THE EVOLUTION OF RELIGION
AND THE
EVOLUTION OF CULTURE

by

TAYLOR THIEL DAVIS

B.Sc., The University of Georgia, 2000
M.A., Tufts University, 2011

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Philosophy)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

July 2014

© Taylor Thiel Davis, 2014
Abstract

The emerging science of religious evolution (the evolution of traits that distinguