field of biology and are used by evolutionary psychologists. The first of these

approaches is the **phylogenetic approach**. Phylogenesis refers to the study of an organism's features, and this approach focuses on identifying shared and different features between different species within a shared lineage, such as primates. When two species inherit a shared trait from an ancestor species, it is known as a homologue (Maestriperi & Roney, 2006). Examples of such homologues include the evolution of wings in insects, which are “ineffective for flight below a certain size, but effective as solar heat exchange panels precisely within that range” (Pinker & Bloom, 1992, p. 459).

By tracing these features across species, we can develop an evolutionary timeline that explains when they emerged and how different species are related to one another (Cosmides & Tooby, 1997). And by understanding the timeline in which different adaptations emerged, we can also better understand the relationship between the environment and adaptive problems that were solved by the adaptations.

A second approach used by evolutionary psychologists is referred to as an **adaptationist approach**. This approach “leads the search for adaptive design, which usually entails the examination of niche-differentiated mental abilities unique to the species being investigated” (Cosmides & Tooby, 1997). This approach is based on the search for adaptations that are driven by natural selection. As discussed above, natural selection refers to the process in which specific genetic mutations are selected and spread throughout a population because they increase reproductive success. This approach thus focuses on identifying a specific adaptation and how it was designed to solve an adaptive problem. When we think about reproductive success, we consider that if one organism or a small set have an adaptation, then, on average, they will have more offspring that

themselves reproduce more until finally that adaptation dominates in the population over the course of hundreds or thousands of generations.

One example of this in humans is bipedal locomotion. Our ancestors were quadrupedal, and over generations we became bipedal because it offered us a number of advantages such as the abilities to carry things, see farther, or walk farther and faster. These abilities are important advantages for individuals on the savannah and not in the trees. Originally, some of our ancestors spent more time upright despite being designed for quadrupedal motion. Those advantages accrued and, over time, the species became bipedal.

Evolutionary biologists further structure their approach by distinguishing between two different types of explanations. Proximate explanations refer to immediate or structural processes while distal explanations refer to those that are ultimate, functional, or operate across time (Cosmides & Tooby, 1997). Applied to psychology, proximate explanations would focus on “genetic, biochemical, physiological, developmental, cognitive, social, and all other immediate causes of behavior” whereas distal or ultimate explanations refer to “causes that operated over evolutionary time” (Cosmides & Tooby, 1997). This is an important distinction in large part because the primary justification for an evolutionary framework is its ability to provide ultimate explanations.

One example of this distinction between proximate and ultimate explanations is humans’ preference for sweet and fatty foods. Proximate explanations are specific and important, but they don’t tell the full story. Feeling hungry and taking pleasure in sweet and fatty foods are sufficient proximate explanations for common eating behaviors, but hunger and pleasure do not provide us with an explanation for how that impulse and

preference emerged. An ultimate explanation for those behaviors would tell us that our impulse to eat and the pleasure we take in sweet and fatty foods are inherited. We get those traits from ancestors whose survival and reproduction were directly tied to their ability to focus on finding the sweet and fatty foods necessary to be strong and healthy. Those who were more driven toward sweet or fatty foods were more likely to survive and more likely to reproduce. Over time, such preferences became species typical, which is why we find ourselves drawn to the sweet and fatty foods despite knowledge that such impulses have problematic consequences to our health in a world of abundance.

Such ultimate explanations are the goal of research in evolutionary psychology and will guide my effort to understand how human beings came to be an ethically minded species. The true value of an evolutionary approach lies in its ability to provide us with ultimate explanations. That opportunity is thus, in large part, the justification for focusing this dissertation on developing an evolutionary understanding of ethical mindedness; because only with such ultimate explanations can we begin to develop a more complete and robust theory of human nature.

Identifying Adaptations

The following sections will outline the seven necessary criteria for identifying species-typical features using phylogenetic and adaptationist approaches. These criteria are designed to rigorously test the hypothesis that something is an adapted feature. They are a critical and necessary framework for an evolutionary perspective because without such criteria there is little means of distinguishing between a retrospective, speculative narrative and the accurate identification of specific trait that can be tested against sound and well-developed evolutionary theory.

One important note about these criteria is that although I have attempted to keep them distinct for the sake of clarity, there is significant overlap between them in many ways. This will become clearer as I apply the criteria in Chapter 3, though it warrants mentioning now as I use examples that may appear to fit within multiple criteria.

Criteria for identifying adaptations. These criteria will ultimately shape how I approach the assessment of existing moral psychology theories in this dissertation and how I will begin making a case for ethical mindedness as an evolved, species-typical feature of human nature. It is important to note here that the criteria for identifying adaptations and thus building evolutionary arguments are based on the work of Cosmides and Tooby (1992, 1997). While there is debate among evolutionary psychologists that focuses on what might qualify as fitting the criteria included in these authors' perspective, there has not been a systematic or direct critique of the importance and use of these criteria, which are relatively well-established.

The criteria for identifying adaptations were developed in response to critiques that evolutionary psychology is ultimately the production of "just-so stories". According to Merriam-Webster, a just-so story is defined as "a speculative story or explanation of doubtful or unprovable validity that is put forward to account for the origin of something (such as a biological trait) when no verifiable explanation is known." Cosmides and Tooby (1997) generated the criteria below in response to this concern, which is indeed a significant one. The criteria are thus designed to assess the validity for assertions of an adapted or evolved features.

Fitness benefits. Citing Hamilton (1964) and Williams and Williams (1957), Tooby and Cosmides (1992) explain that, "Because design features are embodied in

organisms, they can, generally speaking, propagate themselves in only two ways: (1) by solving problems that will increase the probability that the organism they are situated in will produce offspring, or (2) by solving problems that will increase the probability that the organism's kin will produce offspring" (p. 53). The authors use "design features" to refer to an organism's features that have been selected (or discarded) due to their impact on the individual's survival or reproduction (Tooby & Cosmides, 1997). Each individual and species is situated in an environment that presents specific, longstanding challenges to reproduction. A fitness benefit is provided by traits that give the individual an advantage with respect to one of these longstanding challenges or that enhance their reproductive opportunities and success.

This improved reproductive success due to a specific trait is referred to as a fitness benefit. A trait becomes species-typical when, over the course of thousands of generations, individuals who possess the trait enjoy greater reproductive success, which proliferates the trait throughout the population until the trait spreads through the species.

It is important to clarify here that fitness benefits do not refer to an individual organism that maximizes reproductive success as an intentional goal (Tooby & Cosmides, 1992). Instead, fitness benefits refer to an evolutionary process occurring across generations.

Fowers (2015) expands upon this and distinguishes between the relevant proximate or immediate goals and the ultimate explanation in describing that, "Many activities and even goals that individual organisms pursue will tend to enhance their fitness through proximate goals such as staying warm or maintaining contact with conspecifics." The ultimate explanation for pursuing these proximate goals is that

“Individuals pursue these goals because their evolutionary heritage has given them the capacities and inclinations to do so, which does enhance their fitness without any need for the organism to recognize that relationship” (p. 51). In other words, the ultimate explanation for the adaptations that contemporary humans have is that those features contributed to the reproductive success of our ancestors.

Lock-and-key fit between adaptive problem and adaptation. Every environment presents the species in it with specific challenges to survival and reproduction. Adaptive traits provide an advantage by facilitating survival and reproduction in the face of a longstanding problem. Because these problems are solved through changes to the physiological design of a species, the problems are referred to as adaptive problems and are “generally closely related to survival (finding food, avoiding injury or illness) and reproduction (mate selection, providing for offspring)” (Fowers, 2015, p. 43). Such problems must also be longstanding as an adaptive trait can only be selected and spread throughout a population over the course of a long history of survival and reproduction.

A compelling case for evolved traits requires that an adaptation be both identified and linked to the specific adaptive problem it was selected to solve. Cosmides and Tooby (1997) explain the following:

Adaptive problems have two defining characteristics. First, they are ones that cropped up again and again during the evolutionary history of a species. Second, they are problems whose solution affected the reproduction of individual organisms – however indirect the causal chain may be, and however small the effect on number of offspring produced. This is because differential reproduction (and not survival per se) is the engine that drives natural selection.

The identification of adaptations ought to also be considered in terms of what Cosmides and Tooby (1997) refer to as design evidence. That is, adaptations can be understood using a process similar to how one might attempt to understand a human-

made machine. In this, we may look at the specific and coordinated design features in order to understand a machine's function. The authors use as their example the process of identifying a machine as a TV rather than a stove, and explain that one can observe “many coordinated design features (antennae, cathode ray tubes, etc.) that are complexly specialized for transducing TV waves and transforming them into a color bit map (a configuration unlikely to have arisen by chance alone)” but very few features that appear functional for cooking food (Cosmides & Tooby, 1997). Cosmides and Tooby (1997) apply this to the identification of adaptations by explaining that, in order to do so, one must demonstrate the following:

(1) it has many design features that are complexly specialized for solving an adaptive problem, (2) these phenotypic properties are unlikely to have arisen by chance alone, and (3) they are not better explained as the by-product of mechanisms designed to solve some alternative adaptive problem.

Phenotypic refers to an observable, physical property that the authors explain “solves an adaptive problem with ‘reliability, efficiency, and economy’” – they identify the discovery of such properties as “prima facie evidence that one has located an adaptation” (Cosmides & Tooby, 1997).

Cosmides and Tooby (1997) offer that natural selection is a “feedback process that ‘chooses’ among alternative designs on the basis of how well they function. It is a hill-climbing process, in which a design feature that solves an adaptive problem well can be outcompeted by a new design feature that solves it better.” Because this process selects designs based on their effectiveness at solving an adaptive problem, it naturally creates a tight and specific fit between the selected design and the problem it solves. This is referred to as the lock-and-key fit between an adaptive problem and adaptation.

Perhaps most critical to understanding both reproductive success and the lock-and-key fit of adaptations is thus the importance of recalling that “fitness maximization is an evolutionary process that occurs over thousands of generations rather than a goal that individuals pursue” (Fowers, 2015, p. 51). This not only contextualizes the process of natural selection, but it highlights the fact that when we talk about evolution, we are talking about a process that unfolded in an entirely different context and environment. This environment, often referred to as the environment of evolutionary adaptedness (EEA), identifies the roughly 10 million years in which our ancestors lived a seminomadic existence, hunting game and foraging plants in groups of approximately 25 individuals (Fowers, 2015). Despite our radically different existence today, our basic nature is defined by the adaptations that evolved for survival millions of years ago. This is commonly referred to as the problem of “having a stone age brain in a modern world”.

The EEA is thus not only the context in which a lock-and-key adaptation must be located and explained, but it is the primary focus when evolutionary psychologists and biologists begin explaining what it means to be human. In other words, Fowers (2015) explains the following:

When evolutionists set out to identify what it is to be human, in contrast to our hominin forbearers, and, more distantly, other primates, they focus primarily on the Paleolithic period, which stretches from roughly 2.5 million years ago to 10,000 years ago. This is the most commonly referred to EEA because the focus on this time period emphasizes the traits acquired in this time frame that differentiate us from our immediate forbearers and our primate cousins. (p. 51)

Cannot be explained by any other adaptation. Closely related to the necessary evidence of a lock-and-key fit between adaptive problem and adaptation is the idea that an adaptation cannot be better explained as selected to solve an alternative adaptive

problem (Cosmides & Tooby, 1997). This is again the nature of identifying design evidence of adaptations and a critical process in an adaptationist approach.

This also speaks to the authors' basic principle that "different neural circuits are specialized for solving different adaptive problems" (Cosmides & Tooby, 1997). In assuming the perspective of an engineer (i.e., the adaptationist approach and search for evidence by design), the basic engineering principle that the same machine rarely solves multiple problems equally well is assumed to inform this feature of adaptations (Cosmides & Tooby, 1997). Cosmides and Tooby (1997) describe the following:

[...] our minds consist of a large number of circuits that are functionally specialized. For example, we have some neural circuits whose design is specialized for vision. All they do is help you see. The design of other neural circuits is specialized for hearing. All they do is detect changes in air pressure, and extract information from it. They do not participate in vision, vomiting, vanity, vengeance, or anything else.

The authors explain that we have such specialized circuits because the same mechanism, or initially selected mutation, is rarely capable of solving different adaptive problems. Instead, these mechanisms are carefully calibrated to the environmental problem they solve.

As such, an evolutionary perspective suggests that individual adaptations can be identified not only as providing fitness benefits, with a lock-and-key fit between the adaptation and EEA-specific adaptive problem, but the adaptive function must not be derivative from a different adaptation.

Presence/absence in phylogenetic relatives. Evolutionary theory states that the characteristics of related species are inherited from common descent. Common descent clarifies that "closely related species will share many adaptations that can be traced to their common ancestor" (Fowers, 2015, p. 54). Species are considered closely related if

they share a relatively recent common ancestor, and more closely-related species will share more common traits than those with a more distant common ancestor.

One part of identifying adaptations is thus placing them within an evolutionary timeline. Such analysis is conducted by comparative research that works to identify shared features in closely related species, disparate features between those species, and homologous features (those inherited from a common ancestor). By examining these similarities and differences, evolutionary psychologists are better-able to identify the specific periods in which a particular adaptations or species separation occurred.

This phylogenetic approach and the timeline it creates enables us to better understand the environment in which a given adaptation is thought to have emerged. By doing this, we are also able to gain insight into the particular adaptive problems (and existing adaptations) that a species may have faced and can therefore solidify the lock-and-key fit between these problems and the identified adaptation.

One example of this search for the presence or absence of a particular feature in close phylogenetic relatives that is particularly relevant to the current dissertation is Brosnan and de Waal's (2003) study examining what the authors viewed as potential building blocks for morality in Capuchin monkeys. In this study, monkeys were separated in pairs and each was given an identical task in exchange for a reward. When both monkeys were rewarded with a cucumber, only 5% of the rewards were rejected. When one monkey was offered a cucumber after observing another rewarded with a grape, however, their rejections of the cucumber soared to 50%. When they observed the other monkey given grapes without performing a task at all, the rejections grew even higher and more monkeys dissolved into tantrums that the authors describe as mimicking the

behaviors of toddlers who observed their siblings rewarded with whole cookies after receiving only half themselves.

de Waal (2006) refers to this behavior as reflecting “prosocial tendencies” and distinguishes between phylogenetic relatives by explaining that “as opposed to primates marked by despotic [dominance by a single member] hierarchies (such as the rhesus monkeys), tolerant species with well-developed food sharing and cooperation (such as capuchin monkeys) may hold emotionally charged expectations about reward distribution and social exchange that lead them to dislike inequity” (p. 48). In this sense, we see the author beginning to build an evolutionary theory of morality using a phylogenetic approach. De Waal (1996) also makes the controversial assertion of proto-morality in other primates. In doing so, he focuses specifically on observations of empathy, help, and consolation among chimpanzees and refers to these as the “building blocks” of morality. These “prosocial tendencies” and “proto-morality” among primates suggests a relationship between ethical mindedness and moral functioning between humans and other species. I will discuss these shared and different features in greater detail as they relate to cooperation in Chapter 3.

Develops naturally and without explicit learning. It is important to note that evolutionary psychology is not “just another swing of the nature/nurture pendulum” (Cosmides & Tooby, 1997). A longstanding and defining feature of evolutionary psychology has been the rejection of the nature/nurture dichotomy (Cosmides & Tooby, 1997). In evolutionary psychology, it is emphasized that all the features of human psychology have a biological element to them; but the biology in and of itself is not sufficient to explain or create human nature. There has to be an environment that allows

the inherited feature to be properly expressed. In this sense, we have to be born prepared for a natural development that only occurs through natural interaction with our environment.

One example of this is attachment. For attachment to properly occur, babies have to be born ready to attach; but if the necessary nurturing isn't provided, then the natural preparation goes away and the child is left with devastating physiological and psychological deficits. Proper development requires both nature and nurture, and the two cannot be dichotomized or separated. Evolutionary psychology focuses on the architecture and adaptations that define our shared nature while recognizing the richness and diversity of behaviors, experiences, and beliefs that emerge when that nature is expressed through different environments and nurturing experiences.

Although evolutionary psychologists do not dichotomize nature and nurture, it remains important to distinguish between an adaptation and the product of learning or socialization. This is the core point of the current criteria: In order for something to be considered an adaptation, it must emerge naturally over the course of development.

For example, we have the biological gift of developing spoken language. The capacity to develop language is something we are almost universally born with, and spoken language is almost universally developed simply through exposure and emerges without explicit teaching (although explicit teaching often enhances language capacities). Reading, in contrast, does not develop in the absence of explicit learning. We have gone on to create writing, reading, poetry, sagas, and graphic fictions, but all of these are cultural innovations that must be taught and learned, not abilities that emerge naturally through development. Although a complex network of evolved traits contribute to the

capacity to develop reading abilities, some of which relate closely to our natural capacity for spoken language, the trait itself is not one that can be viewed as a specific adaptation in part because it does not emerge without explicit effort and input because we do not see it emerging in infants or without teaching in the same way that we can observe the mimicking of sounds or attention to spoken language.

So what does it mean to develop naturally? It would be easy to confuse this feature with the suggestion that an evolved trait must be present at birth, which is not the case. As put by Cosmides and Tooby (1997), “this is to confuse an organism’s ‘initial state’ with its evolved architecture. Infants do not have teeth at birth – they develop them quite a while after birth. But does this mean they ‘learn’ how to have teeth?” Certainly, they do not. It means that they are born prepared to have teeth; the developmental process of growing teeth is naturally occurring and included in our genome, but it expresses itself only at a certain point.

When we refer to an evolved architecture or trait, we are talking about something that is species-typical – or something that reliably emerges in most individuals of a given species. “In fact, an aspect of our evolved architecture can, in principle, mature at any point in the life-cycle, and this applies to the cognitive programs of our brain just as much as it does to other aspects of our phenotype” (Cosmides & Tooby, 1997). Ethical mindedness and moral functioning, as I will continue to examine in this dissertation, is another example of this. We will observe the building blocks or natural preparation for ethical mindedness in infants while mature moral functioning emerges over time and through social interaction.

Ubiquitous across cultures. In keeping with the idea of natural development, further evidence of an adaptation is that the trait can be found across cultural groups and time. This is similar to the importance of demonstrating natural development, and another way of providing evidence that the trait emerges regardless of what specific social or environmental input is received. In other words, a trait cannot be said to be species-typical unless it can be reasonably observed across cultural groups. Absent this evidence, the trait could be considered the product of culturally-specific learning, and thus lacks evidence of being a species-typical adaptation.

Cosmides and Tooby (1997) introduce the view of evolutionary psychologists by explaining the following:

All normal human minds reliably develop a standard collection of reasoning and regulatory circuits [...]. These circuits organize the way we interpret our experiences, inject certain recurrent concepts and motivations into our mental life, and provide universal frames of meaning that allow us to understand the actions and intentions of others. Beneath the level of surface variability, all humans share certain views and assumptions about the nature of the world and human action by virtue of these human universal reasoning circuits.

This is perhaps, at first blush, a controversial point of view in the context of today's focus on multiculturalism and diversity. There are two important caveats to understand the emphasis of this criteria of ubiquity. The first is that the authors' claim of ubiquity or universal presence is not one that assumes no exceptions can be found for any identified trait. Clearly with a subject matter as variable, wide, and intricate as the human mind and nature there will always be exceptions. By focusing on "universality" in this case, evolutionary psychologists are instead looking for "species-typical" traits. While that will be the term I employ, and the interpretation favored by this dissertation, I have

nonetheless included quotations such as this that make claims of universalism for their applicability to the basic point.

Beyond this, and perhaps far more importantly, is the issue that the authors' statement may be read as a challenge to, or a failure to account for, today's focus on and evidence of rich difference and diversity across cultures. To those of us who so value this focus and have even a cursory understanding of the historic failings in our discipline, it can seem at best dubious, and at worst a perpetuation of past mistakes, to claim species-typical views and experiences. I count myself among this skeptical party and want to make abundantly clear my own reading of this criterion.

While the challenge of claiming universality in a deeply and complexly diverse world presents us with a unique dilemma, the solution to that dilemma lies in identifying what specific level of generality is required to focus on species typical versus culturally variant features. Universals are, by definition, observable only at a high level of abstraction and generality. Cultural elements are much more specific; they specify what these universal generalities mean or how they manifest in a given context.

For example: Cultural context teaches us the basics of behavior and social norms. Culture teaches us how to belong and creates the context for belonging, but our ability to learn those social conventions, our inclination to conform to them, and our instinct to pass them on to the next generation are adaptations. The specific conventions we favor and pass on are culturally variant.

In other words, the focus of evolutionary psychologists is on identifying circuitry and instincts. By noting that an instinct can only be identified as such and counted as an adaptation or species-typical feature of human nature, if it exists across cultures, we

ensure that human nature is not to be described or explained based on the qualities of one cultural group over another. In introducing *The Adapted Mind* (1992), Cosmides and Tooby explain that their central premise by stating,

[...] there is a universal human nature, but that this universality exists primarily at the level of evolved psychological mechanisms, not of expressed cultural behaviors. On this view, cultural variability is not a challenge to claims of universality, but rather data that can give one insight into the structure of the psychological mechanisms that helped generate it. (p. 5)

We seem not to have trouble saying people in all cultures have binocular vision, or are bipedal, but we struggle with the notion of universals when they are psychological or social. The goal of evolutionary psychology, and the one to which this dissertation is intended to contribute, is to meet that challenge, and to continue working towards conceptualizing our basic nature not only in terms of biology, but also psychology. There are certain general and abstract universals to our psychological nature; these include not only that we are designed to learn, follow, and pass on our cultural norms, but that we are all ethically minded.

Neural circuits. Given this enormous challenge to identify ubiquity across cultures, perhaps the strongest case for an evolved trait can be made if a specific neural circuit can be identified for the adaptation. Within evolutionary psychology, however, this is neither typical of identified traits nor required for two reasons. First, the limitations of existing knowledge and technologies generally preclude the decisive identification of specific neural circuits. Second, and as it relates to the current case, the complexity of ethical mindedness mandates the presence of neural circuits designated for many relevant but simpler processes. However inconclusive it may remain at this time, we have begun

identifying evidence of some neural circuits for these simpler processes. This evidence will be discussed in Chapter 3.

Concluding Points

The primary purpose of this chapter was to introduce an evolutionary approach and explain the seven basic criteria for assessing claims about adaptations. Taken together, these features identify the evidence necessary to claim that an identified trait is in fact an evolved feature of human nature. They are outlined here for two purposes. First, they will form the framework from which moral psychologists' existing explanations for ethical mindedness will be examined for theoretical adequacy. Second, they will provide the structure around which I will begin building a more adequate theory of and argument for ethical mindedness as a species-typical feature of human nature.

Enumerated for clarity, the criteria are:

1. Evidence that the adaptation provides fitness benefits
2. Evidence of a lock-and-key fit between the adaptation and adaptive problem
3. The adaptation cannot be better explained as solving an alternative problem
4. Evidence of the adaptation's presence or absence in phylogenetic relatives
5. Evidence the adaptation develops naturally and without explicit learning
6. Evidence of ubiquity across cultures
7. Evidence of related neural circuits

This framework is thus intended to address the research questions of, first, is ethical mindedness a demonstrable characteristic of human nature? And, second, what might an adequate theory of ethical mindedness look like?

CHAPTER 2: MORAL PSYCHOLOGY

As social scientists drawn towards ever-greater and more complex questions, we may all someday look back to find our exploration was most led astray by a perfunctory start, riddled with blind presuppositions about the subject itself. This is perhaps the best argument for the importance of the current dissertation: That without a conceptualization of the basic elements of human nature – one that includes our fundamental ethical mindedness – research across disciplines can be led astray; built around questions and ideas grounded in unchecked, inaccurate assumptions about the human subject. While the field of moral psychology is vast and focused on empirical and theoretical questions that go far beyond the scope of this dissertation, my focus here is to examine how moral psychologists answer this single ontological question: Why are human beings ethically minded? What is it that makes us characteristically and persistently concerned with morality?

As I emphasized in the introduction to this dissertation, it seems reasonable to look to moral psychology as a subdiscipline able to explain why ethical mindedness appears to be a feature of human nature. However, and as I will explore in this chapter, it does not appear that this subdiscipline has yet produced a fully systematic, cogent, and compelling explanation.

Moral psychology is a relatively young subdiscipline that emerged from work across social, personality, developmental, and evolutionary psychology. This fragmentation reflects a similar lack of cohesion across the social sciences and has left the field without clearly agreed-upon definitions for central concepts and terms. Even morality itself is studied at times as it relates to matters of justice and fairness while other

authors focus on caring and harm or social consensus about the evaluation of specific acts (Blasi, 1990; Graham et al. 2011; Tisak, Tisak, & Goldstein, 2006).

Reflecting this fragmentation, the field appears to have structured itself largely by organizing into three theoretical camps or perspectives. These include the Social Cognitive Domain approach, the moral identity perspective, and Moral Foundations Theory. For the purposes of clarity, I have organized this chapter according to these three camps. Each section will begin with a brief explanation of the theoretical perspective, followed by an analysis of quotations demonstrating how or if the basic question of why human beings are ethically minded is addressed (particularly as it relates to an evolutionary perspective). These quotations are selected following a thorough review of the literature and identification of which statements appear to be the best ways that theorists in each group addressed the question of whether and why human beings appear to be ethically minded. This is important in order to give due credit and to avoid presenting straw persons. Each section will conclude with a brief summary of the key concerns and contributions of that perspective.

Social Cognitive Domain Theory

I begin the more detailed exploration in this chapter with a review of the work from psychologists writing from the Social Cognitive Domain perspective because these authors offer what I view as the most circumscribed theoretical framework within moral psychology and because this perspective has been active the longest. They primarily examine morality from a developmental perspective, citing principles of evolution to explain preexisting capacities that enable moral development. The authors in this domain, which is dominated by the work of Turiel (2002, 2003) and those who follow him, view

morality in terms of justice, welfare, and rights. In this sense, their work reflects a universalist view of morality like that seen in Piaget's and Kohlberg's psychology (Killen, Margie, & Sinno, 2006). They therein assume that a basic moral trait exists as a universal feature of human nature, but do not appear to fully explore or explain that trait. Although I refer to this trait as ethical mindedness, I have not observed moral psychologists from any perspective using that term.

Based on Turiel's (2002, 2003) influence, authors writing from a Social Cognitive Domain perspective view morality as including three domains: the moral, conventional, and personal. The moral domain, following the universalist view, is exclusively interpersonal and related to helping others, harming others, or justice, rights, and fairness (Smetana, 2006). This is distinguishable from the personal domain, which occurs independent of interpersonal relationships or interactions and is believed to affect only the individual (Nucci & Turiel, 2000). The conventional domain refers to externally-established rules and regulations. Although such rules are likely interpersonal or social and may overlap with moral expectations, they are contextually determined and subject to change across social environments. The dependence on a specific social context is what distinguishes conventional from moral domains. It is thus the inherent universality of morality that defines the moral domain for Social Cognitive Domain theorists.

Authors writing from a Social Cognitive Domain perspective thus focus on morality as a universal product of human development that emerges through social interaction and focuses primarily on matters of interpersonal helping and harming (Smetana, 2006). This universality of morality suggests that these authors identify morality as part of our basic nature. While this is an inherent element of a sound theory of

ethical mindedness, the authors from this perspective do not appear to have fully explored the question of *why* morality appears to be universal, or how it came to be an essential feature of human nature.

Some authors indicate that this is due to their emphasis on the process through which individual morality develops. Turiel and Killen (2010) argue that the focus on human development, and not the work of evolutionary psychologists, is most likely to provide an adequate understanding of moral functioning. The authors make this claim in the following statement:

Evolutionary psychologists too often provide the circular story that if it exists, it must have evolved and that if it evolved, it exists to further preservation of the species. An alternative is to examine the process developmentally from the perspective of how emotions and thoughts are formed in people's social interactions and reflections on their experiences. (Turiel & Killen, 2010, p. 38)

There are multiple conceptual problems with this statement. First, and perhaps primarily, is the false binary suggesting that one must focus on either evolution or individual development as two entirely separate phenomena. The authors appear to justify this logical fallacy (false dichotomy) using a misunderstanding of what constitutes an evolutionary claim. While it is not uncommon for laypeople to cite evolution in such circular arguments, that is not the case for sound evolutionary reasoning. In contrast, a strong evolutionary argument would identify the specific adaptation that evolved to solve a specific problem and, by doing so, make the case for how this specific adaptation became a core feature of human nature. Evolutionary psychologists would then explain how that feature of human nature influences individual development, or what developmental evidence has been observed to support claims of this feature's ubiquity. As I mentioned in Chapter 1, the necessary criteria for identifying adaptations are

specifically designed to address this concern about just-so stories in evolutionary psychology. Not only are such critiques already addressed by the application of an evolutionary perspective, but it is difficult, if not impossible, to tell a complete developmental story without a solid conceptualization of the being that is developing or how the features of being came about. Similarly, a good evolutionary account includes a description of how an adaptation emerges naturally in development.

Claiming an alternative to evolutionary thinking, Turiel and Killen (2010) focus instead on the developmental process that occurs when individuals reflect on their experiences. In one example of how people learn to be moral through empathy, these authors explain, “Most likely, judgments that involve empathy derive from prior personal experiences with exclusion, which, to some extent, enable individuals evaluating a decision to exclude someone as wrong” (Turiel & Killen, 2010, p. 45). In general, they appear to claim that because we are empathic, we can tell when we’re benefitting and harming others by sensing the effect that we have on them. Let’s assume they are correct that judgments involving empathy derive from past experience with one’s own exclusion. Despite this potentially accurate statement, the authors seem not to realize that they have avoided the evolutionary question, creating only a second question for which evolution remains likely to offer the clearest answer.

In other words, to develop the belief that exclusion is wrong based on experience, the past experience of exclusion must be somehow negative or painful enough to teach us that it is harmful. So if the question “why are human beings ethically minded” is answered with “because human beings experience empathy” then we are left asking “why do human beings experience empathy?” And if this question is answered with “because

past experiences of exclusion teach us empathy” then we are now left to ask “why would negative experiences teach us to be empathic toward others?” and “why do we have a sufficiently negative experience of exclusion to learn from it?” While this repetitive line of questioning may eventually lead us to an adequate answer, the current theorists have stopped with the initial and unsatisfying conclusion that human beings are ethically minded because we learn from experiences of social exclusion. Perhaps even more problematic than this, they have explained in part how it is that we know how other people feel but said nothing of why it is that we care. Why does knowing we have hurt another person often change our behavior? Despite their claim that evolutionary reasoning serves little utility to understanding moral functioning, the arguments they make seem to lead us directly to questions that can be answered only by evolutionary explanations. This is because questions about relatively ubiquitous features of humans require an ultimate explanation—our ancestors acquired these characteristics and were therefore more reproductively successful.

This is notable given that questions about empathy and social exclusion have many clear evolutionary answers and explanations. As I have made reference to when describing our evolved nature in Chapter 1, we evolved to experience social exclusion as deeply painful. The same authors who suggested that evolutionary theory does not provide a fruitful answer and that our focus should instead be on the process of development have perhaps inadvertently gestured at an evolutionary answer for the process of development with their claims about past experiences. These two processes cannot be separated, nor can any question about human nature be answered without the inclusion of evolutionary evidence to explain human nature and natural development. As

I described in the introduction to this dissertation, such conceptual integration is not optional in the social sciences. It is mandatory if we hope to develop adequate theoretical frameworks for understanding human nature.

I referred above to the reasoning used by Turiel and Killen (2010) as having avoided the evolutionary question. We see this issue again in the work of Helwig and Turiel (2011). These authors explain, “children's domain distinctions are based on early social experiences. The development of morality involves constructions of judgments about welfare, justice, and rights through multiple social relationships and experiences” (p. 572). Similarly, Jambon and Smetana (2014) offer that “Perspective-taking abilities, including the capacity to coordinate one’s own perspective with those of another, have long been theorized to play an important role in the emergence of more mature moral reasoning” (p. 2).

In these cases, the authors have answered the question, “Why do children develop morality?” with, “Through social relationships and experiences” and “because they have perspective-taking abilities.” The idea that perspective-taking abilities interacting with social experiences contribute to the development of mature moral reasoning is not an inaccurate narrative. It simply isn’t a complete one, nor is it one that adequately explains the ontological question of why this development ultimately occurs. While the authors have taken steps towards answering the question of how moral development occurs, they have simply shifted the question about our basic nature from, “why do children develop morality?” to “why are children born with perspective-taking abilities and the capacity to learn from social relationships and experiences?”

There is no reason to doubt the importance of social experiences or perspective-taking abilities in the development of individual moral perspectives. But provided no evidence that these capacities are part of our basic nature, and no explanation for how that came to be, their ultimate theoretical perspective falls short of explaining core features of human nature and the uniquely human development that follows.

These broad gestures at the ontological question of ethical mindedness are seen among other Social Cognitive Domain adherents. Nucci and Turiel (2009) state that moral thinking “develops out of constructions through children’s social experiences and interactions” (p. 152). Smetana, Jambon, and Ball (2013) explain that, “[sentiments such as] empathy or avoidance of harm are viewed as constituting the building blocks on which moral knowledge is constructed” (p. 28).

Even the most charitable interpretation of these statements would offer only the idea that humans possess presumably-evolved capacities onto which individual morality develops through social experience; but that is neither stated explicitly nor systematically explained. Here we see vague gestures at an unexplored basic nature and the role of social interaction without any systematic understanding or exploration of what that basic nature is, how it came to be, or how exactly it interacts with the social environment and what happens when it does. These authors appear curious about human nature but unable to grasp onto the evolutionary perspective that would explain that nature and provide them with a starting point for exploration of the social interactions that trigger and contextualize the emergence of moral thinking.

A more articulate view of the idea that humans are characterized by basic evolved capacities, through which individual morality is developed through the repeated

interaction between these capacities and the social world, is presented by Killen and de Waal (2000). These authors move further in the direction of an evolutionary explanation yet engage in the same lack of ultimate explanation previously observed and thus leave the ultimate question open. We see this in their statement:

- (1) there are specific types of behavior demonstrated by both human and nonhuman primates that hint at a shared evolutionary background to morality; and
- (2) there are theoretical and actual connections between morality and conflict resolution in both nonhuman primates and human development. (Killen & de Waal, 2000, p. 352)

The authors further explain that “[...] the phylogenetic origins of human morality can be detected in the social interactions of nonhuman primates [...] and the ontogenetic origins of morality can be observed in the early social interactions of children” (Killen & de Waal, 2000, p. 352). This notion is reiterated in the statement, “[...] we are born with a moral ability and a tendency to be influenced by the moral values of our social environment, but there is little evidence to suggest that we are born with a ready-made moral code in place” (Killen & de Waal, 2000, p. 354).

These are arguably the strongest claims about ethical mindedness that can be found in the Social Cognitive Domain literature. But despite this relative strength, even these quotations remain plagued by two theoretical shortcomings. First, these gestures do not adequately fulfill the basic requirements for an evolutionary argument; if the authors believe that morality develops on or through evolved capacities, much more work can and should be done to demonstrate the existence of such capacities. They provide some phylogenetic evidence (which is to be expected given that de Waal is a primatologist and second author on this paper), but do not address further criteria for evolutionary claims. Second, and subsequently, the authors’ unexplored gesture is presented as a brute

assertion without explanation for how or why such abilities exist. Instead of asking, “why are human beings ethically minded?” we are now left to ask, “why are human beings born with a moral ability and tendency to be influenced by the moral values of our social environment?”

In other words, these authors gesture at the presence of phylogenetic and developmental evidence, thus making a seemingly implicit argument for an evolved predisposition, without doing the work of fully building a case for ethical mindedness as a feature of human nature. They do not clearly identify the evolutionary problem, the adaptation that evolved to solve that problem, or the relationship between that evolved trait and what they presently identify as the propensity to develop individual morality. They begin the work of identifying evolved capacities, but it is unclear to whom the authors leave the task of understanding what specific traits evolved to create this propensity. Without identifying such traits or beginning to make an adequate evolutionary argument to justify their claim, this particular case for moral development as predicated on evolved predispositions remains thin.

Taking a charitable interpretation of this theoretical perspective, we can identify two valuable points in the Social Cognitive Domain for an evolutionary account: The first centers on the notion that humans have essential characteristics that enable our unique capacity for moral development; the second and more prominent point emphasizes that an individual’s morality is constructed through a lifetime of social interaction. Although neither of these points is wholly incompatible with the evolutionary argument for ethical mindedness, it is unclear whether the authors would suggest that they constitute the whole ontological story. It seems that the Social Cognitive Domain perspective has

avoided the evolutionary question. These authors offer rich theoretical perspectives on moral development, but fail to address the basic ontological question that may be the most fitting starting line for their entire subdiscipline: Why are human beings ethically minded?

Summary. Perhaps the Social Cognitive Domain's greatest contribution lies in its basic conceptualization of the moral domain. By identifying morality as universal, and building theory and research upon that idea, the authors in this field have referenced the element of ubiquity that is itself central to a sound evolutionary argument. Despite the inadequacy of their evolutionary claims, they highlight the natural process of individual development and the ubiquity across cultures and times that can and should be included in a more comprehensive theory of ethical mindedness. These authors clearly understand that moral development occurs universally, and likely that it occurs independent of specific cultural context. Some also appear to grasp that we see similar building blocks in our closest phylogenetic relatives. Taken together, they gesture at features of human nature and a more substantial literature on the subsequent human development.

That these basic features are included only as gestures with emphasis placed on individual development, however, is also a significant shortcoming of the Social Cognitive Domain. Without the innate capacities to recognize and respond to ethical issues, no amount of childhood experience could teach children to have a moral framework; nor could we explain much of the evidence explored in Chapter 3 (some of which demonstrates moral functioning and behavior in pre-linguistic infants). Ethical mindedness is analogous to language. If children were not prepared to understand and express language verbally, no amount of talking would teach them how to speak. We

require carefully designed vocal cords, mouths, tongues, and cognitive capacities, which all evolved to make the learning of language possible. Language is not taught from scratch; babies are born prepared to learn it. The same is true of ethical mindedness; we are born prepared to absorb and develop moral perspectives.

These authors may understand that ethical mindedness, or the propensity to form moral perspectives, is innate to humans and has emerged from a constellation of evolved adaptations and traits. But the general perspective from the Social Cognitive Domain has yet to explain or justify what basic features of human nature are at play in moral development. As I cited from Turiel and Killen (2010), the authors mention empathy and perspective taking. These are likely important, species-typical features of human nature that are closely related to moral functioning. However, as I discussed above, the authors' exploration is inadequate. They fail to explain why we care how other people feel. They do so proximately by explaining that our own past experiences inform our awareness of how others feel, but they offer no ultimate explanation for why we, as a species, appear to be generally concerned with the feelings of others (at least some other people). They do not explain how we evolved to be that way or what characteristics enable it.

What this costs us in terms of theoretical sufficiency is significant. By arguing that the basic propensity for morality is an innate feature, but making little or no attempt to understand the emergence of that feature (or in the case with Turiel and Killen (2010), even dismissing the importance of doing so based on a dramatic misunderstanding of evolutionary science), these authors are building a theoretical framework with no foundation, and there is no guarantee that their neglect of ultimate explanations will not lead them astray in their theorizing.

Moral Identity Perspective

I explored the Social Cognitive Domain perspective first as it is the most circumscribed and longest standing of the three theoretical camps in moral psychology. I will discuss the moral identity perspective second because this view, while slightly broader and more varied across individual authors, also remains fairly focused and is closely related to the ideas that have been developed by Social Cognitive Domain theorists.

Moral identity authors are a more loosely knit group of theorists who don't assume a single theoretical perspective but generally share a focus on moral identity, character, and reasoning.¹ These authors define moral identity as a commitment to the well-being of others that is central to an individual's identity or self-perception (Hart, Atkins, & Ford, 1998). They emphasize the need for consistency between this self-perception and the individual's actions, which are typically interpersonal and observable in what they consider "prosocial behaviors" (e.g., Aquino, Freeman, Reed II, Lim, & Felps, 2009) or through an "altruistic personality" (e.g. Carlo, 2006; Hardy & Carlo, 2011).

The emphasis on the need for consistency is likely related to the emergence of moral identity theory having been inspired by Kohlberg's focus on moral reasoning and the gap between reasoning and action; or the consistency or lack thereof between one's moral reasoning and behavior. Moral identity theorists aim to explain that gap. Citing Walker (2003), Darnell, Gulliford, Kristjánsson, and Paris (2019) explain that "Kohlberg's approach was challenged empirically in seminal papers by Augusto Blasi

¹ Because these loosely connected authors don't typically identify "moral identity" as a singular theory, I will not capitalize that label in the current dissertation.

(1980, 1983) in which he reported that moral reasoning only accounts for 10% of the variance in moral behavior” (p. 4). Moral identity theorists suggest that this variance is a measure of the accessibility and centrality of one’s moral identity. If a person’s moral identity is weak or less accessible, they are unlikely to think that morality is as important a factor in their actions.

In keeping with the less cohesive nature of this group, however, there are several exceptions to these descriptions. Narvaez has also built on the work from a moral identity perspective and more recently focused on neurobiological roots (Narvaez, 2008) and attachment parenting (Narvaez, 2014) as potential sources or explanations for individual moral identity. To maintain focus on the subject of the current dissertation, this section will begin by emphasizing the statements made by authors writing from a moral identity perspective that broadly focus on ideas related to why human beings are ethically minded. Narvaez’s complementary contributions in Triune Ethics Theory will be considered after describing the moral identity perspective.

Aquino, Reed, and Diener (2002) compare their perspective to the Social Cognitive Domain by explaining that, "Moral identity does not supplant the cognitive-developmental model or the idea of moral reasoning as a predictor of moral action. Rather, it complements this approach by identifying a social-psychological motivator of moral conduct" (p. 1425). This shared view is also summarized by Narvaez (2010) who states that moral identity is “cultivated within the dynamic interplay among the developing organism; environment; and the timing, duration, and intensity of interaction” (p. 171).

Moral identity theorists herein describe their expansion upon the cognitive-developmental model as primarily based on their view of an individual's moral identity provides a unique motivator for moral intentions and actions. In this sense, we can view their answer to the question of why human beings are ethically minded as grounded in their focus on the centrality of moral identity and its development through social and environmental interaction.

Citing both Blasi (1983) and Festinger (1957), Aquino and Freeman (2009) describe that the “motivational potency of moral identity arises from the human desire for self-consistency” (p. 379). In this, the authors reference the notion that humans experience an essential need for their external self-presentation to accurately reflect their subjective self-perception or, in this case, their moral identity. Put more simply, these authors believe that all human beings possess a moral identity, or a sense of themselves as moral beings; and they view the motivation for moral action as grounded in a need for our behavior to be consistent with our internal identity.

Summarizing their central hypothesis and expanding upon these basic perspectives, Aquino, Freeman, Reed, Lim, and Felps (2009) thus offer that, “Moral intentions and behaviors will be a joint function of (a) the centrality of moral identity to an individual's self-conception and (b) the extent to which situational cues temporarily affect the current accessibility of the moral self-schema within the working self-concept” (p. 124). This initial emphasis on the variable centrality of one's moral identity is reiterated by Hardy and Carlo (2005), who note that, “[...] when one's self is centered on moral concerns, this inclination serves as a key motivating force for moral action” (p. 235). While there are elements of these quotations that present concerns for a developing

theory of ethical mindedness, a significant contribution here is the idea that morality may be more or less central to a person's identity depending on the individual and on the context.

Aquino and Freeman's latter point about the extent to which situational cues affect the accessibility of a moral self-schema is also consistent with a notion frequently reiterated by Narvaez (2013), who notes that, "An individual may learn to habitually use one mind-set or another [... and that] certain situations [...] may evoke a particular mind-set" (p. 244). In their more recent work, Hardy and Carlo (2011) also incorporate this emphasis on situational influences by stating that, "Moral identity formation is associated with certain individual characteristics, developmental contexts, and opportunities for moral action" (p. 214). Walker (2004), concurs that morality is, "a fundamental and pervasive aspect of human functioning that consists of interpersonal and intrapsychic components" (p. 5).

Moral identity theorists are thus in considerable agreement that some basic feature of ethical mindedness is not only a broadly defining element of our humanity, but also a core part of both individual identity and interpersonal behavior. Although these authors therefore expand upon the more narrowly developmental perspective taken in the Social Cognitive Domain, their broader theoretical perspectives nonetheless seem to again leave the ontological question of why human beings appear to be ethically minded unaddressed.

In these quotations, we see the issue of avoiding the evolutionary question occurring in a more simplified way. The question remains of why human beings are born with a moral identity or moral self-schema at all. What, exactly, is it that is elicited or

developed through social interactions? What is accessed? What do we see as variably central to each individual's self-concept and where does it come from? These authors may not be explicitly attempting to answer the question of why human beings appear to be ethically minded, but the remaining questions nonetheless appear to be, "but why do human beings so readily develop these different mind-sets?" "Why does moral identity formation so readily occur?" And, "Why is morality a fundamental and pervasive aspect of human functioning?"

Hardy and Carlo (2005), refer to a self that is, "centered on moral concerns" (p. 235) and later (2011) explain that, "identity is unavoidably moral (i.e., there is no such thing as an amoral identity) in that morality is inherently a part of our world, and our identity is the framework from which we view and relate to the world" (p. 500). Although elucidating some points, the circularity of this perspective ultimately leaves one repeating the same, basic question. If the answer to, "why are human beings moral?" is, "because morality is a part of our world and our identities reflect our world," the obvious follow-up is, "then what makes our world moral?"

Although no theorist is likely to fall into the trap of answering with "because human beings are moral," we again see a complicated form of avoiding the question in the work of Blasi (1999). In one article, Blasi focuses significantly on the relevant contribution of spontaneous emotions and agentic, post-hoc reasoning in his exploration of the development of morality. He explains,

A natural tendency to respond empathetically may predict later development of altruistic concerns; early natural desires for revenge may be related to a later concern for justice (Solomon, 1990). [...] To become moral, however, natural emotions and motives need to be shaped, refined, and directed according to a specific framework which is not naturally given; this work of reorientation demonstrates the moral limits of spontaneity. (p. 14)

He expands upon these claims with the following:

There is another way of looking at emotion and its relation to moral motivation and moral action. Here, *spontaneously* elicited emotions indeed lack the intentionality and the specific moral meaning that would be necessary for them to produce moral motivation, but the *agentic* regulation of spontaneous emotions would then be emphasized and become the focus of the explanation. [...] In this view, there are two ways for emotion to have moral meaning and therefore contribute to moral motivation. In one way, a spontaneous and unintentional emotion may be a result of already existing moral concerns; the emotion and its action tendency, then, would have the same meaning as the concern from which they spring (Montada, 1993). In the second way, the emotion would be shaped and directed by agentic regulatory processes. Since this type of regulation is consciously motivated, it can be guided by concerns that are specifically moral. (p. 14-15)

While Blasi (1999) does not appear to be grappling with the idea that people need to be shaped and directed in unnatural ways if they are to develop any morality at all, he is nonetheless struggling with the basic question of what, when it comes to moral identity and action, is instinctive or automated and what is learned. Blasi suggests that there are two ways for emotion to have moral meaning – spontaneously or as shaped and regulated by agentic processes guided by concerns that are specifically moral. He does not say, however, how spontaneous emotions are related to moral concerns. One reading may be that what's moral is what we have internalized, but is it better to consciously think about morality or for it to be habituated? There appears to be some suggestion here that Blasi believes it is better to be agentic, but the quote ultimately appears to be equivocal.

It also isn't clear what Blasi means by unintentional emotion and whether that is inherited or something that has been internalized through moral education and he does not offer us an explanation of how we inherited these emotions. Blasi (2009) later states that "morally relevant tendencies exercise a crucial function for the survival of the human and other animal species [and are ultimately] a result of evolutionary pressures and

selections” (p. 400), but even this appears to just throw out there that “we must have evolved this way” without offering a meaningful evolutionary explanation. It is ultimately only a gesture.

This idea that morality is ultimately constituted by innate features of human nature, coupled with developmental outcomes, is consistent with an evolutionary perspective. Where Blasi’s conceptualization appears to fall short is in its failure to fully grapple with what the basic feature of human nature is – what those instinctive emotional reactions are and why we have them. This raises but does not answer the question of why we are ethically minded, and what evolutionary evidence supports a theory of that basic feature. He appears to understand that, at a basic level, the building blocks for morality must have been present in matters of survival and reproduction in order to have emerged as universal features of human nature, but he does not expand on that understanding.

In this, we see the moral identity authors moving much closer to theoretical adequacy despite doing so without fully exploring their own hypothesis. There is strong evidence for Blasi’s (2009) hunch about the relationship between moral tendencies and survival. For example, in Chapter 3 I will more thoroughly present evidence that our ventral striatum, or the part of the brain that registers pleasure, activates more strongly when we perform an act of kindness for others than when we do so for ourselves. That is an interesting fact that suggests a form of morality or ethical mindedness may be a basic feature of human nature.

Provided a generous interpretation of his conclusions, Blasi (1999, 2009) may be making a similar claim with reference to instinctive emotional reactions and their relationship to moral identity. In and of itself, however, the finding that humans take

pleasure in caring for others does not explain *why* we are ethically minded. The question remains how it is possible that we became the kind of creature that would find pleasure in that. Therefore, even some of the more nuanced claims by moral identity theorists fall short of an adequate argument about our basic nature. Given this, Blasi's statements leave us to wonder how, specifically, these emotions and cognitions relate to morality or ethical mindedness, as well as how the instinctive emotional reactions and cognitive capacities came to be universal features of human nature.

Another example of moral identity theorists grappling with the relationship between evolution and our moral nature comes from Hill and Lapsley (2009). In contrast to Blasi's focus on moral reasoning and emotion, these authors emphasize gossip and attention to moral and immoral people and actions. They explain the following:

[...] differential attention to morality in the breach is considered adaptive insofar as avoiding those seeking to betray us would lead to greater evolutionary success (Krebs, 2008). [...] One reason is that people generally may be more likely to discuss immoral people than moral people, because it would be more evolutionarily adaptive to point out those individuals who may cause harm. However, this is speculative and a topic for future research. (p. 522)

While this may read as the kind of circular, colloquial evolutionary argument criticized by Social Cognitive Domain authors above, it is not an inaccurate one and it need not be qualified as speculative. I will again explore this evidence in greater detail in Chapter 3, but the type of evolved capacity the authors mention has been well-documented. Cheater detection, or the ability to discern those who violate social exchange rules or norms, is a well-evidenced human capacity and may be closely related to what these authors claim is a natural ability to detect or attend to amoral individuals. This is another example of how sound evolutionary reasoning and argument may or may not change a moral

psychologist's premise, but it would improve their moral psychology and theoretical perspectives if they included such explicitly relevant evidence.

As I mentioned at the beginning of this section, Narvaez has attempted to relate moral identity and morality to neurological features. This may appear to be more closely related to evolutionary evidence, but the approach from this perspective remains limited. For example, Narvaez (2008) explains that “ethical motives and behaviors [arise] out of biological propensities” (p. 96). She does not specify what is meant by these biological propensities, but she does add that, “much of who we are and what we do is traceable to the more ancient parts of the brain [...] The sympathetic, parasympathetic, limbic, and related systems must be regulated by caregivers early on or deficits ensue and moral optimization may not be possible” (Narvaez, 2008, p. 112). Narvaez (2008b) also appears to gesture at the elements necessary for an adequate argument. In one statement, she describes the development of her theoretical perspective with the following:

[It] is derived from psychological, evolutionary, and neurosciences, emphasizing the importance of the limbic system and related structures for moral information processing and behavior. Most research in moral psychology has focused on the work of the neocortex (e.g., deliberate reasoning), often neglecting the motivational structures that lie underneath. [...] First, [I] emphasizes the motivational orientations driven by unconscious emotional systems that predispose one to process information and react to events in particular ways. (p. 313)

Again, offering a charitable interpretation, Narvaez may here indicate the beginning of an understanding of the role evolutionary science will play in an adequate theory of human nature insofar as any effort to understand the relevant biology will closely relate to a sound evolutionary argument. While Narvaez does not clearly explain neuroscientific research and the inclusion of neurological substrates as part of an evolutionary case, she does focus consistently on such biological evidence. In this, she

recommends a focus on biological propensities but does not go on to explain specifically how they account for human morality or how they arose through evolution. This is another instance of avoiding the question and providing only a brute assertion about biological propensities that is put forth without further explanation about how or why such propensities came to be.

This is a similar issue to that which I illustrated above using the example of the ventral striatum and the pleasure derived from acts of service for others. Again, the question remains: How is it possible that we became the kind of creature that would find pleasure in that? How is it possible that we became the kind of creatures whose emotional systems predispose us in this particular way? To reiterate a reference I used in Chapter 1, human beings also derive pleasure from sweet and fatty foods. In this sense, we evolved to find sweet and fatty foods pleasurable because fruits and animal proteins were good for us. Pleasure is a mechanism and a characteristic, but it is not the full story. Narvaez makes a similar case about our predisposition to process certain information or events in ways that contribute to the development of morality. She does not, however, provide a full case for that mechanism or explain how and why it evolved.

Summary. As I noted at the beginning of this section, authors from the moral identity perspective make a significant and important contribution to an adequate case for ethical mindedness in their understanding of the variable centrality of moral identities across individuals. There is good reason to believe that there will be individual differences in how important morality is to each individual as well as in how it is defined, understood, and acted upon. There is no simple inherited template for morality. As I will explore in Chapters 3 and 4, the universality identified within the Social Cognitive

Domain, and the individual variability seen from Moral Identity authors, will be critical features for an adequate theory of ethical mindedness.

Authors writing from the Moral Identity perspective also grapple with similar issues to those seen in the Social Cognitive Domain. Despite clear efforts to include evolutionary references in their claims and to explain their perspective using features of human nature, the ultimate statements continue to fall short of adequate evolutionary argument and at times engage in another avoidance of the basic evolutionary question.

In this sense, the authors do a better job of building sound theory, or at least appearing to understand what is missing from the existing Social Cognitive Domain perspectives, but ultimately continue to gesture vaguely at evolutionary explanations while making limited attempts at providing them in response to questions about human nature. As was the case with Blasi (2009) and Narvaez (2008b), we see that a better understanding of both existing evolutionary research and argument would significantly improve and clarify their developing theories.

Moral Foundations Theory

Finally, I turn to Moral Foundations Theory. While authors from the Social Cognitive Domain and moral identity perspectives emphasized developmental views, Moral Foundations theorists are more firmly grounded in biological and cognitive perspectives. These authors primarily follow the views of Haidt's (2001) social intuitionist model and focus on moral intuitions, reasoning, and judgment. They view moral intuitions as quick, instinctive, and automated reactions and moral behavior as based on these flashes of emotion rather than the deliberative processes we saw emphasized by some authors from the moral identity perspective (Haidt & Bjorklund,

2007), which they portray as post hoc rationalizations of actions rather than as sources of actions. Such retroactive deliberation is viewed as an occasional, but rare, influence on ultimate moral judgments or actions (Haidt 2001, 2004).

In this sense, Haidt and his followers explain morality as having evolved from intuitive emotional instincts that facilitated adaptive cooperation (Haidt & Bjorklund, 2007; Haidt & Joseph, 2004). They identify five foundations or domains of morality: purity/sanctity, ingroup/loyalty, fairness/reciprocity, harm/care, and authority/respect (Haidt & Graham, 2007; Graham et al., 2011). As will be increasingly clear throughout this section, Moral Foundations theorists do a notably better job of beginning to build an evolutionary case for morality and ethical mindedness. There continue to be some issues in their evolutionary reasoning, but they nonetheless appear to grapple with the question and begin the necessary work of explaining why human beings are ethically minded.

In one example, Haidt and Bjorklund (2008a) explain the following:

[...] the human mind has been shaped by evolution so that children can very easily be taught or made to care about harm, fairness, in-groups, authority, and purity; however, children have no innate moral knowledge – just a preparedness to acquire certain kinds of moral knowledge and a resistance to acquiring other kinds (e.g., that all people should be loved and valued equally). (p. 204)

This assertion that the mind has been shaped by evolution such that we readily absorb and are concerned with matters relevant to morality, but without predetermined knowledge, is a strong one. Without using the same language, the author's idea of a "preparedness to acquire certain kinds of moral knowledge" appears consistent with evidence of the universality of ethical mindedness. In this sense, these authors appear to grapple with the question of why human beings are ethically minded. Despite this identification of the core question, however, the authors do not take clear steps towards

answering it. They simply gesture vaguely at evolution as the place where an answer may be located without doing the detailed work of building that explanation.

Similarly, Haidt and Kesebir (2010) cite evidence that I will use in Chapter 3 in their statement that, “moral-perceptual abilities emerge in infancy” (p. 13). This evidence that morality emerges naturally through development and absent explicit learning experiences is one part of a sound evolutionary argument. The authors go on to explain the following:

The enormous and accelerating gains from cooperation in agriculture, trade, infrastructure, and governance are an example of what has been called a “major transition” in evolution (Maynard, Smith & Szathmary, 1997), during which human beings went from being a social species like chimpanzees to being an “ultrasocial” species (Campbell, 1983; Richerson & Boyd, 1998), like bees and ants, able to live in groups of thousands with substantial division of labor. What “inside the head” mechanisms were already in place in pre-agricultural minds such that when early agriculturalists created the right “outside the head” products – such as virtues, institutions, social practices, and punitive gods – ultra-large-scale cooperation (i.e. civilization) materialized so rapidly? Many explanations have been offered, but two of the most widely discussed are reputation and moralistic rule enforcement. (Haidt & Kesebir, 2010, p. 21)

The general direction of this quotation suggests a much stronger understanding of, and effort to, make an evolutionary claim. In this case, the Moral Foundations theorists attempt to tie specific adaptations to specific problems that were present in unique evolutionary contexts. They also make reference to phylogenetic evidence. While it must be noted that there are significant issues with their reasoning (including the idea that there exists some evolutionary boundary with the advent of agriculture), they nonetheless grapple with the question of an evolved nature and endeavor to address some of the necessary criteria for evolutionary argument.

In doing so, the Moral Foundations theorists expand the ontological examination of ethical mindedness further than the other two theoretical perspectives by identifying

specific psychological capacities that emerge through evolution and form the foundation for human morality. This is, again, an indication of the greater theoretical adequacy in this perspective.

This greater adequacy is grounded in large part by focusing on the social intuitionist model, which emphasizes instinctive emotional reactions, and the five foundations, which locate a wider range of morally relevant domains that emerge through the evolutionary process. Haidt and Joseph (2004) explain this relationship between social intuition and the five foundations with the following:

[...] human beings come equipped with an intuitive ethics, an innate preparedness to feel flashes of approval or disapproval toward certain patterns of events involving other human beings. The four patterns for which we believe the evidence is best are those surrounding suffering, hierarchy, reciprocity, and purity. (p. 56)

Similarly, Haidt (2001) draws a parallel between morality and language as an evolved capacity by stating:

Morality, like language, is a major evolutionary adaptation for an intensely social species, built into multiple regions of the brain and body, that is better described as emergent than as learned yet that requires input and shaping from a particular culture. Moral intuitions are therefore both innate and enculturated. (p. 826)

In examining these quotations together, we see that the authors recognize the innateness of morality, and the propensity to develop it as an evolved feature of human nature. They also identify more specific domains for morality with their emphasis on suffering, hierarchy, reciprocity, and purity. These domains were later expanded and described as harm/care, authority/respect, fairness/reciprocity, and purity/sanctity (Haidt & Joseph, 2007). Subsequently, the authors later added a fifth moral domain identified as ingroup/loyalty (Haidt & Graham, 2007). Such clarified domains are not only far more adequate than the simple focus on interpersonal experience identified by most Social

Cognitive Domain and moral identity authors, but they are also more directly related to potential and existing evolutionary arguments.

Given this, the governing focal points that direct Moral Foundations Theory readily lend themselves to evolutionary argument. The first of these focal points is on quick, instinctive flashes of judgment (moral intuitions) and the related propensity to develop moral viewpoints. The second is on the specific domains that relate to selection pressures across evolutionary time. These domains refer to the sources of such quick reactions, such as being disgusted by impurity, angered by illegitimate authority, sensitive to fairness, and so on. What directly links these foundations to evolution is that our reactions to them appear to be built-in and instinctive; the connection is in the automated nature, which suggests that we evolved to react to these issues.

In this sense, Moral Foundations theorists are probably correct in their view that different moral questions or issues have different evolutionary explanations and sources. That Moral Foundations theorists clearly lay out multiple moral domains distinguishes them from other moral psychologists, who tend to focus exclusively on domains of harm/care and justice/injustice. The basic distinction here is between moral pluralism, or the idea that there are many different moral values, and ethical monism, which would refer to the belief that all of our different values can be reduced to a singular “supervalue” (Mason, 2018). It does not appear that Social Cognitive Domain and moral identity theorists have done the work of reducing multiple values into a single supervalue; rather, they seem to simply not recognize the full scope of the moral domain. We see this expanded scope as a primary focus in the work of Moral Foundations theorists, and it is a significant contribution to moral psychological theory.

Moral Foundations theorists are also likely to be correct in their position that if ethical mindedness is an evolved feature, much of our relevant reactions should be automatic. An adequate evolutionary argument can and will demonstrate that children develop such ethical mindedness before they are taught and that we are able to make relevant decisions without significant conscious deliberation. In all of these ways, the authors provide a beginning to both address the question of why we are ethically minded, and to indicate the importance of answering this question using an evolutionary perspective.

In telling this evolutionary story, Haidt (2010) explains the following:

Human beings have long been ultrasocial creatures embedded in webs of accountability (Lerner & Tetlock, 2003) with a constant need to justify our own actions, monitor the actions of others, and persuade third parties to trust and support us. We are all descended from ancestors who excelled at those tasks. We are quite skilled at reasoning in those contexts and for those purposes (Mercier & Sperber, 2009). Reasoning is a great lawyer but a clumsy dance partner. (p. 183)

These authors are on the right track; they simply haven't moved very far down that road. While their basic direction is not inaccurate, it is also not complete. They have not fully explored the evolutionary evidence for these foundations, which would confirm the evolved tendency for such quick reactions and locate the domains in which they occur.

We see this issue in more detail in Graham et al. (2013). The authors outline the five Moral Foundations and begin the work of connecting those foundations to an evolutionary explanation. Their description of the sanctity/degradation foundation (here referred to synonymously as purity/sanctity) is highly specific. The authors explain the following:

Hominid history includes several turns that exposed our ancestors to greater risks from pathogens and parasites [...]. Individuals whose minds were structured in advance of experience to develop a more effective "behavioral immune system"

(Schaller & Park, 2011) likely had an advantage over individuals who had to make each decision based purely on the sensory properties of potential foods, friends, and mates. Disgust and the behavioral immune system have come to undergird a variety of moral reactions. (p. 71)

In this, the authors make a case for the evolution of a disgust response, which is likely related to the quickness of moral judgments and reactions. Their identification of the risks associated with pathogens and parasites presents a highly specific adaptive problem, and an evolved aversion is described as the fitness benefitting solution that spread throughout the species. Although brief, this is the beginning of a strong evolutionary case.

The authors make similar cases for the care/harm foundation. They explain the following about the care/harm foundation:

All mammals face the adaptive challenge of caring for vulnerable offspring for an extended period of time. Human children are unusually dependent, and for an unusually long time. [...] Rather, mammalian life has always been a competition in which females whose intuitive reactions to their children were optimized to detect signs of suffering, distress, or neediness raised more children to adulthood than did their less sensitive sisters. Whatever functional systems made it easy and automatic to connect perceptions of suffering with motivations to care, nurture, and protect are what we call the Care/harm foundation. (p. 67)

The authors state the following about the fairness/reciprocity foundation:

All social animals face recurrent opportunities to engage in nonzero-sum exchanges and relationships. Those whose minds are organized in advance of experience to be highly sensitive to evidence of cheating and cooperation, and to react with emotions that compel them to play “tit for tat” (Trivers, 1971), had an advantage over those who had to figure out their next move using their general intelligence. (p. 69)

In each of these examples we see specific adaptations identified in relation to adaptive problems and connected to moral foundations. This is a good beginning of a moral case.

The authors also reference phylogenetic evidence in describing the loyalty/betrayal (above described as ingroup/loyalty) foundation. For this they explain the following:

Chimpanzee troops compete with other troops for territory (Goodall, 1986); coalitions of chimps compete with other coalitions within troops for rank and power (de Waal, 1982). But when humans developed language, weapons, and tribal markers, such intergroup competition became far more decisive for survival. Individuals whose minds were organized in advance of experience to make it easy for them to form cohesive coalitions were more likely to be part of winning teams in such competitions. (p. 70)

They also acknowledge that there is some debate regarding group-level selection, but the relevant point for the current dissertation is to note that this is an example of a group of Moral Foundations theorists appropriately citing phylogenetic evidence in their effort to make an evolutionary case for a particular moral foundation. That said, this example also demonstrates where their use of evolutionary evidence begins to fall short. The argument about individual minds fails to sufficiently explain the adaptation they are describing. It is a gesture at what could become a strong evolutionary case but saying “individuals whose minds were organized” does not actually describe what uniquely human feature distinguished these individuals from their phylogenetic relatives or the fitness benefits that it provided them with. In other words, what was the specific characteristic that evolved? How did it specifically solve the selection pressure that gave rise to it?

This issue is repeated in what appear to be even more vague and poorly explained terms when the authors introduce the authority/subversion (also described as the authority/respect) foundation. They explain,

Many primates, including chimpanzees and bonobos, live in dominance hierarchies, and those whose minds are structured in advance of experience to navigate such hierarchies effectively and forge beneficial relationships upwards and downwards have an advantage over those who fail to perceive or react appropriately in these complex social interactions (de Waal, 1982; A. Fiske, 1991). The various modules that comprise the Authority/subversion foundation are often at work when people interact with and grant legitimacy to modern institutions such as law courts and police departments, and to bosses and leaders of many kinds. (p. 70)

Again, these authors appear to understand the relevance of evolutionary explanation in identifying the moral foundations, but their application is fairly thin. In this case, they use a trait identified in chimpanzees to explain a shared feature of human nature. They do not explain the fitness benefits that feature provided for human beings, or necessarily what that feature is beyond gesturing at the structure of modern (particularly Western) societies. They also fail to recognize that the authority/respect foundation is significantly different among humans than among chimpanzees (cf. Boehm, 2008). They have begun the work of developing an evolutionary explanation based on moral pluralism, but their descriptions often fall short of theoretical adequacy.

This issue may in part be related to how Moral Foundations theorists approach the task of identifying the moral foundations, and the criteria they require for the identification of each. Moral Foundations theorists identify five criteria required to establish a moral foundation. While these criteria have some relationship to the necessary criteria for evolutionary explanations, they also lack some of the basic features. Moral Foundations theorists identify the following criteria:

1. A common concern in third-party normative judgments
2. Automatic affective evaluations
3. Culturally widespread
4. Evidence of innate preparedness
5. Evolutionary model demonstrates adaptive advantage

The last three of these criteria fit closely with an evolutionary model. The criteria of being culturally widespread is essentially synonymous with cultural ubiquity. The idea of innate preparedness describes a feature that emerges naturally through development

and without explicit learning. And evidence of an adaptive advantage could, if interpreted and evaluated adequately, state the requirement of fitness benefits.

To say that a moral foundation must be identified with an automatic, affective evaluation does not conflict with an evolutionary perspective. To the degree that the moral foundations describe actual adaptations, there may well be automatic affective responses to situational cues. From an evolutionary perspective, such affective responses are part of the adaptation, not a criterion for identifying an adaptation. The emotional response cues attentiveness and possible behavioral responses, making it integral to the adaptation. In calling this a criterion for identifying adaptations, Moral Foundation Theorists have committed a category mistake that betrays the limits of their understanding of evolutionary science.

The biggest issue with this approach may be the idea that the first criteria for identification of a moral foundation is common concern in third-party normative judgment. Graham et al. (2013) describe this criterion by explaining the following:

One of the most significant steps in the evolution of morality may have occurred when human beings developed “shared intentionality” – the ability of multiple people to hold a shared mental representation of what they are trying to do together (Tomasello, Carpenter, Call, Behne, & Moll, 2005). [...] But when humans developed the capacity for shared intentionality, our capacity to recognize norms began to grow into a passion for enforcing them on each other (Boehm, 2012). Humans began to live in “moral matrices” – the “consensual hallucinations” that provide a common normative framework against which people can and do judge the actions of others, even when those actions have no direct implications for the self (Haidt, 2012). (p. 109)

The issue here is not that third-party judgment (or third-party punishment in much of the research) is not an evolved feature of human nature that is closely related to moral functioning. The issue here is in the theoretical approach these authors take to developing a theory of human morality. Third-party judgment is itself an adaptation, which means

that the authors approach the subject using criteria that fall on different levels of analysis. This is another category mistake from an evolutionary point of view and reveals a critical limitation of Moral Foundations Theory.

These confusions in criteria for identifying moral foundations further complicates the ability of Moral Foundations theorists to approach the question of why human beings are ethically minded. They agree with evolutionary thinkers that the answer to that question is that we evolved to be moral creatures. How we evolved to be moral creatures thus requires a more well-informed evolutionary approach, wherein the criteria for moral characteristics are based on the criteria for evolutionary scientists use to identify adaptations.

Summary. Moral Foundations theorists deserve significant credit for their initial steps towards answering the question of why human beings are ethically minded. They appear to grapple with the question itself, do more to explore and include evolutionary evidence, and expand the domain of morality. These authors also provide a relatively detailed account of the purity foundation, though much of that work had already been done prior to the publication of their theory (e.g., Tooby & Cosmides, 1997).

Where this work continues to fall short of adequate evolutionary claims is in the authors' failure to be fully systematic in their exploration of the other four foundations. In other words, they provide a slightly more useful framework for building an evolutionary case, but the connections continue to be missing. The authors don't make a clear case for the evolutionary pressures, they do not demonstrate a clear link between those pressures and the instinctive reactions, and they do not clearly relate those reactions to a theory of ethical mindedness.

That said, their contributions are clear. The most significant of these contributions is the expansion of the moral domain, and the evidence that questions of morality relate to a wider range of issues than previously presented. By understanding that not every moral question will relate to purity, loyalty, fairness, harm, or authority, we can begin building a more adequate case for ethical mindedness and the evolutionary process and pressures through which it emerged as a basic feature of human nature.

Concluding Points

I began this chapter with a warning about social scientists being led astray by a perfunctory start. I highlighted that issue because, at the core of this dissertation, is a deep concern about moral psychology research and theory built upon an inaccurate understanding of human nature.

I have concluded my review of moral psychology with Moral Foundations Theory because I believe these authors have the strongest theory of why ethical mindedness appears to be a basic feature of human nature. It is because of the strength of their work that I found it surprising, when reviewing this literature, to read one of the most egregious statements about human nature made by one of the most prominent authors from the Moral Foundations Theory perspective. In this, Haidt and Kesebir (2010) state that moral systems are, “sets of interlocking values, practices, institutions, and evolved psychological mechanisms that work together to suppress or regulate selfishness and make social life possible” (p. 800).

Why is this statement so egregiously inaccurate? Because it presents an unsubstantiated view of human beings as innately selfish and in need of regulation. It suggests that we are only moral when we overcome our fundamental amorality. This

claim is deeply at odds with the authors' ultimate point that human beings are programmed to develop morality. It also flies in the face of a preponderance of evidence for ethical mindedness, which defines us as naturally predisposed to be moral beings or concerned with moral functioning. After all, if human nature is fundamentally amoral, how does ethical mindedness ever get off the ground? Perhaps if adequate evolutionary evidence were present in the current moral psychology research, such seemingly obvious errors in thinking would be less likely. In other words, had these authors begun their work by grappling with the centrality of ethical mindedness to our basic nature, or with a sound theory of human nature to build upon, it would be difficult to make such claims about an essential amorality.

I bring this up to emphasize the importance of the current work and the significant consequences of building theoretical frameworks without adequate conceptualizations of human nature. If we do not account for the centrality of ethical mindedness as a basic feature of our nature, then even the most sound and rigorous theorists among us are at risk of producing work that is peppered with inaccurate and contradictory assumptions.

In reviewing the literature from moral psychology, I understand that there are likely exceptions to the concerns I have identified, and almost certainly alternative interpretations to my own. Those possibilities deserve attention and warrant discussion among the relevant theorists. I do not presume to be best-equipped for singlehandedly producing an integrated and comprehensive moral psychology.

What I can offer without qualification, however, is the statement that an adequate theory of human nature is not optional for these theorists – nowhere is that clearer than in this final example of from Haidt and Kesebir (2010) in Moral Foundations Theory. And

what I find missing across all three perspectives is clear: First, these authors, like many others, appear to have not yet come fully to terms with the centrality of ethical mindedness despite their ultimate focus on morality. Second, those who do attempt to explain or make cases for that centrality fall short in their efforts to build an adequate evolutionary case, often by simply avoiding the evolutionary question and offering only brute assertions about human moral functioning that provide no explanation for how or why an element of ethical mindedness came to exist. And finally, even those who do come closer to theoretical adequacy in their evolutionary arguments are not fully systematic or thorough, nor have they focused exclusively on a sound evolutionary case for ethical mindedness as a basic feature of human nature.

Despite these shortcomings, however, we also see significant contributions coming from each existing theoretical perspective. From the Social Cognitive Domain, we gain an understanding that morality and moral development appear to be universal across cultures and times, thus suggesting that this is a basic feature of human nature. From the moral identity perspective, we gain a clearer understanding that the centrality of morality to identity and behavior varies significantly across individuals, suggesting that morality is subject to individual differences and contextual experiences. And from the Moral Foundations Theory, we understand that morality likely refers to domains far wider than that of interpersonal interaction or relationships and is broader than helping and harming.

While a comprehensive evolutionary case for a concept as abstract and complex as ethical mindedness is beyond the scope of this dissertation, the next chapter will demonstrate how evolutionary theory can be used to begin building that case. This will

consist of a demonstration of how a widely recognized adaptation—cooperation—is inextricably tied to the ethical issue of fairness. The first step in that demonstration will be to apply the evolutionary criteria for assessing adaptations to cooperation in order to illustrate that cooperation is a species-typical feature of human nature. Chapter 4 will elaborate on this argument by exploring the evolution of cooperation in order to connect it with ethical mindedness. This theoretical perspective will prepare us for understanding how the existing observations and contributions from moral psychologists fit into a more adequate theory of ethically minded human nature.

CHAPTER 3: A PARADIGM ACCOUNT OF COOPERATION AS AN ELEMENT OF ETHICAL MINDEDNESS

The ultimate purpose of this dissertation is to address two basic questions: Is ethical mindedness a demonstrable characteristic of human nature? And what might an adequate theory of ethical mindedness look like?

In Chapter 1, I presented and explained the basic criteria for sound evolutionary argument, given that evolutionary theory provides us with the most long-standing and comprehensive scientific direction for making claims about human nature. In Chapter 2, I provided examples of how moral psychologists have attempted such claims or made reference to the ontology of ethical mindedness. Although these authors offer significant contributions to a larger theory of moral functioning, I have attempted to demonstrate how moral psychologists fall short of providing an adequate evolutionary argument for the emergence of ethical mindedness as a species-typical feature of human nature. As the subdiscipline of psychology that is arguably most responsible for demonstrating this feature, should there be one, this appears to be an important gap in the development of theoretical perspectives on moral functioning. Without a species-typical feature of ethical mindedness, it is unclear how the universal morality these scholars attempt to explain can get off the ground.

The current dissertation is ultimately only the beginning of what would constitute a full evolutionary case for ethical mindedness. This is in large part because the concept of ethical mindedness is itself so broad, abstract, and inclusive of so many evolved traits that a comprehensive review of a wide range of relevant evolved characteristics is simply beyond the scope of this dissertation. Some of these include not only the connection between cooperation, trust, and the instinct or interest in fairness, but also the

complicated relationship between group belonging and shared identity, or hierarchy and permissions and prohibitions. What is within the scope, however, is a clear demonstration of what such a review may look like, or of how the framework of evolutionary theory can be used to identify and understand basic features of human nature.

In order to complete an adequate examination of evolutionary evidence, I will explore cooperation in this chapter for two purposes: First, as an example of how the criteria for identifying adaptations can be applied; and, second, to begin answering the questions of whether ethical mindedness is a species-typical feature of human nature, and what an adequate theory of ethical mindedness might look like. By doing so, I will demonstrate both one feature of ethical mindedness as an example of how other relevant features can be identified and discussed using the framework presented in Chapter 1. In Chapter 4, I will connect cooperation to the presence of ethical mindedness through the related adaptations of cheater detection, indirect reciprocity, and third party punishment. While there are other features and domains that can be discussed, cooperation provides a seamless connection between an evolved characteristic and ethical mindedness as an unmistakably evolved feature of human nature.

Cooperation thus offers an excellent opportunity to examine a widely-studied concept for which the existing evidence illustrates the basic nature and circumstances from which ethical mindedness emerged. Cooperation can be linked to the domains identified in existing theories of moral psychology while providing us with evidence that meets the necessary criteria for evolutionary argument. By honing in on studies related to matters of cooperation, we can use evolutionary reasoning to demonstrate a species-typical feature that indicates the inevitability of ethical mindedness. I will begin this

chapter with a brief discussion of cooperation, followed by a more detailed examination of the relevant evolutionary evidence according to the criteria outlined in Chapter 1, and finally a section that explores the links between cooperation, trust, and fairness in order to demonstrate a seamless connection between cooperation and ethical mindedness.

Cooperation

Human societies “are based on a detailed division of labor and cooperation between genetically unrelated individuals in large groups” (Fehr & Fischbacher, 2003, p. 795). This pervasiveness of cooperation has been identified not only in modern societies characterized by individuals, organizations, and large nation states, but it is true of hunter-gatherers who “typically have dense networks of exchange relations and practice sophisticated forms of food-sharing, cooperative hunting, and collective warfare” (Fehr & Fischbacher, 2003, p. 795). We find evidence of these sophisticated practices of cooperation despite evidence that the majority of individuals in hunter-gatherer bands (generally 25-30 individuals) are genetically unrelated (Hill et al., 2011). Cooperation has thus been identified as a species typical adaptation with features of “antiquity, universality, and cross-cultural elaboration” (Cosmides & Tooby, 1992, p. 206). Focusing on the specific form of cooperation of “social exchange” to describe cooperation, Cosmides and Tooby (1992) explain the following:

Wherever human beings live, their cultural forms and social life are infused with social exchange relations. Such relations appear in an enormous range of difference guises, both simple and highly elaborated, implicit and explicit, deferred and simultaneous, practical and symbolic. The magnitude, variety, and complexity of our social exchange relations are among the most distinctive features of social life, and differentiate us strongly from all other animal species. (p. 206)

Colloquially, we may often think of cooperation as a more explicit and effortful exchange between two people, and this conceptualization can make it difficult to observe the overwhelming prevalence of cooperation in our societies. That issue is easily resolved by considering the inherent cooperation required by simple tasks like shared routines, driving in traffic, monetary transactions, or the roles we play in larger professional, social, or governmental institutions (Fowers, 2015). Not only are our day-to-day lives infused by cooperative efforts and arrangements, but our basic survival depends upon cooperation. Driving is an excellent example of this. We tend to follow the basic rules of traffic and doing so ensures that we are generally able to drive safely on a daily basis. Following these basic rules is also generally a communal practice which typically and mutually ensures the safety of everyone involved.

I turn now to the basic criteria for evolutionary claims as identified in Chapter 1. I will use those criteria to structure the current section and argue that cooperation appears to be a demonstrable feature of human nature. I have organized subsections according to the criteria for evolutionary argument and will include research from relevant disciplines. This chapter uses cooperation to begin the process of building a systematic, cogent argument for ethical mindedness that both supports my claim that it is a basic, evolved feature of human nature and begins answering the question of what an adequate theory of ethical mindedness might look like.

Fitness benefits. Fowers (2015) identifies three primary fitness benefits of cooperation. The first of these is that cooperation can increase the quantity of a particular good. Hunters who cooperate within a band, for example, are more likely to succeed in securing larger prey. Speaking more broadly, food sharing in general provides a more

consistent diet than that which can be obtained by an individual or small group. The second fitness benefit provided through cooperation occurs when cooperators share or trade different goods (as opposed to increasing the quantity of a single good). This benefit refers to acts of sharing food across diverse food-gathering methods in order to increase an individual's variety of available foods. Finally, Fowers (2015) identifies shared goods as those that are only made possible through cooperative activities. He explains that "there are many shared goods, including playing games and sports, singing in a choir, and scholarly conferences. Government, laws, and public security are also shared goods that are only possible through cooperation" (Fowers, 2015, p. 164). This idea of a shared good is critical to my ultimate conclusions about cooperation. When we make the mistake of viewing cooperation as an individual's action or choice, what emerges is a view of cooperation that is fundamentally individualistic, with the individual responding only to egocentric interests and external contingencies. A shared good, by contrast, is collective. Shared goods can only be achieved with the participation of the community, and thus can only be had together with others (Fowers, 2015). This is an important distinction from aims that can be possessed or pursued individually (e.g., knowledge, wealth or fame) because it frames cooperation not only in terms of individual interests and actions, but also in terms of a widely present communal focus in human activity (Clark & Aragon, 2013; Fowers, 2015). Shared goods can only exist collectively, and thus require a focus on the group to fully understand. This collectivity of cooperation will be critically important in its ultimate connection to ethical mindedness in Chapter 4.

The two primary adaptive problems that are important for understanding cooperation in terms of food sharing are the inconsistency of food availability and the

lengthy juvenile period for human beings. The inconsistency of food availability emerged as a significant adaptive problem when our ancestors moved toward bipedalism and relocated into the African savanna where food sources are not as plentiful and are more scattered than in the forest. I will explore these issues in greater detail below, when discussing the lock and key fit between adaptive problem and adaptation, but it is important note here that this locates the emergence of food sharing (from which cooperation develops) in a particular EEA. Following this move to the savanna, food sharing emerged around hunting, which is where we see this behavior differentiating from the type of sharing observed in other species. In this, humans share food beyond just sharing directly with offspring. Human hunters bring large game back —wholly or largely intact – and share it with the entire band. In this sense, we see the adaptation of food sharing as a cooperative act that emerged when our human ancestors were faced with the problem of food inconsistency and the need for consistent sources of food.

Returning to the specific fitness benefits, however, I begin by looking at an example of food sharing in an alternative species. This is an example of what has been termed direct reciprocity. Direct reciprocity is described as behavior in which “individuals provide benefits to one another reciprocally” (Fowers, 2015, p. 166). In the case of direct reciprocity, it is important to understand that the benefits do not need to be immediate but can be received over time. We see this in research on vampire bats, who cannot survive two consecutive nights without feeding (Wilkinson, 1988). Variance in whether individual bats are able to feed on a given night creates a significant selection pressure, and food sharing among them has been well documented. Wilkinson (1988, 1990) observed that the strongest predictor of food sharing was whether the recipient of

shared food had in the past shared food with the sharer. In this sense, we see not only the benefits observed over time, but we see the fitness-enhancing nature of this behavior. In other words, by sharing one's food with an individual who has shared before, the sharer not only saves the recipient's life (at minimal cost to themselves), but also ensures that an individual who might one day save their own life survives in order to do so. A species whose survival depends on frequent feeding, but whose likelihood of finding food is variable, would have been unlikely to survive without an instinct to share and the capacity to track that form of cooperation among those around him or her.

What we can observe in the dynamic and circumstances of vampire bats is an adaptive problem (variable food supply and a need for frequent feeding to survive) and solution (food sharing) that can be viewed as an evolution-based instinct to cooperate in ways that relate to the general concept of ethical mindedness by preparing the individual for cooperative interpersonal behavior and a concern about the wellbeing of others. In order to make this connection to ethical mindedness and adequately address the evolutionary criterion of fitness benefits, however, we have to look at all three of the fitness benefits for cooperation in humans. Where in bats, such food sharing appears to be strictly *quid pro quo*, in humans it appears to be a shared activity.

If we look at simply an increase in quantity as a benefit of cooperation, the findings of Kaplan and Hill (1985) are compelling. These authors study the food acquisition among Ache foragers and find that individually hunted game is shared just as much as collectively hunted game. They acknowledge the possibility that "cooperative acquisition is a consequence rather than a cause of food sharing" (p. 236) but recognize the overwhelming evidence that the increased amount of hunted game and the

inconsistency of successful hunts are the primary predictors for extensive food sharing among predatory animals, including human foragers. The key point here is that when a successful hunt results in more than a few pounds of meat, food sharing is the norm among these groups. It is an automatic and pervasive group behavior, which leads to the communal nature of cooperation that I emphasize in my connection to ethical mindedness.

This focus on inconsistency of food resources could be interpreted similarly to the *quid pro quo* of food sharing observed in vampire bats. In these bats, the fitness benefits of food sharing focused primarily on increasing quantity, or otherwise ensuring the consistency of an individual's diet. In humans, we observe cooperative food sharing as not only contributing to the quantity and consistency of available food, but also to the variety.

Kaplan, Hill, Lancaster, and Hurtado (2000) note two key differences between the chimpanzee and human diet. First, human hunter-gatherers focus primarily on vertebrate meat, which makes up between 30 and 80 percent of their diets. Chimpanzees spend only two percent of their time hunting. The significant commitment to hunting observed among human males thus represents a dramatic division of labor that requires a high level of cooperation between male hunters and females, and the latter have been observed to focus on plant food gathering. The important development here is how the intersexual cooperation and food economy occurs when male hunting provisions women of child-bearing age and children. Such provisioning enables allomaternal care, or when young are cared for by individuals who aren't their mother, which has been observed only in humans, marmosets, and tamarins (Burkart, Hrdy, & van Schaik, 2009).

This is where a second difference in the diets of humans and chimpanzees has been observed. Chimpanzees primarily forage ripe fruits, whereas hunter-gatherer females engage in complex, extractive foraging of roots. These roots require more skill and knowledge to obtain but offer far greater nutrient density. Kaplan et al. (2000) note that the human focus on more challenging but nutrient-rich diets has made the extended juvenile developmental period necessary to learn the relevant skills for both this foraging behavior as well as the complicated practice of hunting. Such a long juvenile period is made possible by the intersexual cooperation of male and female humans.

Kaplan et al. (2000) study diet, intelligence, and longevity of human life. In this work, the authors “propose that the long human life span co-evolved with lengthening of the juvenile period, increased brain capacities for information processing and storage, and intergenerational resource flows, all as a result of an important dietary shift” (Kaplan et al., 2000, p. 156). This is a much larger claim about human nature, and one that connects more directly to the idea of cooperation as a way of life and an evolved feature of human nature. I bring this problem up because while food sharing offers a lock-and-key fit to the adaptive problem of food inconsistency, and food cooperation is likely the beginning of cooperation, cooperation has been generalized to every aspect of human life. This generalization is a critical point, and one part of the process through which that generalization occurred may be related to a second adaptive problem for cooperation: the lengthy juvenile period of human beings.

Food sharing is often considered a key element of the transition from apes to humans. For example, there is some evidence suggesting that communal butchering sites and shared tools for food processing existed among hominin populations at least 2.5

million years ago (Ambrose, 2001; Klein, 2000). The uniqueness and pervasiveness of human food sharing and cooperation appear inseparable, with some authors suggesting that more generalized cooperation may follow food sharing due to the significant fitness benefits that were incurred (Kaplan et al. 2000). What this evidence demonstrates, however, illustrates a far bigger point about cooperation than simply the fitness benefits supporting evidence that it is an adaptive feature of human nature. Not only do these findings support the idea that cooperation is a basic feature of human nature, but they begin to indicate the scope of human cooperation and distinguish it from the basic quid pro quo that can be observed among other species.

The key points for the current section are the clear fitness benefits of food sharing, how food sharing can be viewed as the beginning of cooperation, and how cooperation has been generalized and become characteristic across the human species. In terms of this dissertation as a whole, however, it is important to understand that humans evolved to cooperate in far greater extent and depth than our primate relatives. Cooperation in this light is not simply a behavior, but a defining element of our intensely social and communal nature. Once we begin viewing this pervasive cooperation from this communal perspective, we are again directed toward the idea of group-level functioning because the foraging band cooperates as a band, not just between individuals. I will expand upon this focus on the group when I discuss third party punishment in Chapter 4, but the key point that leads us there is the idea of cooperation as such a basic feature of human groups that it can ultimately be taken for granted by group members, indicating a level of trust in the group, and thus a basic investment in fairness. This idea of our investment in fairness is what will ultimately lead us to one of ethical mindedness.

Lock-and-key fit between adaptation and adaptive problem. Turning back to another specific criterion for evolutionary argument, however, I focus now on the lock-and-key criterion. This criterion makes clear that identification of a specific adaptation requires a tight fit between the identified adaptation and a longstanding challenge to survival and reproduction (the adaptive problem it was selected to solve) (Cosmides & Tooby, 1997). The lock-and-key fit is thus defined by the adaptive problem. In the case of cooperation, and particularly cooperation as it emerged from food sharing, this problem was food inconsistency. Our ancestors faced with the problem of less-condensed food resources and access to large game that provided an excess of food sources in one moment, with a potential lack thereof in the next. Food sharing solves this problem in a lock-and-key fashion and sharing hunting proceeds is a specifically strong fit because food inconsistency is likely the only reason why one would engage in such widespread acts of food sharing.

One way of looking at how cooperation evolves from this lock-and-key fit of food sharing is in the findings from Hammond and Axelrod's (2006) study of cooperation and in-group favoritism. These authors use a game theory simulation study and find that "a predisposition to favor in-groups can be easily triggered by even arbitrary group distinctions and that preferential cooperation within groups occurs even when it is individually costly" (Hammond & Axelrod, 2006, p. 926). They do so by searching for a very simple evolutionary model of cooperation and competition (Hammond & Axelrod, 2006).

Hammond and Axelrod (2006) used a simple, one-move prisoner's dilemma game in which actors had only three traits: the ability to distinguish another agent's group

membership among four groups; the ability to distinguish whether that agent cooperated or defected when meeting an ingroup member; and the ability to distinguish whether that agent cooperated or defected when meeting an outgroup member. The strategy of cooperating with ingroup agents and defecting with outgroup members (the ingroup favoritism strategy) was one of four possible approaches. The highest immediate payoff was to defect when the other agent cooperates. The authors included conservative parameters, which they made even more conservative when retesting to confirm their results. These were a 12 percent reproduction rate, random “choice” of exchange partner for each round, and a 10 percent mortality rate. The act of cooperation lowered the reproduction rate by 1 percent while receiving cooperation raised the reproduction rate by 3 percent.

After running the simulation ten times for 2000 iterations, the authors found that 76 percent of agents utilized the strategy of cooperation with ingroup members (Hammond & Axelrod, 2006). This cooperation proved to be a dominant strategy in part because agents appeared to congregate in ingroups through migration while the ingroup favoritism caused groups to grow faster despite “free riders” who took advantage of cooperation within the group.

So how does this evidence suggest a lock-and-key criterion for the evolution of cooperation? The primary conclusion that we are interested from these studies is the finding that cooperating with cooperators is extremely adaptive; so much so that it quickly becomes dominant within a group. In other words, when a group-living species hits on the opportunities afforded through cooperation, this is an adaptation that is very likely to be adopted. In Hammond and Axelrod’s (2006) study, what we see is evidence

that cooperation can evolve as a species-typical feature over time. The clear lock-and-key fit between the problem of food inconsistency and the emergence of food sharing helps to clarify how the practice of cooperation propagated throughout the species with relative ease. What this lock-and-key fit does not provide us with, however, is adequately verified evidence that this particular adaptation emerged specific to the adaptive problem I have identified, and not as a byproduct of an alternative adaptation. Research demonstrating how this criterion may be met is explored below.

Cannot be explained by any other adaptation. The criterion that an adaptation not be better explained by any alternative adaptation or adaptive problem is a difficult one to demonstrate insofar as it presents the unique challenge of essentially proving a negative. This criterion is also closely related to the lock-and-key fit, and one way around the challenge of meeting it may lie in narrowing the level of specificity.

One way of narrowing the level of specificity when looking at cooperation is examining the relationship between human cooperation and kin selection. I have suggested that food sharing emerged because it helped solve the problem of food inconsistency. It could also be argued that food sharing occurred because it is a way to care for one's kin and promote common genes. With this claim, we can see how we would have two relatively parallel explanations for the emergence of food sharing, or two ways to explain cooperation in terms of food sharing. Thus, to demonstrate how this criterion can be addressed, I will show how an alternative explanation can be addressed.

Kin selection refers to the propensity to help those who are genetically most closely related. While this could be interpreted as altruistic for the individual who may pay a cost to provide such help, from a genetic perspective, kin selection facilitates the

propagation of one's own genes by protecting the individuals with whom those genes are shared. Fowers (2015) explains that, in this sense, "kin selection is a(n) (implicitly) collective effort to pass shared genes on to the next generation" (p. 165). It should be noted that kin selection is not considered an adaptation, but that it operates at a higher level and can be used to explain adaptations. It is often used to explain cooperation in eusocial insects. For example, why would worker bees, who cannot reproduce themselves, die for the sake of their hive? Although Darwin was never able to fully explain this phenomenon, the answer that ultimately emerged was kin selection. In other words, worker bees will die for their hive because the queen is their sister, and all of the babies are therefore their half brothers and sisters (Hamilton, 1964). The current question is whether kin selection offers us a similar explanation for food sharing. That is, does food sharing emerge simply from a drive to protect one's kin?

In a study that I examine below, Olson and Spelke (2008) found that young children allocate more resources to family and friends than to strangers, but the authors did not find a significant difference in the resources that are allocated between family and friends. In this sense, we see an ingroup favoritism that may relate to the findings explored above in the evidence for lock-and-key fit, but not sufficient evidence that such findings can be explained simply by kin selection.

More specifically than this, however, is evidence from Hill et al.'s (2011) study of the social structure of hunter gatherer societies. In one example, these authors found that "in the Ache we estimate the mean genetic coefficient of relatedness (Hamilton's r) between adults in 58 precontact bands to be only 0.054 (note that the r for siblings is 0.5, suggesting very little shared genetic material between the adults among the Ache). This

agrees with the Ache informants who reported that during the precontact period they often lived with people described as ‘friends, not relatives’” (Hill et al., 2011, p. 1288).

What this means in terms of the current section, however, is that it appears questionable to view cooperation as a byproduct of kin selection because it appears to have emerged within groups that were composed primarily of non-kin. In other words, it appears that food sharing and cooperation cannot be alternatively explained by kin selection because we do not see a strong preference for biologically-related kin driving these practices.

A further way of demonstrating this criterion would be to explore the possibility that food sharing could be explained by prestige or hierarchy, such as the hunter who shares game with the group gaining power in the group, as opposed to a means of addressing food inconsistency. This too, however, is easily challenged as hunters virtually never take credit for what they provide, and research suggests a strong egalitarian hierarchy ethic in hunter gatherer groups (Henrich & Gil-White, 2001). Power is never granted for hunting, strength, beauty, bossiness, or other traits (Henrich & Gil-White, 2001). Thus, we find again that the only viable explanation for food sharing appears to be its adaptive fit with the problem of food inconsistency.

Presence/absence in phylogenetic relatives. The identification of adaptive features in phylogenetic relatives is based on the notion that a species’ characteristics are inherited from common descent, or the idea that “closely related species will share many adaptations that can be traced to their common ancestor” (Fowers, 2015, p. 54). By mapping out species based on their shared characteristics and tracing those characteristics to their common ancestors, we are able to develop an evolutionary timeline that shows us

with reasonable certainty how specific features came to be basic parts of a species' nature. When examining cooperative behavior and looking for those features that demonstrate the groundwork for ethical mindedness, there is significant evidence of similarities and differences in studies of our close phylogenetic relatives.

A particularly interesting example of this research uses a matched experiment to compare the cooperative behaviors of chimpanzees to those used by human children. Melis, Hare, and Tomasello (2006) conducted experiments on eight semi-free-range chimpanzees. In these, the authors investigated whether chimpanzees “(i) know when collaboration is necessary and (ii) choose the more effective of two potential collaborators, based on previous experience with each of them” (Melis et al., 2006, p. 1298). The authors conducted multiple trials in which chimpanzees were required to simultaneously pull two separate ropes in order to obtain food that was placed on a board behind a fence. They found that when the food was placed in two separate piles, one for each chimpanzee, the animals collaborated successfully; but when the food was placed in a single pile in the center of the platform, the dominant chimpanzee would often monopolize all of the food after collaborating for access. When this happened, the cooperation would collapse over multiple trials as the subordinate chimpanzee eventually ceased to collaborate.

Warneken, Lohse, Melis, and Tomasello (2011) conducted a similar experiment with 3-year-old children and conditions termed “dispersed” and “clumped” to identify how a pile of gummy bears or stickers was placed on the board in front of them (p. 268). In stark contrast to the chimpanzees, these authors found that children split the rewards almost equally regardless of whether they were clumped or dispersed on the board. As

such, the children demonstrated well-founded trust that their efforts would be rewarded, and they collaborated successfully across conditions. In a surprising finding, the authors also found that children in the clumped condition were more likely to share equally when the reward was gummy bears than when it was stickers. While a purely speculative interpretation, this is an interesting discovery given the relationship between food sharing and cooperative behavior.

In a third study, Hamann, Warneken, Greenberg, and Tomasello (2011) found that children around three years old share more equitably following collaborative efforts than after what they term “windfall” or “parallel-work” situations (p. 328). Chimpanzees, in contrast, do not share regardless of collaboration. The authors analyzed the difference in how resources that were initially distributed unequally are shared based on three conditions. In the first condition, or the collaboration condition, two ends of a rope had to be pulled simultaneously in order to move a board and gain access to an unequally distributed reward (reward distributed in a 3:1 ratio for each participant). In the second condition, or the “no-work” or “windfall” condition, access was given without any effort or work (reward still distributed in a 3:1 ratio). In the third or “parallel-work” condition, two small blocks were pulled toward the access point but did not require cooperation or collaborative effort to retrieve (again with a 3:1 distribution).

In the windfall condition, in which participants walk into a room and find three rewards at one end with only one at the other, neither the children nor the chimpanzees tended to share the rewards equally. In the parallel-work condition, in which participants simply pulled their own ropes to gain access to a reward (one rope leading to three and the other to one), chimpanzees almost never shared but roughly one third of the “lucky”

children shared their rewards to create an equal distribution. While this third of children who shared their good fortune is noteworthy and extremely relevant to the current dissertation, it was in the collaboration condition that these authors discovered the most striking difference between children and chimpanzees. In this, they found that following a cooperative and equal effort, chimpanzees still tended not to share while the “lucky” child appeared to recognize the equal effort and shared in almost 80% of the trials.

What we see here is exactly the twofold pattern that phylogenetic analysis should produce: A common trait that we share with a close phylogenetic relative, as well as one that we appear not to share. That is, both chimpanzees and humans were easily able to engage in cooperative efforts in order to gain access to a reward. This has been demonstrated across multiple studies and conditions. Where we differ, however, is in the human instinct to share. Not only do we appear to trust in that effort (as evidenced by the consistency of collaboration despite clumped distribution that was typical in children but fell apart in chimpanzees), but human trust appears well-founded. While chimpanzee behavior can be explained as having the capacity to cooperate while nonetheless using dominance to solve a problem of food competition and distribution, humans appear incentivized to cooperate and maintain cooperation through sharing. This apparent and apparently well-founded trust suggests a building block for ethical mindedness in our concern about fairness and equity, and the research demonstrates one way in which phylogenetic evidence can be used to identify species-typical traits by examining their presence or absence in close phylogenetic relatives.

Develops naturally and without explicit learning. Building on this phylogenetic evidence and research on how young children engage with issues of cooperation and

sharing, the next criteria for adequate evolutionary argument is that the specific trait appears to develop naturally and without explicit learning. In order to claim that a trait is hard-wired in the human condition, we must be able to demonstrate that it emerges naturally over the course of development. As I specified in Chapter 1, evolutionary psychologists insist that any feature of human nature requires an environment that allows that feature to properly emerge or be expressed. I reiterate that here to emphasize that this criterion is not a claim that a feature will express itself under any developmental circumstances. This criterion simply states that provided typical environmental input, a certain feature will emerge over the course of development and without needing to be expressly taught to a growing child.

Provided a typical environment, however, the questions remain of how we see building blocks of ethical mindedness emerge naturally in young children, and what evidence can we turn to in order to meet this criterion for making an evolutionary argument about cooperation and ethical mindedness as species-typical traits? For this we look primarily to the work of developmental psychologists and research examining early attention to matters of harm, benefit, and early signs of cooperative behavior.

Hamlin, Wynn, and Bloom (2007) examined 6- and 10-month old infants' intuitions about social interactions. In the first of these approaches, infants were exposed to a form of puppet show in which a small "climber" (made of wood with large eyes glued onto it) repeatedly attempted to climb a hill. On the third attempt, the climber was either helped by a character who pushed it from behind, or it was hindered by a character who pushed it down. The infants were exposed to alternating trials of helping or hindering until their measured looking time indicated that they had processed the events.

The infants were then exposed to a test phase in which they were encouraged to choose between the helper and the hinderer. The infants overwhelmingly chose the helper (14 out of 16 10-month-olds and 12 out of 12 6-month-olds), indicating a strong and apparently unlearned preference for the helping behavior. An alternative condition ruled out the possibility that infants simply prefer pusher-uppers to pusher-downers. In this, a separate set of infants was shown a similar display in which an inanimate object (the piece of wood without any eyes) was alternately pushed up and pushed down the hill by a pusher-upper and a pusher-downer (the same helper and hinderer from the previous condition). In this case, the infants showed no preference and reached for each at a similar rate when encouraged to choose between them.

In a study of even younger infants, Hamlin, Wynn, and Bloom (2010) employ a measure of “preferential looking” to measure visual attention of 3-month-olds and determine the preference of children who are too young to indicate preference by reaching. In this study, infants were again shown a character who tried but failed to climb a hill and was then alternately helped or hindered. The infants again demonstrated a strong preference for the helper (mean preferential looking to helper was 13.12 seconds versus 6.22 for the hinderer) and, again, showed no preference between the pusher-upper versus pusher-downer of an inanimate object (with mean preferential looking of 7.89 seconds versus 10.93 seconds respectively). The authors interpret this evidence suggesting that “three-month-olds’ preference for helpers over hinderers appears to reflect a specifically social evaluation” (Hamlin, Wynn, & Bloom, 2010, p. 926).

Finally, Hamlin, Wynn, Bloom, and Mahajan (2011) studied how infants and toddlers react to antisocial others. The authors introduce this explaining that, “although

adults generally prefer helpful behaviors and those who perform them, there are situations (in particular, when the target of an action is disliked) in which overt antisocial acts are seen as appropriate, and those who perform them are viewed positively. The current studies explore the developmental origins of this capacity for selective social evaluation” (Hamlin et al., 2011, p. 19331). In these studies, the authors find that while five-month-old-infants indicated uniform preference for those who act positively toward others regardless of the behavior of the target, eight-month-old infants demonstrated a more selective preference for both those who acted positively toward prosocial individuals as well as for those who acted negatively toward antisocial individuals. They view these findings as evidence that the kind of pervasive social judgment and action observed in adults emerges early in developing children and without explicit teaching or encouragement from others.

When we examine these findings in the context of cooperation, however, what we find here appears to be something like a fairness instinct. While this evidence sets an important developmental baseline for the examination of third party punishment that I discuss below, there is additional research that looks more specifically at cooperation.

In order to examine cooperation more specifically, several sets of authors have conducted studies with young toddlers. Warneken, Chen, and Tomasello (2006) conducted a study with 18-24 month old children and young chimpanzees interacting in four different cooperative activities with a human adult partner. In each experiment, the adult stopped participating and the authors observed the subsequent behavior of the children and chimpanzees. While every one of the child participants responded by attempting to reengage their adult partner (either with language or efforts to physically

remind them how to participate), none of the chimpanzees ever made such an attempt. The authors interpreted these findings as evidence for a “uniquely human form of cooperative activity involving shared intentionality that emerges in the second year of life” (Warneken, Chen, & Tomasello, 2006, p. 640). The key here is to understand that although the chimpanzees appear to enjoy the game, or will continue playing until the adult stops, they interpret the adult stopping as simply the end of the game. The human children, by contrast, appear uniquely motivated to engage in this type of cooperative play, and they make clear efforts to reengage their counterpart.

In a study of such cooperative behavior or preferences among slightly older children, Olson and Spelke (2007) examine the three pillars of mature cooperative behavior: preferentially sharing resources with close relations (kin selection), people who have shared in the past (reciprocity), and those who have shared with others (indirect reciprocity). While these tendencies are well documented in adults, the authors explore them in the behavior of 3.5 year old children.

In these studies, children were first introduced to a doll who possessed desirable resources, plus six additional dolls who were described as either siblings, friends, or strangers (Olson & Spelke, 2007). The children were then asked to help the first doll distribute its resources among the other dolls. In this study, children distributed significantly more resources to siblings and friends than to strangers.

In a second variation of this study, the children were asked to help a doll distribute its resources among dolls who had previously shared directly with the original doll (direct reciprocity), or those who had shared with other dolls (indirect reciprocity). This exploration found that children distributed more resources to the dolls who had

directly shared with the original doll than those who had shared indirectly. Finally, a third variation compared the child-led distribution to indirect sharers and non-sharers and found that more resources were distributed to those who had given indirectly than those who had not shared at all.

Taken together, this research suggests several relevant findings. First, pre-language infants appear to be both drawn to helpers and cooperators and averse to those who hinder or harm others (even when those others are only loosely representative of human figures) without the explicit encouragement of caregivers. These preferences may be the prerequisite for what we observe during the second year of life as an explicit form of cooperative activity and a motivation to engage in both problem-solving and social games. Finally, in the third year of life, human children can be observed distributing resources in ways that specifically reward cooperation and sharing. As I mentioned above, I view these findings as indicative of a strong developmental base for third party punishment and an investment in fairness, which I will discuss in greater detail when connecting cooperation to the presence of ethical mindedness. Second, this developmental research demonstrates a specific inclination toward cooperation that appears to develop naturally and prior to exposure to complex networks of cooperation or significant input from caregivers.

I briefly noted when introducing the idea of these criteria that while I have here discussed them in more distinct terms, there is nonetheless significant overlap between them. One example of this would be the research I have included in this section that also includes evidence related to chimpanzees and the presence or absence of cooperative behaviors in children as compared to close phylogenetic relatives. Another example of

this overlap would be a comparison of how cooperation emerges developmentally across cultures.

Some work has been done in this area as well, and while an extensive review of this evidence is beyond the scope of this dissertation, findings suggest that young, preschool aged children, despite being less well-represented in the relevant literature, demonstrate little difference in their sense of fairness or relevant cooperative behaviors across cultures (Blake et al., 2015; Corbit, McAuliffe, Callaghan, Blake, & Warneken, 2017). This similarity has been shown to shift, however, as school-aged children appear to begin differing across cultures and can be presumed to respond increasingly to the social norms of their respective cultures (Zeidler, Herrmann, Haun, & Tomasello, 2016; Rochat et al., 2009; House et al., 2013). One potential interpretation of these findings that is particularly relevant to the current focus is that young children demonstrate a more basic, instinctive focus and attunement while in older children we begin to see the effects of an evolved instinct to learn cultural norms of cooperation and fairness.

Ubiquitous across cultures. I turn now to examining the criteria of ubiquity across cultures. I dedicated significant time to qualifying and explaining this criterion in Chapter 1 and did so in large part because it is one that presents a particular challenge to the field and to evolutionary argument. In order to claim ubiquity across cultures, and to do so both accurately and without risking the erasure of cultural diversity, the level of specificity must be one that can reasonably be observed in the context of that diversity. Cooperation provides us with one opportunity to focus on that ubiquity.

When I described cooperation above, I used examples such as driving in traffic to highlight the pervasiveness of cooperation in modern societies. While that particular

example may easily provide us with evidence of ubiquity across industrialized societies, it does not necessarily offer us a clear evolutionary context or evidence of ubiquity across periods of time and more widely diverse societies.

One way that evolutionary psychologists address this issue is by turning to the research of anthropologists examining hunter-gatherer cultures. These cultures are of particular interest because they are viewed as the closest proxy we have for the ancestral environment. While surviving hunter gatherer societies cannot be seen as representative of our ancestors for many reasons, they are the closest proxy we have for what those groups may have looked like. There is also significant variability between such surviving groups, so it is important to note that hunting for one group does not necessarily reflect hunting for another (some survive primarily on whales or sea life, for example, while others may hunt large game in savanna or forest environments). While this complicates the interpretation of this research, it also bolsters its strength when common practices are observed across such societies.

In one example, Hill (2002) studies cooperative food acquisition by Ache foragers. The author does so in part to address what he views as a gap in data on foraging peoples that is necessary for determining the evolution of cooperative behaviors. In order to examine cooperative behavior, Hill coded data from field notes and tapes including a total of 21,315 minutes following 50 men and 6,173 minutes following 20 women. After identifying cooperative events, the author totaled the amount of time spent during each and established a sum for each day. He defined cooperation as “expending time and energy in an activity that seems mainly designed to increase the food return rate of another individual” and clarifies that in some cases these behaviors might increase the

foraging gain rate of both the cooperators and recipient, though these latter cases were coded separately (Hill, 2002, p. 113). The author identified an extremely wide range of cooperative activities. These included the following:

[...] cuts a trail for others to follow; makes a bridge for others to cross a river; carries another's child; climbs a tree to flush a monkey for another hunter; allows another to shoot at prey when ego has first (best) shot; allows another to dig out an armadillo or extract honey or larva when ego encountered it; yells whereabouts of escaping prey; calls the location of a resource for another individual to exploit while ego continues searching; calls another to come to a pursuit of a peccary, paca, monkey, or coati; waits for others to join a pursuit, thus lowering own return rate; tracks peccaries when ego has no arrows (for other men to kill); carries game shot by another hunter; climbs fruit trees to knock down fruit for others to collect; cuts down palms (for others to take heart or fiber); opens a "window" in a tree to test for palm starch (for others to come take); carries the palm fiber others have collected; cuts down fruit trees for others to collect the fruit; brings a bow, arrow, ax, or other tool to another in a pursuit; spends time instructing another on how to acquire a resource; lends bow or ax to another when it could be used by ego; helps to look for another's arrows; prepares or repairs another man's bow and arrows in the middle of a pursuit; goes back on trail to warn others of wasp nest; walks toward other hunters to warn of fresh jaguar tracks or poisonous snakes; removes dangerous obstacles from the trail before others arrive. (p. 114)

I include this list in its entirety to emphasize the particular observation of such pervasive cooperation in a foraging culture. This is particularly notable given that the data Hill (2002) analyzed was not originally collected in an effort to identify cooperative behavior, and these behaviors and findings were thus simply background observations from data that was collected in differently focused studies. The author notes that the estimates of cooperative time presented are therefore a minimum estimate (Hill, 2002, p. 114).

Findings indicate that men spent an average of 41 minutes per day in food-acquisition activities that could be viewed as cooperative, and women spent an average of 33 minutes per day in cooperative foraging (Hill, 2002, p. 114). What is also noteworthy here is that this evidence of cooperation is specific to food gathering. As I referenced above, there is also significant data that food sharing was a key element of the transition

from apes to humans with some evidence suggesting that communal butchering sites and shared tools for food processing existed among human populations at least 2.5 million years ago (Ambrose, 2001; Klein, 2000). In this research, authors find that large game is always distributed among the group regardless of individual hunting success while vegetable matter is typically kept within families (Ambrose, 2001; Klein, 2000). While the relationship between cooperation and food sharing or gathering is significant, it is likely that cooperative acts are a much more pervasive part of daily life than captured through even this specific focus.

These rich examples suggest that anthropological research can show that cooperation is ubiquitous across human societies, including hunter-gatherer cultures. While an area of research that offers fascinating insight into the evolution of cooperation and the particular building blocks of what we consider ethical mindedness, the central purpose of the current section is to begin demonstrating the ubiquity necessary for making an evolutionary claim. I shift now to the final criterion for evolutionary argument, or the evidence of neural circuits.

Neural circuits. As I stated in Chapter 1, perhaps the strongest case for an adaptive trait can be made if a specific neural network can be identified. Given the current state of the science, this is neither typical of evolutionary arguments nor necessary. Below I will discuss some evidence that may be viewed as tangentially related to evolutionary claims of cooperation and ethical mindedness. That said, however, it should be stated clearly that no neural circuits have been identified specifically for cooperation.

Greene and Paxton (2009) examined patterns of neural activity associated with honest and dishonest moral decisions. In this research, participants underwent functional magnetic resonance imaging (fMRI) studies and were offered the opportunity to earn money by accurately predicting computerized coin-flips. The conducted trials in which participants had to record their predictions ahead of time as well as trials in which they did not, which allowed them to potentially earn the monetary reward by lying about their predictions. The authors found that many of the subjects behaved dishonestly in order to gain more of the reward (and based on their improbable levels of accurate predictions, which fell well above the expected level of chance).

Where their findings suggest a relevant neural circuit is in the neural activity associated with both honest and dishonest behavior. Individuals who acted honestly showed no control-related activity in the prefrontal cortex when choosing to behave honestly. In contrast, those who acted dishonestly with some frequency demonstrated increased activity in the prefrontal cortex when they chose to behave honestly. In addition to this, the response latency for dishonest people who were pretending to answer incorrectly increased, which supports evidence that hesitation is indicative of unaccustomed honesty. How can we interpret these results as they relate to ethical mindedness? While they fall slightly outside the present focus on cooperation, these findings indicate that dishonest people deploy honesty to maintain the illusion that they are honest, and that this action requires effort. The participants were observed hesitating only when dishonest people acted honestly. That hesitation was not found when honest people reported honestly but emerged when dishonest people were making an effort to deceive others. That is the key observation: maintaining a plausible deception required

explicit effort and activation. While the authors found a natural, instinctive tendency towards honesty among many individuals, they also observed a more labored effort when taking advantage of or exploiting others. That exploitation may relate closely to a discussion of cooperation, but the hypothesis and connection would be purely speculative given the focus of the current study.

In addition to this research, Keltner, Kogan, Piff, and Saturn (2014) note that other fMRI studies have found the following:

Images of suffering and need (e.g., images of physical malnutrition in children) activated a midbrain region known as the periaqueductal gray, which when stimulated in mammalian species triggers caregiving behavior (Simon-Thomas et al. 2012). Within the social psychological literature, cues related to need, harm, vulnerability, and weakness trigger more prosocial tendencies (Penner et al. 2005, Piff et al. 2010). (p. 430)

Keltner et al. (2014) also highlight findings that several neurotransmitters are associated with prosocial behavior. That these neurotransmitters include oxytocin (Kogan et al., 2011; Kosfeld, Heinrichs, Zak, Fishbacher, & Fehr, 2005; Rodrigues, Saslow, Garcia, John, & Keltner, 2009), serotonin (Crockett, Clark, Hauser, & Robbins, 2010; Crockett, Clark, Tabibnia, Lieberman, & Robbins, 2008), and dopamine (Bachner-Melman et al., 2005; Reuter, Frenzel, Walter, Markett, & Montag, 2011) indicates that we may have evolved to experience pleasure associated with prosocial behaviors. Such programming would suggest that we do have these neurological building blocks and an evolved propensity to be concerned with moral functioning and ethics. While there are multiple findings that appear related to the possibility of a neural circuits for behaviors related to cooperation, it should be made explicit that no circuit has been identified for cooperation itself. As I stated in Chapter 1, this is not a requirement for evolutionary argument, and it is unlikely that a particular circuit for cooperation, let alone for ethical mindedness will

be identified in the foreseeable future. I include these relevant results, not to suggest that they provide such evidence, but simply because their relevance is of interest, and they do suggest directions for future research that may one day reveal a more specific neural substrate.

Concluding Points

I began this chapter by contextualizing my focus on cooperation as one way of demonstrating the use of the criteria for evolutionary argument to begin building an explanation for ethical mindedness from an evolutionary perspective. It is important to note how I have explored cooperation through the seven criteria that followed. Initially, I focused heavily on identifying the adaptive problem of food inconsistency and the solution of food sharing. This provided an EEA for the emergence of food sharing, which occurred as humans transitioned to bipedalism and were able to move greater distances, ultimately migrating from the more densely packed wooded areas and into the Savanna. This environment offered our ancestors adequate resources, but not the consistent availability of those resources. Food sharing thus emerged as the solution to the adaptive problem of food inconsistency. Above I focused on food sharing in my demonstration of fitness benefits and the lock-and-key fit between adaptive problem and adaptation. In this chapter I also discussed how food sharing is the likely basis for food cooperation and began a discussion of cooperation as it has been generalized from this early adaptation. In examining research related to alternative explanations for cooperation, I presented evidence indicating how cooperation can evolve and spread throughout a population; or the quickness with which cooperation becomes a dominant practice within a group. Although I also discussed food sharing specifically in this section (as an adaptation not

better explained by kin selection or hierarchy), one of my main points was to make clear that cooperation with cooperators is a more general adaptation shown to quickly spread throughout a population without reference to a specific form of cooperation.

In my research on cooperation in our phylogenetic relatives and as it naturally develops from infancy, I thus continued this broad focus on cooperation. I showed how human cooperation both resembles and differs from that of our closest phylogenetic relatives, and I included research ranging from infancy up to 3.5 years of age to highlight evidence that such specifically human cooperation appears to emerge naturally over the course of development.

Following this, I explored anthropological research in hunter gatherer societies to demonstrate the ubiquity of cooperative activity across cultures and to include research on societies that are generally viewed as a proxy for the ancestral environment. And, finally, I included relevant neurological findings which, despite not offering a specific neural substrate for cooperation or ethical mindedness, do ultimately provide some evidence of related neurological activity.

I review all of this to emphasize that this chapter has included research from evolutionary psychologists, evolutionary biologists, primatologists, social psychologists, developmental psychologists, anthropologists, and cognitive psychologists. I have reviewed findings across disciplines in the social sciences in order to demonstrate, according to the criteria set forth by evolutionary psychologists, that there appears to be consistent and well-documented evidence that cooperation meets the criteria for an evolved adaptation. However, while I believe these findings adequately answer the question “is cooperation a species-typical feature of human nature”, the purpose of this

dissertation is to examine ethical mindedness. I will continue addressing this question in Chapter 4 by using an evolutionary perspective to connect the idea of cooperation to the emergence of ethical mindedness. This will lead further toward an ultimate explanation of how and why humans came to be ethically minded and offer a view of moral functioning that is grounded in an ontological explanation of human nature.

CHAPTER 4: THE CONNECTION TO ETHICAL MINDEDNESS AND A FINAL NOTE ON MORAL PSYCHOLOGY

The purpose of Chapter 3 was to utilize criteria from evolutionary psychology in order to demonstrate that cooperation is an evolved, species-typical feature of human nature. In Chapter 4, I will now transition to addressing the ultimate questions put forth in this dissertation: Is ethical mindedness a demonstrable feature of human nature, and what might an adequate theory of ethical mindedness look like?

I will continue to use an evolutionary approach in order to connect the adaptation of cooperation to the idea of ethical mindedness. I will do this by examining cooperation as it was made possible in increasingly large groups and among strangers through cheater detection, indirect reciprocity, and third party punishment; which are adaptations that I will use to illustrate the emergence of trust, a fairness instinct, and, ultimately, ethical mindedness. I will focus significantly on the instinct to be attuned to fairness (fairness instinct), and therefore ethical mindedness, as it emerges from this evolutionary viewpoint. As I have suggested in this chapter, I believe a cohesive evolutionary approach leads us to a view of fairness that is wholly communal.

Finally, I will contrast this communal view of fairness with the existing views put forth by moral psychologists in order to demonstrate what can be missed by not addressing such ontological questions from a perspective of how the various features of our nature emerged and became species-typical over time.

Cooperation, Trust, and the Connection to Ethical Mindedness

In Chapter 3, I demonstrated that cooperation clearly meets the necessary criteria to be considered an evolved feature of human nature. As an adaptation, patterns and practices of cooperation are characterized by clear fitness benefits; a lock-and-key fit

between the adaptation and the adaptive problem; lack of explanation as a byproduct of an alternative adaptation; presence and absence in our close phylogenetic relatives; natural development from birth; and ubiquity across cultures. Although it cannot be tied to specific neural substrates, there are promising directions for future research and relevant findings. Taken together, there is little argument against cooperation as an evolved, species-typical feature, and cooperation was demonstrated to have three key adaptive benefits. These benefits included increasing the quantity of a particular good, the opportunity to increase the diversity of goods through trade or sharing, and the creation of shared goods that are only made possible through cooperative activities (Fowers, 2015).

But what then makes cooperation ethical? Without a clear connection to the idea of ethical mindedness, we may be able to view this in simple terms; or consider that the idea of cooperation begins and ends with the demonstration of a pattern of behavior that indicates no deeper features about our basic nature. In other words, cooperation could be viewed as simple quid pro quo, which does not necessarily rise to the level of ethical or moral functioning.

What is interesting, then, is the seemingly paradoxical context in which cooperation appears to have spread throughout the species. That is, widespread cooperation creates the opportunity for non-cooperators to exploit cooperators by gaining benefits without paying the cost of cooperation. This would seem to provide a particular advantage for exploitative non-cooperators, making it curious that it was cooperation, not exploitative non-cooperation, which emerged as a dominant species trait. The advantages

of exploitation thus create a specific dilemma around a system of widespread cooperation. West, Griffin, and Gardner (2007) reiterated this dilemma by explaining,

A behavior is cooperative if it provides a benefit to another individual and if it has evolved at least partially because of this benefit. Such behaviors pose a problem to evolutionary theory because — all else being equal — they would reduce the relative fitness of the performer of that behavior and hence be selected against. (p. R661)

It is interesting, then, that we find ubiquitous cooperation despite the seemingly clear benefits of exploitation. Such ubiquity indicates that as a species, cooperation has become normalized in many contexts. While kin selection, direct reciprocity, and indirect reciprocity may explain cooperation in dyads and family groups, they do not adequately explain cooperation in groups of unrelated individuals, let alone cooperation in large groups made up primarily of strangers (Fowers, 2015). This dilemma of the proliferation of cooperation in large groups has a number of potential resolutions. One possible resolution is that the willingness to cooperate among strangers requires trust that one's cooperation will be reciprocated, and such trust depends upon the individual's confidence that one has the abilities to both detect and sanction cheaters. This first ability to detect cheaters is appropriately referred to as cheater detection. It is a critical adaptation to the emergence of widespread cooperation.

Cheater detection. Cosmides and Tooby (1992) review extensive research that was designed to identify and document the idea of cheater detection. In doing so, they not only document a lock-and-key adaptation to an adaptive problem, but they also demonstrate one means of ruling out explanations by alternative adaptations. The authors identify the Wason selection task (Wason, 1966) as “one of the most intriguing and widely used experimental paradigms for exploring people's ability to detect violations of

conditional rules” (Cosmides & Tooby, 1992, p. 181). In this task, subjects are asked to determine whether the conditional statement such as “if P then Q” was violated in any of four instances that are represented by cards. A statement in this form is thus violated only when P is true but Q is false. Wason originally expected that people would be relatively good at identifying such simple violations, but research over the past 25 years has found that “few people actually give this logically correct answer” (Cosmides & Tooby, 1992, p. 181).

Where these findings become particularly interesting is in how researchers have manipulated the conditional statement to assess how people respond based on varying degrees of familiarity. For example, a more familiar conditional statement may be “If a person goes to Boston, then he takes the subway” or “If a person eats hot chili peppers, then he will drink a cold beer” (Cosmides & Tooby, 1992, p. 181). Less familiar statements of relationships may include “If you eat duiker meat, then you have found an ostrich eggshell” or “If there is an ‘A’ on one side of a card, then there is a ‘3’ on the other side” (Cosmides & Tooby, 1992, p. 181).

The authors note that in many of these experiments, performance on familiar rules is comparable to performance on unfamiliar ones. Those that are focused on food, for example, “have never elicited logical performance higher than that elicited by unfamiliar rules, even though the typical sophomore in such experiments has had about 22,000 experiences in which he or she has had both food and drink, and even though recurrent relations between certain foods and certain drinks are common—cereal with orange juice at breakfast, red wine with red meat, coffee with dessert, and so on” (Cosmides & Tooby, 1992, p. 181). Even the example above about going to Boston, which has the best record

for eliciting logically correct responses, was answered correctly by only 48% of 96 respondents in a best-performing study by Cosmides (1989, cited in Cosmides & Tooby, 1992).

Where these findings become especially interesting and relevant to the current focus is when the subjects are asked to look for violations of conditional rules about social contracts. For example, the statement “If a person is drinking a beer, then he must be over 20 years old” expresses a social contract insofar as it communicates that “one is entitled to receive a benefit (beer) only if one has satisfied a requirement (being a certain age)” (Cosmides & Tooby, 1992, p. 182). When respondents are provided with this conditional statement, the logically correct responses soar to roughly 75% (e.g. Griggs & Cox, 1982; Cosmides, 1985, cited in Cosmides & Tooby, 1992).

Cosmides and Tooby (1983) assumed the task of analyzing the total literature on the Wason task to develop a theory that explained what effect the content of the statement had on logical performance. What they found was striking and specific to content about social contracts. The authors explain the following:

Robust and replicable content effects were found only for rules that related terms that are recognizable as benefits and cost/requirements in the format of a standard social contract. ... Moreover, most of the content effects reported for non-social contract rules were either weak, clouded by procedural difficulties, or had some earmarks of a social contract problem. All told, for non-social contract thematic problems, 3 experiments had produced a substantial content effect, 2 had produced a weak content effect, and 14 had produced no content effect at all. The few effects that were found did not replicate. In contrast, 16 out of 16 experiments that fit the criteria for standard social contracts—i.e., 100%—elicited substantial content effects. (Cosmides & Tooby, 1992, p. 183)

For authors in search of evidence that demonstrates an evolved capacity, these findings are notable and exciting. They are not, however, independently sufficient. The authors caution first that the social contract feature may be confounded with familiarity, and

second that it is also possible that the social contract content “simply facilitates logical reasoning” instead of suggesting a special design; that is the existence of an adaptation specific to social exchange (Cosmides & Tooby, 1992, p. 184).

To address these concerns, Cosmides and Tooby demonstrated that even when using statements that were “unfamiliar and culturally alien, the social contract problems elicited a high percentage of P & not-Q responses” (Cosmides & Tooby, 1992, p. 186). More specifically, the performance on unfamiliar social contracts was found to be comparable to that on familiar social contracts and even higher than familiar descriptive statements (such as that about Boston and taking the train). Given this, the authors reasonably concluded that familiarity cannot explain the pattern of reasoning and that “social contract performance is not a by-product of familiarity” (Cosmides & Tooby, 1992, p. 187).

In assessing whether the improved performance occurred because social contract content facilitated logical reasoning, the authors conducted studies in which the social contract of the statements were left identical, but the modus ponens, or the logical “if p then q” statement of antecedent and consequent inference, was reversed. For example, the statement “if you give me your watch, I’ll give you \$20” is compatible with modus ponens (q and p) and the violation is “I take your watch, but don’t give you \$20” (q and not p). The statement that does not follow the modus ponens form was a reversed social contract form of “If I give you \$20, you give me your watch” and the violation for this statement was, “I don’t give you \$20, and I take your watch” (not-p and q). In these cases, the social contract and violation remain the same but the rule of logic is different in the latter form. The authors found that respondents continued to perform well on the

social contract problems (with correct responses from 67% and 75% of respondents on social contract problems) and poorly (answering correctly just 4% of the time) when provided with q and not-p conditions for non-social contract statements. This again suggests that the significantly improved performance is related to the social contract, not the form of logical processing.

A third set of studies from Gigerenzer and Hug (1992) analyzed whether the respondent's perspective influenced their perception of cheating. In these studies, the authors used the social contract "If an employee gets a pension, then that employee must have worked for the firm for at least 10 years." Participants were categorized in two conditions, as either the employer or the employee. From the employer's perspective, cheating would be receiving the pension without working for ten years, whereas from the employee's perspective, cheating would be working for ten years but not receiving the pension. Participants in these studies again responded consistently with the existing data. On the switched social contract (the employee perspective), respondents were correct about 60% of the time, while from the employer's perspective, 75% of respondents answered correctly. This is a particularly important finding when looking for evidence of a capacity for cheater detection because such detection is only useful if it can be applied from each individual's perspective.

Finally, Cosmides and Tooby examined whether "altruist detection" was as robust a capacity as cheater detection. In terms of the social contracts, an altruist was defined as one who would pay the cost, but not accept the benefit. In this research, Cosmides (1989) found that performance plummeted back to the rate observed when solving familiar descriptive problems that were unrelated to social contracts. This is a critical finding.

What we see here is that the capacity for detection is not simply related to the violation of a cost-benefit equation, nor is it generalized to any social contract. It is instead highly specialized for the detection of cheating.

In a later follow-up, Cosmides and Tooby (2005) compared rule violations in two conditions: one in which the violation was committed due to inattentiveness, and a second in which it was intentional. They found that participants correctly identified the violation 68% of the time when it was intentional but only 27% of the times when it was unintentional. We see here further evidence that the cheater detection capacity is highly specialized and designed to detect cheaters who knowingly violate social contracts.

What we ultimately see here is this: a highly specialized adaptation that makes cooperative social life possible in face-to-face interaction. We also see how this adaptation has been examined as a lock-and-key fit for an adaptive problem (the threat of exploitation in cooperative interaction) and studied for whether it could potentially be explained as the byproduct of an alternative adaptation. This is the precise meaning of the special design of features that solve adaptive problems very specifically and it demonstrates that cooperation cannot be explained by any alternative adaptations.

In and of itself, however, even cheater detection does not bring us to the level of widespread cooperation that is observed among human beings. What cheater detection alone accomplishes, rather, is widespread cooperation among dyadic pairs. Cheater detection gets us to dyadic, interpersonal cooperation, and it suggests a basic investment in the trustworthiness of those one interacts with, but it does not provide us with the full story of how the adaptation of cooperation has contributed to the emergence of ethical mindedness. For ethical mindedness, or a larger investment in morality and ethics and the

propensity to evaluate the actions of even those with whom I am not directly interacting, cooperation must be linked to broader social functioning. To make this case, the next step is a focus on indirect reciprocity as an evolved adaptation that shifts cooperation and cheater detection one step further toward a demonstration of ethical mindedness.

Indirect reciprocity. The idea of direct reciprocity provides us with an explanation for cooperation in dyadic interactions. In other words, one is likely to reciprocate the cooperation they receive from another. Indirect reciprocity is the idea that explains cooperation with strangers who are known to have cooperated with others; or the idea that one is also likely to cooperate with someone with whom they have never interacted if they know that person has cooperated with others in the past. In exploring whether cooperation can be shown to develop naturally and without specific teaching or socialization in Chapter 3, I discussed a study from Olson and Spelke (2008) in which 3-4 year old children were found to demonstrate preference toward individuals who had acted generously toward others. This idea that we show preference toward cooperative others, and develop trust based on reputation and reports of past cooperation, is a critical development when tracking the connections between cooperation and ethical mindedness.

Cheater detection is an adaptation that made it feasible for individuals to cooperate with a degree of confidence that they won't be taken advantage of. Within a dyadic context, cheater detection provides the confidence that individuals have some information indicating that they should trust their cooperative partner or expect fair treatment from them. Indirect reciprocity provides another form of this evidence or information. With indirect reciprocity, individuals can now trust another because they know that person treated someone else fairly. In this sense, the willingness to cooperate is

based on the judgment of trustworthiness, which is based on the history of fair behavior in some context.

Indirect reciprocity is thus about reputation. If someone is seen cooperating with a third party, that person is then presumed to be safe for cooperation, and these observations need not ultimately be direct. Human language affords us the possibility of communicating about what happens in different times and places, with people who weren't present for those events. This appears to be a unique human capacity, and it is reflected in indications that human beings rely heavily upon gossip and reputation to determine the likelihood of fairness and cooperation from others (Keltner et al., 2014). Keltner et al. (2014) view cooperation as one mode of prosociality, and argue that “concern over gossip is a powerful check against free riding – that is, gossip is a form of cost for not acting prosocially and thereby motivates others to prosocial action” (p. 439). In other words, gossip and its related influence on reputation are powerful motivators when it comes to cooperation and social action. Not only can one person tell a second that a third party cooperated with them, thereby enhancing the reputation of the third party as trustworthy; but concern over the opposite, or the development of a negative or untrustworthy reputation also serves as motivation for cooperative behavior.

What we gain from this form of information is significant. Where cheater detection explained cooperation in dyads, indirect reciprocity explains cooperation in a triadic context. Beyond this, even, those triads need not be immediate or closely related. Take, for example, the case of internet reviews, which have become a powerful means of both locating trustworthy cooperators and encouraging trustworthy cooperation. These

reviews are nothing more than electronically mediated gossip that informs cooperation with strangers.

When we consider this extension of indirect reciprocity, what we observe is the way a phenomenon that worked extremely well in a given band of people, or a specific group context, has since been generalized. Indirect reciprocity now works whether one is familiar with the other two actors or not. This generalization is key. Where we see the transition from cheater detection and dyadic cooperation to indirect reciprocity and triadic cooperation, we now see how these basic features are applied outside of the group context in which they emerged and now inform much larger and more complicated social interaction.

That said, even this connection between cooperation, cheater detection, and indirect reciprocity does not bring us to the level of ethical mindedness. As I stated in the beginning of this section, such widespread cooperation among strangers requires not only trust that one's cooperative actions will be reciprocated, but a system in which the identified cheating and exploitation of cooperation will be sanctioned not just by oneself, but also by others. Cheater detection can be viewed as supporting direct and indirect reciprocity, but in order to maintain cooperation in large groups, we must have the ability to punish cheaters at the group level (Fowers, 2015).

Punishment and third party punishment. It is first important to note that punishment is not a uniquely human form of behavior. We find evidence that animals may punish fitness-lowering behavior or in efforts to maintain hierarchical status (cf. Clutton-Brock & Parker, 1995; de Waal, 1996). Among most non-human animals, however, these behaviors appear to typically be limited to dyadic interactions in which an

aggrieved animal punishes one that violated it. While there is some evidence that chimpanzees and other primates sanction norm violations (de Waal, 1996), human use of punishment differs from most of our primate ancestors in the scope and enforcement of punishment. While we do engage in direct punishment against those we view as violating our prerogatives, we also frequently sanction individuals who harm others even when we ourselves are unaffected by their transgression. This behavior is known as third-party punishment, and it presents an interesting dilemma as it can appear to be costly to those who impose it without benefitting them directly. Fowers (2015) identifies three questions that have been studied closely in research on third party punishment: “Are people willing to incur its costs? Can it help to sustain cooperation in a population? And is it associated with reward centers in the brain” (p. 178)?

A closer examination of this research reveals a clear affirmative answer to all three questions. Game theory researchers have utilized the ultimatum game to examine whether humans across cultures act as rational maximizers (i.e., whether they are motivated simply to maximize their personal gain). The alternative to such rational maximization in these studies is to sacrifice personal gain in order to enforce a fairer outcome. The researchers use the ultimatum game, in which one individual (the proposer) is provided with a sum of money and asked to propose how the money be divided between themselves and a responder. The responder can then accept the proposed distribution, receiving whatever amount was allocated to them, or reject the proposal and neither participant will receive any amount of money. Fehr and Fishbacher (2003) find that responders show a high rate of rejection when they are offered less than 25 percent of the money, and such findings have been replicated across cultures (Henrich & Henrich,

2007; Roth, Panikar, Okuno-Fujiwara, & Zamir, 1991). In other words, people tend to reject offers, sacrificing a personal gain. They do this even though it requires a sacrifice of personal gain (i.e., they receive nothing instead of, say, 20% of the endowment).

Beyond this, Sanfey, Rilling, Aronson, Nystrom, and Cohen (2003) use an fMRI to study the neural activation of responders participating in these ultimatum games. They replicated the previous studies and showed that receiving low offers from the proposers was associated with increased activation in the bilateral anterior insula, which is associated with negative emotional arousal and specifically with anger or disgust. This activation was stronger with increasingly inequitable offers from the proposer. The authors also found activation of the dorsolateral prefrontal cortex (DLPFC), which they interpreted as the cognitive process of determining whether to accept or reject an offer, and the anterior cingulate cortex (ACC), which may be related to the conflict between a desire to maximize gain and the powerful emotional reaction to what was viewed as an unfair offer. These studies suggest that neural activations can be associated decision making about the distribution of resources.

Nowak, Page, and Sigmund (2000) conducted a simulation study to explore whether humans will act as rational maximizers and accept any amount that is offered to them, or whether they will prioritize fairness and be willing to sacrifice a small gain in order to ensure a fairer outcome. The authors examined the ultimatum game from an evolutionary perspective by studying the effects of making information about the responder's choices available to the proposer. In this, they establish that when no information is given to the proposer about the responder's options, the proposers may favor rational maximization, or tend to make unfair proposals, which are ultimately likely

to be rejected. In contrast, if information is available about the responders' choices (i.e., that the responder either accepts the previous offer given or rejects the offer and neither player gets anything), fairness emerges as the stable solution over time. The authors thus find that while proposers may favor rational maximizing when given no information about the responder's rejection of offers, making information available about their choices "readily leads to the evolution of fairness" (p. 1774). This research can reliably indicate the evolution of trust and fairness but does not include a third party. Therefore, it is important to examine research that includes the possibility of third party punishment.

Bowles and Gintis (2004) studied the evolution of third party punishment with a simulation that evaluates the benefits of "joint projects such as hunting and common defense as well as sharing of food, valuable information, and other sources of survival" (p. 18). They include "actors" who are cheaters or selfish (rational maximizers), cooperators who cooperate with everyone but *do not* punish cheaters, and reciprocators who cooperate and also *do* punish the cheaters. The highest individual benefits are given to the cheaters, suggesting that this behavior should hypothetically drive out cooperation, and the reciprocators pay a cost for enacting punishment, which gives them reduced fitness relative to both cooperators and cheaters. The punishment enacted was to increase the probability of ostracism from the group, which reflects the punishment enacted by hunter-gatherers (Boehm, 2000).

The authors ran simulations on 20 groups with 20 members each and allowed individuals to migrate between groups. The groups were initially 100 percent rational maximizers with cooperators and reciprocators entering only through a low rate of mutation, and the passing of one's actor type onto one's children. The authors present

several key findings. First, they demonstrate a long-term stabilization of 37.2 percent reciprocators, 24.6 percent cooperators, and 38.2 percent cheaters. This is a substantial emergence of cooperators and reciprocators despite entering through a rare random mutation, and in the face of reproduction benefits for cheaters. They also found that a high level of cooperation need only emerge in one group because that group would increase in size faster, resulting in migration to other groups and growing population in otherwise abandoned areas. Interestingly, the cheaters were found to be disproportionately without a group membership (93% of the actors not affiliated with a group were cheaters), which indicates that ostracism functioned as expected to remove such behavior from all groups. In other words, these authors find reason to believe that the evolution of cooperation does occur through third party punishment, even when using assumptions that appear to be more stringent than what has been observed in actual hunter-gatherer practices (cf. Boehm, 2000).

The purpose of including this research, and of examining third party punishment so closely is to follow an explanation of cooperation and a resolution of the dilemma of cooperation as a species-typical and beneficial adaptation. The connection from cooperation to cheater detection provided us with insight into how cooperation emerged in dyadic contexts. Similarly, the connection to indirect reciprocity illustrated a triadic context and introduced the importance of reputation and gossip to understanding cooperation among not only immediate triads, but strangers as well. This suggests a specific attention and attunement to fairness, and the fair treatment of others. By examining third party punishment, we see how cooperation leads us to group norms and the idea of strong reciprocation. This is a key point made by Bowles and Gintis (2004).

The authors examined why people engage in third party punishment and find that although third party punishment may be initially costly to the individual who inflicts it, the process evolves. If there is a common thread of third party punishment within a group then group members are less likely to cheat in general. If anyone and everyone can punish someone for cheating, cheating simply becomes too costly within the group. What emerges from this practice of third party punishment is the creation of communal fairness. That is, the focus of fairness maintenance is on the entire group, not just on interactions between dyads.

There are several caveats that are important to note here. The first is a reminder that when I discuss fairness, or communal fairness, in these terms, I continue to emphasize fairness in terms of the instinctive interest in and concern for fairness (not any specific version or manifestation of fairness that emerges from that instinct). The second note here is that communal fairness, or the idea of it as I have discussed here, is not a phenomenon that emerged all at once.

I have traced this emergence from food sharing, which became more general with cooperation, and I have used West et al.'s (2011) definition of cooperation as “a behavior which provides benefit to another individual (recipient) and which is selected for because of its beneficial effect on the recipient” (p. 235). Continuing this process, cooperation was then generalized and spread through cheater detection, indirect reciprocity, and third party punishment. Within that process, gossip also emerged as a communal practice, or a way of manifesting this communal cooperation by speaking freely about third parties and their praise- or blame-worthy attributes. We are effectively communicating to our communities and in doing so we establish and enforce group norms.

In this sense such norms are also a communal practice – we set up reasonable expectations that we expect ourselves and others to live up to. It is a process, not an outcome, and it is instantiated through a myriad of interactions over the entire lifespan of a group. This communal act of cooperation is an ethos because it is embodied in the group, not the individual. It exists in the way the group is structured, the way it interacts, and the expectations that are built into the group's life.

Finally, it is important to also understand that all of this has occurred through an evolutionary process that emerged over thousands of generations, but when I refer to something as communal, I am locating that within the context of the EEA in which it emerged. In this case, communality refers to groups of 25-30 people. While our globalized world is now inconceivably larger and more interconnected, it is important to understand that the basic features of our nature emerged in an entirely different context. Cooperation may not be drastically different than it used to be, but it is certainly more generalized. People have set ways of operating that depend upon others doing their part, and that communal practice and nature has been around for a long time.

Communal fairness and the connection to ethical mindedness. This chapter thus began with a focus on the dilemma of cooperation; or the question of how cooperation emerged as an apparently species-typical adaptation despite it appearing beneficial to take advantage of cooperators. I have examined cheater detection, indirect reciprocity, and third party punishment as adaptations that support positive cooperation (e.g. turn taking, easy reciprocity). Cheater detection involves identifying both cheaters and cooperators so that individuals can choose good cooperation partners. Similarly, indirect reciprocity functions to enable the selection of good cooperative partners as well

as demonstrating both the positive and negative patterns of gossip (which can spread either praise- or blameworthy information about others). Finally, the purpose of third party punishment is ultimately to maintain positive norms. I focus on these particular explanations not only as a way of resolving the dilemma of cooperation, but also to explain how the resolution of that dilemma, and the emergence of cooperation as a species-typical feature of human nature, lead to the idea of ethical mindedness.

I began this examination with a focus on cheater detection, which provided an explanation for cooperation in dyadic contexts. This explanation included evidence of a self-interested human attunement to evaluating the trustworthiness of others. While cheater detection provides some evidence of trustworthiness, that evidence is incomplete.

I thus followed this discussion of cheater detection with a focus on indirect reciprocity as it offers an explanation for cooperation in triadic contexts and evidence for the roles that gossip and reputation play in wider circles of trustworthiness. Indirect reciprocity also illustrates how a basic capacity has been generalized to strangers and individuals outside of one's group (e.g. internet reviews). From there, I focused on third party punishment in order to examine cooperation in the context of increasingly large groups and among strangers.

Third-party punishment is itself an adaptation that can be studied according to the relevant evolutionary criteria. Although doing so in great detail is beyond the scope of the current dissertation, much of that work has been done. There is also neuroscientific evidence that humans have negative emotional reactions to perceived violations of fairness that does not affect them directly (cf. Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003; Tabibnia, Satpute, & Lieberman, 2008) and, relatedly, compelling evidence

that we find third-party punishment rewarding (cf. Rilling et al., 2002; de Quervain et al., 2004).

The real outcome of third party punishment appears to be the relatively frictionless exchanges that are made possible by being a member of a trusting group. Third party punishment makes fairness an enforced group norm and, by doing so, reduces the need for surveillance, lowers the conflict and friction around exchanges, and enhances the harmony of the group. Provided these practices and communal features, individuals need to devote almost no energy to the enforcement of cooperation.

These group norms and enforcements lead to the concepts of a fairness instinct and communal fairness that facilitate that outcome. An evolutionary analysis thus leads to the increasingly clear conclusion that cooperation (a demonstrable adaptation) is seamlessly connected to ethical mindedness, or the instinct to learn an ethical viewpoint and the basic tendency to be concerned about the morality and ethics of others.

In other words, one way of resolving the dilemma of cooperation is to view third-party punishment as operative in maintaining cooperation, which benefits all individuals in the group, making such behavior an act of group-level cooperation in itself (i.e., metacooperation). In this sense, we see the idea of fairness as operating at a larger, communal level instead of being simply an individual concern. This means that individuals are not simply concerned with how someone treats them, but also how that person treats others. We are concerned about whether cooperation can be reasonably expected in our groups, and with this concern comes a focus on how individuals act regardless of the immediate impact those actions have on ourselves.

As stated by Fowers (2015), “the extent of cooperation is naturally limited by the degree to which trust is complete, which, in turn, is dependent on reasonable expectations of fairness” (p. 185). We saw an initial focus on trust in the example of cheater detection. That idea of trust could be strengthened through indirect reciprocity, or through gossip and reputation. But what we gain from third party punishment in terms of the expectation of trust and fairness is significant. It is only through the widespread practice of third party punishment that a communal fairness emerges, as opposed to one that was simply dyadic or specific to a particular exchange. What an evolutionary perspective on third party punishment offers us is instead the view that through the widespread practice of third party punishment, cheating becomes more costly than cooperation within a group. This explains how cooperation emerges as a species-typical feature and not one that can be beneficially taken advantage of by cheaters.

What we see here is a connection from cooperation to third party punishment, and a connection from third party punishment to trust and the expectation of fairness. Once we arrive at this reasonable expectation of fairness, and evidence that humans evolved to react emotionally to violations and confirmations of that expectation, we begin to see stronger evidence of ethical mindedness. Through cheater detection, indirect reciprocity, third party punishment, and the ubiquitous practice of cooperation that these adaptations explain, it becomes increasingly clear that human beings did in fact evolve to be instinctively concerned with the ethics and morality of others. Cheater detection (or an instinctive evaluation of trustworthiness), indirect reciprocity (or an attunement to reputation and gossip), and third party punishment (or the enforcement of group norms of fairness) are evaluations of moral acceptability that lead to a communal-level social

concern and enforcement. It is through this process, and the inextricable links between observable cooperation and an understanding of its evolutionary emergence that we see how using an evolutionary framework for demonstrating an evolved adaptation (cooperation) leads us in the direction of an answer for the question of whether ethical mindedness is a species-typical feature of human nature.

By taking an evolutionary approach, this chapter has attempted to build a view of our instinct for fairness based on how it emerges from our basic human nature, and that view differs greatly from what is frequently focused on by the existing theories of moral psychology. To a large degree, these theories take an interpersonal perspective that does not lead seamlessly to the communal nature of fairness. But it is only when fairness becomes a communal level concern that can be taken for granted that it can become a way of life. In that sense, fairness (which emerges directly from our evolved cooperation) appears not only to be a shared good (i.e., something only made possible through cooperative activities) that is far greater and more pervasive than can be captured by a simple focus on harm and care, but it appears to be a good with which humans evolved to be instinctively concerned. In that sense, we are led seamlessly from an evolutionary view of cooperation to evidence of our basic ethical mindedness.

Taken together, this chapter is intended to provide one example of how a theory of ethical mindedness can be built by following its evolutionary history. This work could be followed up with other examples. Social norms and belonging, to give two examples, are demonstrable adaptations that, like cooperation, can be clearly explored according to evolutionary criteria and are likely connected to the presence of ethical mindedness in human beings. The purpose of this section, however, was to provide an example of how

the evolutionary argument for cooperation that was presented in Chapter 3 can be extended to provide evidence that ethical mindedness is a species-typical feature of human nature.

Communal Fairness and Moral Psychology

If the primary conclusion of an evolutionary examination of cooperation is the presence of a communal-level fairness and its relationship to ethical mindedness as an evolved feature of human nature, then perhaps a key question becomes what would we miss if we didn't do this analysis? In other words, if my primary concerns about the existing work of moral psychologists were related to the incompleteness of their evolutionary reasoning, then it bears examining what can be missed when such analysis is not completed.

What is clear is that the existing moral psychologists understand the importance fairness and justice to moral functioning. Below I will briefly provide some examples of how the authors from each camp refer primarily to the ideas of fairness and justice. Having provided this evidence, I will then discuss how these statements conflict with ideas that were found by taking an evolutionary approach.

Social Cognitive Domain Theory. Social Cognitive Domain theorists rely heavily on Turiel's (1983) definition of morality as "prescriptive judgments of justice, rights, and welfare" (p. 3). We see this definition in the research across theorists from this perspective. Nucci and Gingo (2011) explain the following:

What has become clarified is that the domain of morality constitutes a basic knowledge system for regulating social interactions around issues of fairness and human welfare. This basic knowledge system stems from patterns of social interaction that are inherent to social relations, and happen to have connection to a common set of affective experiences. (p. 439)

Similarly, Nucci (2001) explains that, "Within the domain theory of social development, morality refers to conceptions of human welfare, justice, and rights, which are a function of the inherent features of interpersonal relations" (p. 7).

In a widely-cited claim, Turiel and Killen (2010) explain the following: Research on moral reasoning provides ample evidence of how seemingly formal, abstract concepts about welfare, justice, and rights are applied flexibly in social situations (see Helwig, 1995, on rights; Killen, Margie, & Sinno, 2006, on exclusion; and Turiel & Perkins, 2004, on honesty and trust). Moral reasoning does involve general concepts about welfare, justice, equality, and rights. As such, moral concepts are not monolithic, and moral ends can at times come in conflict with each other (e.g., welfare vs. rights). (p. 35)

Even within this acknowledgement of the complexity of moral functioning, however, the authors continue to distinguish morality as specific to the questions of welfare, justice, equality, and rights while also situating those concerns in specifically social situations. In this way, Social Cognitive Domain theorists make repeated reference to morality as based on matters of justice, rights, welfare, harm, and care (offering us only a limited view of moral functioning) while also specifying that these issues are primarily to be viewed in terms of the application of abstract moral principles to interpersonal relationships and functioning. This interpersonal domain is typically limited to person-to-person interactions that are often dyadic, but also includes the actions of an individual that may affect others more generally. The implicit commitment is to the idea that morality is a regulator of the individual's actions. This regulation of individual behavior is a dramatically different focus from the metacooperation revealed by an evolutionary examination of the evolution of cooperation. This illustrates the difference between the idea that morality has to control our amoral, selfish, individual nature and the idea that morality is an expression of our ethical, cooperative, and communal nature.

The issue here is not that the authors are necessarily incorrect in what they include or where they locate moral concerns, but that they are incomplete. By situating their view of justice within interpersonal relations, these authors lose the entirety of an evolutionary narrative in which justice emerges at the group or communal level. I will explore this issue in greater detail in the final section of this chapter, but reference it here as it is the specific concern that is present across moral psychologists.

Moral Identity. As I discussed at length in Chapter 2, moral identity theorists are a more loosely-knit group within moral psychology. Because of both this and their general focus on individual development, it is difficult to identify a specific domain and definition for moral functioning within this group. We do, however, see some common themes that speak to the domain of moral functioning.

Lapsley and Hill (2009) explain that, “even among those who value morality as a source of self-definition, there are different ways of living a moral life well – some might orient to justice, some to care, and still others to utility or virtues (and so on)” (p. 207). Here, the authors appear to understand that there are different ways of living a moral life (and presumably therefore different domains of morality), and they later acknowledge that these claims have not been adequately tested. Their brief discussion, however, is focused on dispositional development with an emphasis on personality and the possibility of different moral personality types. Although this is consistent with the basic idea of a moral identity, it fails to track and contextualize moral functioning in terms of its evolved, communal nature.

Similarly, Hart, Atkins, and Ford (1998) conceptualize moral identity in terms of actions that promote or protect the welfare of others. They define moral identity as, “a

self-consistent commitment to lines of action benefitting others” (Hart, Atkins, & Ford, 1998, p. 513). Colby and Damon (1993) state that, “people's moral choices are inevitably construed in the context of actual interpersonal relations, and moral ideas have behavioral meaning only insofar as they are actively interpreted in such contexts” (p. 153).

Many of the authors writing from a moral identity perspective associate morality with altruism, prosociality, and beneficence. We see this in Hardy and Carlo’s (2011) idea about moral identity and an altruistic personality. While they specify that this does not refer to all moral identity theorists, they describe the view of many with the following:

Some see moral identity as the heart of, or nearly synonymous with an “altruistic personality” (Carlo, Pytlik, Zillig, Roesch, & Dienstbier, 2009; Staub, 2005). The altruistic personality is a set of other-oriented tendencies or traits, such as empathy, social responsibility, and moral reasoning, which motivate prosocial behaviors and mitigate anti-social behaviors (whether deliberately or automatically; for reviews, see Carlo et al., 2009; Dovidio, Piliavin, Shroeder, & Penner, 2006). Care exemplars are individuals known for such altruistic personalities (Colby & Damon, 1992; Oliner & Oliner, 1988; Walker & Frimer, 2007). Driving or unifying the other-oriented tendencies of such individuals is a personal investment in helping others – indeed a sense of moral or altruistic identity. (p. 497)

In this, Hardy and Carlo (2011) cite a wide range of moral identity theorists in explaining the role of an altruistic personality. They explicitly describe this personality as “other-oriented” and reference traits including “empathy, social responsibility, and moral reasoning.” The other who can be benefitted may be an individual or a more general “other”.

None of this appears to be entirely inaccurate with regard to moral functioning or ethical mindedness. Empathy, social responsibility, and moral reasoning are certainly

features of human nature that relate to our fundamental ethical mindedness. Beyond this, so too may be our care for others and its basis in beneficence and altruism (Carlo, 2006).

What is missing here, however, is again a clear path from this view of moral functioning to a communal understanding of fairness as not emergent from individual relationships or concerns, but as a product of coordinated group behaviors. Again, morality is seen as the regulation of the individual's behavior, but in this view, behavior is also portrayed as being regulated by an identity that resides within the individual.

Moral Foundations Theory. As I explored more thoroughly in Chapter 2, Moral Foundations theorists have what is decidedly the broadest and most multifaceted view of what constitutes the domain of moral functioning. Unlike the focus on harm and care that we find in the Social Cognitive Domain and moral identity literatures, Moral Foundations Theory is based in moral pluralism and its most significant contribution is likely its inclusion of five distinct moral foundations. As a reminder, these foundations are harm/care, authority/respect, fairness/reciprocity, purity/sanctity, and ingroup/loyalty (Haidt & Graham, 2007).

As I also described in Chapter 2, the issue with these foundations is not that they appear to be inaccurate, but that they are likely incomplete. Beyond this, the difficulties the authors have with making adequate arguments for the five moral foundations, the discussion in Chapter 2 suggested that they may be confused about how to build an appropriate evolutionary argument, or how to properly conduct evolutionary research and reasoning. This confusion has contributed to the development of a theoretical framework that is not fully justified, and for which the existing reasoning appears based on significant category mistakes related to evolutionary evidence and research.

There are, however, strong statements from authors in this camp that indicate their understanding of the broader scope of moral functioning. Haidt and Joseph (2004) suggested that intuitive judgments that “pop into consciousness without our being aware of the mental processes that led to them” (p. 56) offer us a connection between sociological facts and systems and individual biology or functioning. Haidt (2007) also reflected upon sociological theory as he cites Durkheim’s idea in explaining that morality “binds and builds; it constrains individuals and ties them to each other to create groups that are emergent entities with new properties” (p. 1000). Note here the implicit commitment to individuals as somehow predating groups. This is, again, the type of mistake that can be made when attention isn’t paid to how individuals evolved.

Approaching these ideas more concretely, Graham, Haidt, and Rimm-Kaufman (2008) explained that they believe “it is better for social scientists to ignore philosophers and just examine morality as an empirical phenomenon. We think the best approach is a functionalist approach that is anchored in what people and societies are trying to do when they regulate, punish, gossip, praise, and otherwise engage in moral life, broadly speaking” (p. 271). Put even more simply, Haidt and Bjorklund (2008) state “we believe that moral judgment is best understood as a social process, not as a private act of cognition” (p. 181). These authors could simply mean transactions between individuals, or they could perhaps mean the development of norms of cooperation within groups; but they don’t specify or elaborate what is meant by social process so it is difficult to credit them with the group level of analysis.

If we take all of these statements together, what we find among the Moral Foundations theorists is this: They clearly understand that fairness and justice are one of

the key foundations of moral functioning. They also appear to gesture explicitly at the larger social nature of morality. But without a sound evolutionary approach, the authors seem not to understand that it is the group or communal nature of fairness from which individual morality emerges. They appear to struggle with piecing together an understanding of moral functioning that is based in individual processes and instincts, instead of viewing those instincts as emerging from a fundamentally communal nature. They also focus explicitly on morality as the regulation of individual behavior.

A final note on what is missing. In Chapter 2, I detailed the limitations in how authors from each of the main three theoretical perspectives of moral psychology have accounted for ethical mindedness according to an adequate evolutionary argument. What I show here, with regard to the relatively narrow definitions of morality that each camp appears to rely on, is one consequence of that failure to track moral functioning through its evolutionary origins. Without a full picture of how ethical mindedness, and therefore moral functioning, emerged as a species-typical feature of human nature, it has been impossible for these authors to fully understand the moral domain in which that functioning occurs.

So what is wrong with these conceptualizations of moral functioning and their emphasis on fairness and justice? It is not that fairness and justice aren't interpersonally relevant, or that they are not central issues in moral functioning and ethical mindedness. The issue also is not that the authors have not identified evolved features of human nature, despite not having provided an adequate account for those features, or that regulation of individual behavior irrelevant. The problem here may be due to that failure, however, because in order to understand fairness as an inherent concern for human

beings, you have to understand how that concern evolved. None of the moral psychologists appear to treat humans as ultrasocial beings. They treat us as transactional beings, or those with strong social connections and who learn and develop through social interaction but they appear to be so focused on how moral concerns play out in our individual behavior and development that they don't contextualize that moral functioning in terms of our ultrasociality.

After looking at the research and literature across moral psychology, it seems that the authors from each theoretical perspective make the same mistake in their view of morality. They are limited in their accounting for human beings as socially constituted, with social identities, and as profoundly attuned to group norms. These authors appear to arrive at their emphasis on fairness without making an adequate evolutionary case. Such a case would necessarily include an emphasis on cooperation and that it emerged in such a way that everyone has an interest in cooperation, which makes such cooperation a communal concern, not simply an individual or dyadic one. In other words, human beings are not simply concerned about how another individual treats them, but also how that individual treats others and therefore whether cooperation is something they can expect in their group. In addition, that communal nature helps us to see that morality is about group norms, not just about individual behavior. The difference is that these norms help to define and structure the group as a unit, not just to regulate individual behavior.

Even Moral Foundations theorists see the source of moral functioning as instinctual, which is a bit more correct, but they don't describe that instinct in terms of its deeply social nature. In other words, it is not just that humans are paying attention to whether there is fairness in their own interpersonal exchanges; it is that humans are

paying attention to whether their entire group and world are constituted by an ethic of fairness. A similar statement could be made about the purity foundation, as well as about hierarchy and authority. It is, of course, accurate to state that we have to figure out whether we are complying with the norms and fitting in with our social group correctly, but there is significantly more to that process.

That we develop and maintain those norms together as collective groups is what is important. Without this view, one could see the norms, ethics, rules, or morality as things that are imposed onto human beings either *tabula rasa* or as necessities for shaping inherently evil and nasty beings into moral ones or for the sake of a social contract that arises among pre-existing, isolated individuals. None of these interpretations are accurate. Most human beings appear desperate to be moral; they go to great lengths to make ethical choices and to live in a way they see as good and consistent with the beliefs and standards set by their group. Humans seem to want to be moral if we can possibly manage it.

I have argued in this dissertation that one way of beginning to build an adequate explanation of ethical mindedness is to examine cooperation; the connections between cooperation, cheater detection, indirect reciprocity, and third party punishment; and to ultimately connect this to a communal understanding of fairness. From this view, ethical mindedness is a product of our intense sociality; and our propensity to be concerned about morality at every social level because it emerges from our connections to communities and groups.

If it is true, and I believe I have presented promising evidence that suggests it may be, this understanding of the communal nature of our ethical mindedness is critical for

understanding moral functioning. While this view is only one shift in perspective, and there may be many more to be found, it certainly offers a strong justification for studying such ontological questions from an evolutionary perspective. If we are to truly understand our nature, we must understand how that nature emerged.

CONCLUSION AND DIRECTIONS FOR FUTURE RESEARCH

This dissertation began with the basic question of whether human beings are ethically minded and a search for the most appropriate way to answer that fundamentally ontological question about the morality of human nature. Because ethical mindedness describes the propensity to be concerned with morality and ethics at all, it would appear that the first step to understanding moral functioning is to understand ethical mindedness. Because ethical mindedness describes an instinct to learn or an emergent property of the human mind, the first step to understanding it has therefore been to understand how and why this feature of human nature came to be. And to understand how a basic feature of human nature evolved, we must first understand the basics of how one adopts an evolutionary perspective and develops an evolutionary case. Although it seems reasonable to look first toward moral psychologists as the researchers with the most at stake in understanding and building a case for ethical mindedness, it appears that none of these theorists, regardless of their particular theoretical framework, have fully taken on board the ontological concept of ethical mindedness or therefore provided a fulsome evolutionary explanation of moral functioning.

The purpose of this dissertation has been to begin both of those tasks. I began with an examination of how an evolutionary argument can be built based on the necessary criteria for identifying adaptations as indicated by evolutionary psychologists. The criteria utilize an adaptationist approach to explore the reproductive benefits of the proposed adaptation and to demonstrate how it could reliably spread throughout a population; as well as a phylogenetic approach that places the adaptation within an evolutionary timeline by exploring its presence and absence in the species' phylogenetic

relatives. These criteria were specific. They are designed to prevent the sort of just-so stories that evolutionary psychologists have been accused of presenting. If effectively applied, they leave little room to question whether a particular adaptation is in fact an evolved, species-typical feature. Beyond this, they direct us toward the development of an ultimate explanation. Although many existing research identifies proximate explanations for moral functioning, we have yet to establish a full, ultimate account for how and why human morality emerged over evolutionary time.

After clarifying the purpose and criteria for an evolutionary assessment of adaptations, I returned to my original question of ethical mindedness and examined how moral psychologists currently answer whether it appears to be a species-typical feature of human nature. In Chapter 2, I highlighted the limitations of existing explanations and emphasized the key contributions from each theoretical perspective in moral psychology. From Social Cognitive Domain theorists, we saw the initial recognition that the moral domain is ultimately universal, and an adequate theory must be built upon the idea that what counts as ethical mindedness or moral functioning must be present across cultures and time. If we are to understand any human function at such a scope, it almost certainly directs us toward the fact that morality and ethical mindedness are present far beyond interpersonal interactions. This is not to say that human nature obviates culture, but, as I specified in Chapter 1, that culture presents us with a different level of analysis. Evolution leads to culture and culture defines specific group norms, but both the emergence of cultural groups and the explication of group norms are things that humans appear instinctively driven to do. And humans are also impelled to take the particular norms of their culture very seriously.

From moral identity theorists, we gained a strong understanding that the centrality of morality to each individual is ultimately variable, and likely a product of development that is informed by contextual variables and individual choices. The idea that such individual differences will exist despite claims of universality provides us with a framework for understanding ethical mindedness at a more immediate level – as something that emerges at both the largest social level and is present within and throughout the development of each individual.

And finally, from the Moral Foundations theorists, we gained the invaluable contribution of a pluralistic view of moral functioning; or the idea that there are multiple moral domains, each of which must be accounted for from an evolutionary perspective. These theorists have begun doing even this specifically evolutionary work, though their framework for doing so will ultimately require some adjustment to meet the basic criteria for evolutionary argument.

Beyond this, however, the foundations identified by Moral Foundations theorists offer what may be some of the most promising directions for future research. I applied the evolutionary criteria to cooperation and ultimately connected that adaptation to fairness in my work beginning to build a theory of ethical mindedness, but cooperation is just one adaptation. For example, we could also examine hierarchy and how it is that humans have some opposing tendencies toward vertical and egalitarian hierarchy (cf. Boehm, 2008). This idea of an egalitarian hierarchy, also referred to as a reverse dominance hierarchy, is used to describe “an apparent absence of hierarchy [as a] result of followers dominating their leaders rather than vice versa” (Boehm, 1993, p. 228). Although Moral Foundations theorists discuss hierarchy and identify it as one of the five

foundations, they really only discuss vertical hierarchy. In doing so, they likely miss what is uniquely human, which appears to be the tension between vertical and egalitarian hierarchy. This tension is a potentially rich direction for future research and one that is extremely well-suited for an evolutionary approach to how it emerged and why it appears to characterize human behavior while not being well-documented in other species. It is likely that each of the foundations identified in Moral Foundations Theory will offer similarly rich and promising insights about our basic nature when viewed through the lens of evolutionary psychology.

Perhaps the real work and contribution of this dissertation, however, was in Chapters 3 and 4. In Chapter 3, I applied the criteria for evolutionary assessment of adaptations to cooperation. In doing so, I used cooperation as a paradigm account to demonstrate the use of evolutionary criteria as they can be applied to one element of ethical mindedness. Having used these criteria to assess cooperation as a species-typical adaptation, I then followed the evolution of cooperation in order to examine not only its emergence, but its proliferation throughout the species.

In Chapter 4, I followed this evolutionary process from the dilemma of how cooperation emerged as a species-typical adaptation despite the appearance that it is more beneficial to exploit than to cooperate. I suggested that this dilemma may be resolved through the emergence of cheater detection, indirect reciprocity, and third party punishment. These adaptations shifted cooperation from its origins in food sharing and dyadic cooperation to the triadic cooperation of indirect reciprocity and ultimately to the idea of communal, group-based nature of cooperation that manifests in third party punishment. Put in other words, this process illustrated the shift from cooperation to trust

among others and ultimately an expectation of fairness in one's group, which opened the doors to a case for ethical mindedness.

Ethical mindedness as a propensity to be concerned with morality and ethics, or with matters of fairness, thus emerges from the communal nature of cooperation and the ultrasociality of the human species. This is a potentially significant contribution to our understanding of moral functioning because it frames human morality as being fundamentally about our group, not about each of us as individuals or, perhaps more importantly, not about regulating a fundamentally amoral human nature. What matters here is our communal nature, our attachment and commitment to our group, and the deep concern we have for the fairness we can expect within that group. It suggests the possibility that we are not just moral because our groups demand it of us; but our groups may also be moral because we demand that of them.

If the ultimate goal for this research was to begin thinking about ethics in a richer way and to understand that our morality and ethics are ultimately as biological as are our vision and opposable thumbs, then this is the promising possibility that emerges: We may be as fundamentally ethically minded as we are bipedal, and that ethical nature may have emerged through the evolution of our deep connections and dependence upon one another. There is reasonable evidence against an evolutionary story that leads to a case that humans are selfishly individualistic beings, and this suggests we may not ever flourish in a social or political system that requires us to act against our nature.

I mentioned in my introduction the issue of moral education, or that without an adequate understanding of own inherited ethical mindedness, the development of effective moral education is seriously in question. Perhaps this is an issue worth

exploring: That so often we view moral education as an imposition of moral expectations and rules; one that exists to regulate a fundamentally amoral nature. But what if human beings are not fundamentally amoral? What if, instead, human beings are conflicted – too often experiencing basic needs as pitted against a fundamental social connectedness and ethical mindedness? Perhaps we see today the consequences of pitting our basic survival against our desire for connection and egalitarianism. Both are essential to human welfare, and the tendency to set up a zero sum game between them is likely to be fraught with pain and turmoil.

While a thorough exploration is beyond the scope of this dissertation, it bears mentioning that many critics of psychology have focused heavily on the Western influence and excessive focus on individuals to the neglect of communal features of humans (Burge, 1986; Parker, 2007). These same critics have also made powerful arguments for focusing on the communal nature of human beings (Burge, 1986; Parker, 2007). It seems likely that these same critiques could be made of moral psychology, but these critiques would almost certainly be well-supported by this view of ethical mindedness and the outcome of evolutionary reasoning.

This dissertation began with a basic, ontological question: Are human beings ethically minded? The answer, it appears, is not simply yes. It is yes, but not without shifting our focus to understand how we came to be that way. And one conclusion that we may derive from this reasoning is this: We may be ethically minded because we are deeply and inextricably connected to our social worlds. Perhaps the issues we face today exist less because individuals are poorly regulated, and instead because we are poorly connected; because our reality has shifted so far from our evolved nature that it has

become difficult for us to express that nature in the ways we evolved to do. It seems telling, after all, that those of us who have self-selected to study that nature seem so confused by the possibility that it is good.

REFERENCES

- Ambrose, S. H. (2001). Paleolithic technology and human evolution. *Science*, *291*, 1748–1753. <https://doi.org/10.1126/science.1059487>
- Aquino, K., & Freeman, D. (2009). Moral identity in business situations: A social-cognitive framework for understanding moral functioning. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, identity, and character: Explorations in moral psychology* (pp. 375-395). New York, NY: Cambridge University Press. <https://doi.org/10.1017/CBO9780511627125.018>
- Aquino, K., Freeman, D., Reed, I. I., Lim, V. K., & Felps, W. (2009). Testing a social-cognitive model of moral behavior: The interactive influence of situations and moral identity centrality. *Journal of Personality and Social Psychology*, *97*(1), 123. <https://doi.org/10.1037/a0015406>
- Aquino, K., Reed, A., & Diener, Ed. (2002). The Self-Importance of Moral Identity. *Journal of Personality and Social Psychology*, *83*(6), 1423-1440.
- Bachner-Melman, R., Gritsenko, I., Nemanov, L., Zohar, A. H., Dina, C., & Ebstein, R. P. (2005). Dopaminergic polymorphisms associated with self-report measures of human altruism: a fresh phenotype for the dopamine D4 receptor. *Molecular Psychiatry*, *10*(4), 333. <https://doi.org/10.1038/sj.mp.4001635>
- Blake, P. R., McAuliffe, K., Corbit, J., Callaghan, T. C., Barry, O., Bowie, A., ... & Wrangham, R. (2015). The ontogeny of fairness in seven societies. *Nature*, *528*(7581), 258. <https://doi.org/10.1038/nature15703>
- Blasi, A. (1990). How should psychologists define morality? Or, the negative side effects of philosophy's influence on psychology. In T. E. Wren (Ed.), *The moral domain: Essays in the ongoing discussion between philosophy and the social sciences* (pp. 38-70). Cambridge, MA: The MIT Press.
- Blasi, A. (1999). Emotions and moral motivation. *Journal for the Theory of Social Behaviour*, *29*(1), 1-19. <https://doi.org/10.1111/1468-5914.00088>
- Blasi, A. (2009). The moral functioning of mature adults and the possibility of fair moral reasoning. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, identity, and character: Explorations in moral psychology* (pp. 396-440). New York, NY: Cambridge University Press. <https://doi.org/10.1017/CBO9780511627125.019>
- Boehm, C. (2000). Conflict and the evolution of social control. *Journal of Consciousness Studies*, *7*, 79–101. <https://doi.org/10.1177/1069397108320422>

- Boehm, C. (2008). Purposive social selection and the evolution of human altruism. *Cross-Cultural Research*, 42(4), 319-352.
- Boehm, C., Barclay, H., Dentan, R., Dupre, M., Hill, J., Kent, S., . . . Rayner, S. (1993). Egalitarian behavior and reverse dominance hierarchy. *Current Anthropology*, 34(3), 227-254. <https://doi.org/10.1086/204166>
- Bowles, S., & Gintis, H. (2004). The evolution of strong reciprocity: Cooperation in heterogeneous populations. *Theoretical Population Biology*, 65, 17–28. <https://doi.org/10.1016/j.tpb.2003.07.001>
- Bunge, M. (1977). Emergence and the mind. *Neuroscience*, 2(4), 501-509. [https://doi.org/10.1016/0306-4522\(77\)90047-1](https://doi.org/10.1016/0306-4522(77)90047-1)
- Burge, T. (1986). Individualism and psychology. *The Philosophical Review*, 95(1), 3-45. <https://doi.org/10.2307/2185131>
- Burkart, J. M., Hrdy, S. B., & van Schaik, C. P. (2009). Cooperative breeding and human cognitive evolution. *Evolutionary Anthropology*, 18, 175–186. <https://doi.org/10.1002/evan.20222>
- Buss, D. M. (Ed.). (2005). *The handbook of evolutionary psychology*. Hoboken, NH: John Wiley & Sons.
- Brinkmann, S. (2009). Facts, values, and the naturalistic fallacy in psychology. *New Ideas in Psychology*, 27, 1-17. <https://doi.org/10.1016/j.newideapsych.2007.10.001>
- Brosnan, S. F., & de Waal, F. B. (2003). Monkeys reject unequal pay. *Nature*, 425(6955), 297. <https://doi.org/10.1038/nature01963>
- Bruner, J. (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.
- Carlo, G. (2006) Care-based and altruistically based morality. In M. Killen & J. G. Smetana (Eds.), *The handbook of moral development* (pp. 551-579). Mahwah, NJ: Lawrence Erlbaum Associates.
- Clark, M. S., & Aragón, O. R. (2013). Communal (and other) relationships: History, theory development, recent findings, and future directions. In J. A. Simpson & L. Campbell (Eds.), *The Oxford handbook of close relationships* (pp. 255-280). New York: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780195398694.013.0012>
- Clutton-Brock, T. H., & Parker, G. A. (1995). Punishment in animal societies. *Nature*, 373, 209–216. <https://doi.org/10.1038/373209a0>

- Colby, A., & Damon, W. (1993). The uniting of self and morality in the development of extraordinary moral commitment. In Noam, G. G., & Wren, T. E. (Eds), *The moral self* (pp. 149-174.). Cambridge, MA: MIT Press.
- Corbit, J., McAuliffe, K., Callaghan, T. C., Blake, P. R., & Warneken, F. (2017). Children's collaboration induces fairness rather than generosity. *Cognition*, *168*, 344-356. <https://doi.org/10.1016/j.cognition.2017.07.006>
- Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, *31*, 187-276. [https://doi.org/10.1016/0010-0277\(89\)90023-1](https://doi.org/10.1016/0010-0277(89)90023-1)
- Cosmides, L. & Tooby, J. (1997). *Evolutionary Psychology: A Primer*. Retrieved from <https://www.cep.ucsb.edu/primer.html>.
- Cosmides, L. & Tooby, J. (1992). Introduction: Evolutionary psychology and conceptual integration. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 3-15). New York, NY: Oxford University Press.
- Cosmides, L. & Tooby, J. (1992). The psychological foundations of culture. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 19-136). New York, NY: Oxford University Press.
- Cosmides, L., & Tooby, J. (2005). Neurocognitive adaptations designed for social exchange. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 584-627). Hoboken, NJ: Wiley. <https://doi.org/10.1002/9780470939376.ch20>
- Crockett, M. J., Clark, L., Hauser, M. D., & Robbins, T. W. (2010). Serotonin selectively influences moral judgment and behavior through effects on harm aversion. *Proceedings of the National Academy of Sciences*, *107*(40), 17433-17438. <https://doi.org/10.1073/pnas.1009396107>
- Crockett, M. J., Clark, L., Tabibnia, G., Lieberman, M. D., & Robbins, T. W. (2008). Serotonin modulates behavioral reactions to unfairness. *Science*, *320*(5884), 1739-1739. <https://doi.org/10.1126/science.1155577>
- Cushman, P. (1990). Why the self is empty. *American Psychologist*, *45*, 599-611. <https://doi.org/10.1037/0003-066X.45.5.599>
- Danziger, K. (1990). *Constructing the subject: Historical origins of psychological research*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511524059>

- Darnell, C., Gulliford, L., Kristjánsson, K., & Paris, P. (2019). Phronesis and the knowledge-action gap in moral psychology and moral education: A new synthesis? *Human Development*, 1-29. <https://doi.org/10.1159/000496136>
- Darwin, C. (1936). *The descent of man* (2nd ed.). New York: Random House.
- Darwin, C. (1871/2010). *The descent of man, and selection in relation to sex*. Mineola, NY: Dover Publications. <https://doi.org/10.5962/bhl.title.24784>
- de Quervain, D. J.-F., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., & Fehr, E. (2004). The neural basis of altruistic punishment. *Science*, 305, 1254–1258. <https://doi.org/10.1126/science.1100735>
- de Waal, F. (1996). *Good natured: The origins of right and wrong in humans and other animals*. Cambridge, MA: Harvard University Press.
- de Waal, F. (1997). The chimpanzee's service economy: Food for grooming. *Evolution and Human Behavior*, 18(6), 375-386. [https://doi.org/10.1016/S1090-5138\(97\)00085-8](https://doi.org/10.1016/S1090-5138(97)00085-8)
- de Waal, F. (2006). Morally evolved: Primate social instincts, human morality, and the rise and fall of “vener theory”. In S. Macedo & J. Ober (Eds.), *Primates and philosophers: How morality evolved* (pp. 1-80). New York, NY: Princeton University Press. <https://doi.org/10.1515/9781400830336>
- de Waal, F. (2007). With a little help from a friend. *PLoS Biology*, 5(7), e190. <https://doi.org/10.1371/journal.pbio.0050190>
- Dewey, J. (1926). *Experience and nature*. LaSalle, Ind.: Open Court.
- Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, 425(6960), 785. <https://doi.org/10.1038/nature02043>
- Fowers, B. J. (2010). Instrumentalism and psychology: Beyond using and being used. *Theory & Psychology*, 20(1), 102-124. <https://doi.org/10.1177/0959354309346080>
- Fowers, B. (2015). *The evolution of ethics: Human sociality and the emergence of ethical mindedness*. New York, NY: Palgrave Macmillan.
- Gibbs, J. C. (2019). *Moral development and reality: Beyond the theories of Kohlberg, Hoffman, and Haidt* (3rd ed.). New York, NY: Oxford University Press. <https://doi.org/10.1093/oso/9780190878214.001.0001>
- Gigerenzer, G., & Hug, K. (1992). Domain-specific reasoning: Social contracts, cheating, and perspective change. *Cognition*, 43(2), 127–171. [https://doi.org/10.1016/0010-0277\(92\)90060-U](https://doi.org/10.1016/0010-0277(92)90060-U)

- Graham, J., Haidt, J., & Rimm-Kaufman, S. E. (2008). Ideology and intuition in moral education. *International Journal of Developmental Science*, 2(3), 269-286.
- Graham, J. P., Koleva, S. H., Iyer, R., Haidt, J., Motyl, M., Wojcik, S., & Ditto, P. (2013). Moral Foundations Theory: The pragmatic validity of moral pluralism. In *Advances in Experimental Social Psychology* (Vol. 47) (pp. 55-130). Academic Press. <https://doi.org/10.1016/B978-0-12-407236-7.00002-4>
- Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, 101, 366-385. <https://doi.org/10.1037/a0021847>
- Greene, J. D., & Paxton, J. M. (2009). Patterns of neural activity associated with honest and dishonest moral decisions. *Proceedings of the National Academy of Sciences*, 106, 12506–12511. <https://doi.org/10.1073/pnas.0900152106>
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814-834. <https://doi.org/10.1037/0033-295X.108.4.814>
- Haidt, J. (2004). The emotional dog gets mistaken for a possum. *Review of General Psychology*, 8, 283-290. <https://doi.org/10.1037/1089-2680.8.4.283>
- Haidt, J. (2007). The new synthesis in moral psychology. *Science*, 316(5827), 998-1002. <https://doi.org/10.1126/science.1137651>
- Haidt, J. (2010). Moral psychology must not be based on faith and hope: Commentary on Narvaez (2010). *Perspectives on Psychological Science*, 5(2), 182-184. <https://doi.org/10.1177/1745691610362352>
- Haidt, J., & Bjorklund, F. (2008). Social intuitionists answer six questions about moral psychology. In W. Sinnott-Armstrong (Ed.), *Moral psychology: Vol. 2. The cognitive science of morality: Intuition and diversity* (pp. 181-218). Cambridge, MA: MIT Press.
- Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, 20, 98-116. <https://doi.org/10.1007/s11211-007-0034-z>
- Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Dædalus*, 133, 55-66. <https://doi.org/10.1162/0011526042365555>

- Haidt, J., & Joseph, C. (2007). The moral mind: How five sets of innate intuitions guide the development of many culture-specific virtues, and perhaps even modules. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind: Vol. 3. Foundations and the future* (pp. 367-392). New York, NY: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780195332834.003.0019>
- Haidt, J., & Kesebir, S. (2010). Morality. In S. Fiske, D. Gilbert, & G. Lindzey (Eds.). *Handbook of social psychology* (5th Ed., pp. 797-832). Hoboken, NJ: Wiley.
<https://doi.org/10.1002/9780470561119.socpsy002022>
- Hamann, K., Warkeken, F., Greenberg, J. R., & Tomasello, M. (2011). Collaboration encourages equal sharing in children but not in chimpanzees. *Nature*, *476*, 328–331.
<https://doi.org/10.1038/nature10278>
- Hamilton, W. D. (1964). The genetical evolution of social behavior I. *Journal of Theoretical Biology*, *7*, 1-16. [https://doi.org/10.1016/0022-5193\(64\)90038-4](https://doi.org/10.1016/0022-5193(64)90038-4)
- Hamilton, W. D. (1964). The genetical evolution of social behavior II. *Journal of Theoretical Biology*, *7*, 17-52. [https://doi.org/10.1016/0022-5193\(64\)90039-6](https://doi.org/10.1016/0022-5193(64)90039-6)
- Hammond, R. A., & Axelrod, R. (2006). The evolution of ethnocentrism. *The Journal of Conflict Resolution*, *50*(6), 926–936. <https://doi.org/10.1177/0022002706293470>
- Hardy, S. A., & Carlo, G. (2005). Identity as a source of moral motivation. *Human Development*, *48*(4), 232-256. <https://doi.org/10.1159/000086859>
- Hardy, S. & Carlo, G. (2011). Moral identity: What is it, how does it develop, and is it linked to moral action? *Child Development Perspectives* *5*(3), 212 – 218.
<https://doi.org/10.1111/j.1750-8606.2011.00189.x>
- Hardy, S. A. & Carlo, G. (2011). Moral identity. In S. J. Schwartz, K. Luyckx, & V. L. Vignoles (Eds.), *Handbook of identity theory and research* (pp. 495-513). New York, NY: Springer. https://doi.org/10.1007/978-1-4419-7988-9_19
- Hart, D., Atkins, R., & Ford, D. (1998). Urban America as a context for the development of moral identity in adolescence. *Journal of Social Issues*, *54*(3), 513-530.
<https://doi.org/10.1111/j.1540-4560.1998.tb01233.x>
- Helwig, C. C., & Turiel, E. (2011). Children's social and moral reasoning. In P. K. Smith & C. H. Hart (Eds.), *The Wiley-Blackwell handbook of childhood social development* (2nd ed., pp. 567-58). Malden, MA: Wiley-Blackwell.
<https://doi.org/10.1002/9781444390933.ch30>

- Henrich, N., & Gil-White, F. J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior*, *22*, 165–196. [https://doi.org/10.1016/S1090-5138\(00\)00071-4](https://doi.org/10.1016/S1090-5138(00)00071-4)
- Henrich, N., & Henrich, J. (2007). *Why humans cooperate: A cultural and evolutionary explanation*. Oxford, England: Oxford University Press.
- Hill, K. (2002). Altruistic cooperation during foraging by the Ache and the evolved human predisposition to cooperate. *Human Nature*, *13*, 105–128. <https://doi.org/10.1007/s12110-002-1016-3>
- Hill, P. L. & Lapsley, D. K., (2009). The ups and downs of the moral personality: Why it's not so black and white. *Journal of Research in Personality*, *43*(3), 520-523. <https://doi.org/10.1016/j.jrp.2009.01.014>
- Hill, K. R., Walker, R. S., Bozicevic, M., Elder, J., Headland, T., Hewlett, B., ... & Wood, B. (2011). Co-residence patterns in hunter-gatherer societies show unique human social structure. *Science*, *331*, 1286–1289. <https://doi.org/10.1126/science.1199071>
- House, Bailey R., Silk, Joan B., Henrich, Joseph, Barrett, H. Clark, Scelza, Brooke A., Boyette, Adam H., . . . Laurence, Stephen. (2013). Ontogeny of prosocial behavior across diverse societies. *Proceedings of the National Academy of Sciences of the United States*, *110*(36), 14586-14591. <https://doi.org/10.1073/pnas.1221217110>
- Jambon, M., & Smetana, J. G. (2014). Moral complexity in middle childhood: Children's evaluations of necessary harm. *Developmental Psychology*, *50*(1), 22. <https://doi.org/10.1037/a0032992>
- just-so story. 2019. In Merriam-Webster.com. Retrieved June 3, 2019, from <https://www.merriam-webster.com/dictionary/just-so%20story>
- Kaplan, H., & Hill, K. (1985). Food sharing among Ache foragers: Tests of explanatory hypotheses. *Current Anthropology*, *26*, 223–246. <https://doi.org/10.1086/203251>
- Kaplan, H., Hill, K., Lancaster, J., & Hurtado, A. M. (2000). A theory of human life history evolution: Diet, intelligence, and longevity. *Evolutionary Anthropology*, *9*, 156–185. [https://doi.org/10.1002/1520-6505\(2000\)9:4<156::AID-EVAN5>3.0.CO;2-7](https://doi.org/10.1002/1520-6505(2000)9:4<156::AID-EVAN5>3.0.CO;2-7)
- Keltner, D., Kogan, A., Piff, P. K., & Saturn, S. R. (2014). The sociocultural appraisals, values, and emotions (SAVE) framework of prosociality: Core processes from gene to meme. *Annual Review of Psychology*, *65*, 425-460. <https://doi.org/10.1146/annurev-psych-010213-115054>

- Kiley Hamlin, J., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*(7169), 557. <https://doi.org/10.1038/nature06288>
- Kiley Hamlin, J., Wynn, K., & Bloom, P. (2010). Three-month-olds show a negativity bias in their social evaluations. *Developmental Science*, *13*(6), 923-929. <https://doi.org/10.1111/j.1467-7687.2010.00951.x>
- Kiley Hamlin, J., Wynn, K., Bloom, P., & Mahajan, N. (2011). How infants and toddlers react to antisocial others. *Proceedings of the National Academy of Sciences*, *108*(50), 19931-19936. <https://doi.org/10.1073/pnas.1110306108>
- Killen, M., & de Waal, F. B. (2000). The evolution and development of morality. In F. Aureli & F. B. M. de Waal (Eds.), *Natural conflict resolution* (pp. 352-372). Berkeley and Los Angeles, CA: University of California Press.
- Killen, M., Margie, N. G., & Sinno, S. (2006). Morality in the context of intergroup relationships. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 155-183). Mahwah, NJ: Lawrence Erlbaum Associates Publishers. <https://doi.org/10.4324/9781410615336>
- Klein, R. G. (2000). Archeology and the evolution of human behavior. *Evolutionary Anthropology*, *9*, 17-36. [https://doi.org/10.1002/\(SICI\)1520-6505\(2000\)9:1<17::AID-EVAN3>3.0.CO;2-A](https://doi.org/10.1002/(SICI)1520-6505(2000)9:1<17::AID-EVAN3>3.0.CO;2-A)
- Kogan, A., Saslow, L. R., Impett, E. A., Oveis, C., Keltner, D., & Saturn, S. R. (2011). Thin-slicing study of the oxytocin receptor (OXTR) gene and the evaluation and expression of the prosocial disposition. *Proceedings of the National Academy of Sciences*, *108*(48), 19189-19192. <https://doi.org/10.1073/pnas.1112658108>
- Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature*, *435*(7042), 673. <https://doi.org/10.1038/nature03701>
- Lapsley, D. K. (1996). *Moral psychology*. Boulder, CO: Westview Press.
- Lapsley, D. K., & Hill, P. L. (2009). The development of the moral personality. In Narvaez, D., & Lapsley, D. K. (Eds.), *Personality, identity and character: Explorations in moral psychology* (pp. 185-213). <https://doi.org/10.1017/CBO9780511627125.009>
- Maestripieri, D., & Roney, J. R. (2006). Evolutionary developmental psychology: Contributions from comparative research with nonhuman primates. *Developmental Review*, *26*, 120-137. <https://doi.org/10.1016/j.dr.2006.02.006>

- Mainzer, K. (2008). The emergence of mind and brain: An evolutionary, computational, and philosophical approach. *Progress in Brain Research*, 168, 115-132.
[https://doi.org/10.1016/S0079-6123\(07\)68010-8](https://doi.org/10.1016/S0079-6123(07)68010-8)
- Mason, E. (2018). Value pluralism. In E. Zalta (Ed.), *Stanford encyclopedia of philosophy*. Retrieved from <https://plato.stanford.edu/>
- Melis, A. P., Hare, B., & Tomasello, M. (2006). Chimpanzees recruit the best collaborators. *Science*, 311(5765), 1297-1300.
<https://doi.org/10.1126/science.1123007>
- Narvaez, D. (2008). Human flourishing and moral development: Cognitive and neurobiological perspectives of virtue development. In L. P. Nucci & D. Narvaez (Eds.), *Handbook of moral and character education* (pp. 310-327). New York, NY: Routledge.
- Narvaez, D. (2008). Triune ethics: The neurobiological roots of our multiple moralities. *New Ideas in Psychology*, 26, 95-119.
<https://doi.org/10.1016/j.newideapsych.2007.07.008>
- Narvaez, D. (2010). Moral complexity: The fatal attraction of truthiness and the importance of mature moral functioning. *Perspectives on Psychological Science*, 5(2), 163-181. <https://doi.org/10.1177/1745691610362351>
- Narvaez, D. (2013). Natural morality, moral natures and human flourishing. In B. Musschenga & A. van Harskamp (Eds.), *What makes us moral? On the capacities and conditions for being moral* (pp. 239-254). New York, NY: Springer.
https://doi.org/10.1007/978-94-007-6343-2_14
- Narvaez, D. (2014). *Neurobiology and the development of human morality: Evolution, culture, and wisdom*. New York, NY: WW Norton & Company.
- Nowak, M. A., Page, K. M., & Sigmund, K. (2000). Fairness versus reason in the ultimatum game. *Science*, 289, 1773-1775.
<https://doi.org/10.1126/science.289.5485.1773>
- Nucci, L. P. (2001). *Education in the moral domain*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511605987>
- Nucci, L., & Gingo, M. (2011). The development of moral reasoning. In Goswami, U. (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development* (2nd ed., pp. 420-445). Malden, MA: Wiley-Blackwell.
<https://doi.org/10.1002/9781444325485.ch16>

- Nucci, L. & Turiel, E. (2000). The moral and the personal: Sources of social conflicts. In L. Nucci, G.B. Saxe, & E. Turiel (Eds.), *Culture, thought, and development* (pp. 115-137). New York, NY: Psychology Press.
- Nucci, L., & Turiel, E. (2009). Capturing the complexity of moral development and education. *Mind, Brain, and Education*, 3(3), 151-159.
<https://doi.org/10.1111/j.1751-228X.2009.01065.x>
- Olson, K. R., & Spelke, E. S. (2008). Foundations of cooperation in young children. *Cognition*, 108, 222–231. <https://doi.org/10.1016/j.cognition.2007.12.003>
- Parker, I. (2007). *Revolution in psychology: Alienation to emancipation*. Ann Arbor, MI: Pluto Press.
- Pinker, S. & Bloom, P. (1992). Natural language and natural selection. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 451-494). New York, NY: Oxford University Press.
- Richardson, F. C., Fowers, B. J., & Guignon, C. (1999). *Re-envisioning psychology: Moral dimensions of theory and practice*. San Francisco: Jossey-Bass.
- Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, 300, 1755–1758. <https://doi.org/10.1126/science.1082976>
- Reuter, M., Frenzel, C., Walter, N. T., Markett, S., & Montag, C. (2010). Investigating the genetic basis of altruism: the role of the COMT Val158Met polymorphism. *Social Cognitive and Affective Neuroscience*, 6(5), 662-668.
<https://doi.org/10.1093/scan/nsq083>
- Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S., & Kilts, C. D. (2002). A neural basis for social cooperation. *Neuron*, 35, 395–405.
[https://doi.org/10.1016/S0896-6273\(02\)00755-9](https://doi.org/10.1016/S0896-6273(02)00755-9)
- Rochat, P., Dias, Guo Liping, Broesch, Passos-Ferreira, Winning, & Berg. (2009). Fairness in distributive justice by 3- and 5-year-olds across seven cultures. *Journal of Cross-Cultural Psychology*, 40(3), 416-442.
<https://doi.org/10.1177/0022022109332844>
- Rodrigues, S. M., Saslow, L. R., Garcia, N., John, O. P., & Keltner, D. (2009). Oxytocin receptor genetic variation relates to empathy and stress reactivity in humans. *Proceedings of the National Academy of Sciences*, 106(50), 21437-21441.
<https://doi.org/10.1073/pnas.0909579106>
- Roth, A., Panikar, V., Okuno-Fujiwara, M., & Zamir, S. (1991). Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh and Tokyo: An experimental study. *American Economic Review*, 81, 1068–1095.

- Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science, 300*, 1755–1758. <https://doi.org/10.1126/science.1082976>
- Smetana, J. (2006). Social–cognitive domain theory: Consistencies and variations in children's moral and social judgments. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 137-172). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Smetana, J. G., Jambon, M., & Ball, C. (2013). The social domain approach to children's moral and social judgments. In M. Killen & J. G. Smetana (Eds.), Killen, M., & Smetana, J. G. (2013). *Handbook of moral development* (2nd ed., pp. 23-45). New York, NY: Psychology Press.
- Tabibnia, G., Satpute, A. B., & Lieberman, M. D. (2008). The sunny side of fairness: Preference for fairness activates reward circuitry (and disregarding unfairness activates self-control circuitry). *Psychological Science, 19*, 339–347. <https://doi.org/10.1111/j.1467-9280.2008.02091.x>
- Tisak, M. S., Tisak, J., & Goldstein, S. E. (2006). Aggression, delinquency, and morality: A social-cognitive perspective. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 611-629). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Turiel, E. (2002). *The culture of morality: Social development, context, and conflict*. New York, NY: Cambridge University Press. <https://doi.org/10.1017/CBO9780511613500>
- Turiel, E. (2003). Resistance and subversion in everyday life. *Journal of Moral Education, 32*(2), 115-130. <https://doi.org/10.1080/0305724032000072906>
- Turiel, E., & Killen, M. (2010). Taking emotions seriously: The role of emotions in moral development. Emotions, aggression, and morality in children: Bridging development and psychopathology. In W. F. Arsenio & E. A. Lemerise (Eds.), *Emotions, aggression, and morality in children: Bridging development and psychopathology* (pp. 33-52). Washington, DC: American Psychological Association. <https://doi.org/10.1037/12129-002>
- Walker, L. J. (2004). Progress and prospects in the psychology of moral development. *Merrill-Palmer Quarterly, 50*(4), 546-557. <https://doi.org/10.1353/mpq.2004.0038>
- Walker, L. J. (2004). What does moral functioning entail? In T. A. Thorkildsen & H. J. Walberg (Eds.), *Nurturing morality* (pp. 3-17). New York, NY: Springer Science. https://doi.org/10.1007/978-1-4757-4163-6_1

- Warneken, F., Lohse, K., Melis, A. P., & Tomasello, M. (2011). Young children share the spoils after collaboration. *Psychological Science, 22*, 267–273. <https://doi.org/10.1177/0956797610395392>
- Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development, 77*, 640–643. <https://doi.org/10.1111/j.1467-8624.2006.00895.x>
- Wason, P. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology* (pp. 135–151). Harmondsworth, England: Penguin.
- West, S. A., El Mouden, C., & Gardner, A. (2011). Sixteen common misconceptions about the evolution of cooperation in humans. *Evolution and Behavior, 32*, 231–262. <https://doi.org/10.1016/j.evolhumbehav.2010.08.001>
- West, S. A., Griffin, A. S., & Gardner, A. (2007). Evolutionary explanations for cooperation. *Current Biology, 17*(16), R661-R672. <https://doi.org/10.1016/j.cub.2007.06.004>
- Wilkinson, G. S. (1988). Reciprocal altruism in bats and other mammals. *Ethology & Sociobiology, 9*, 85–100. [https://doi.org/10.1016/0162-3095\(88\)90015-5](https://doi.org/10.1016/0162-3095(88)90015-5)
- Wilkinson, G. S. (1990). Food sharing in vampire bats. *Scientific American, 262*, 76–82. <https://doi.org/10.1038/scientificamerican0290-76>
- Williams, G. C., & Williams, D. C. (1957). Natural selection of individually harmful social adaptations among sibs with special reference to social insects. *Evolution, 77*, 249-253.
- Wilson, E. O. (1975). *Sociobiology*. Cambridge, MA: Harvard University Press. <https://doi.org/10.1111/j.1558-5646.1957.tb02873.x>
- Zeidler, H., Herrmann, E., B. M. Haun, D., & Tomasello, M. (2016). Taking turns or not? Children's approach to limited resource problems in three different cultures. *Child Development, 87*(3), 677-688. <https://doi.org/10.1111/cdev.12505>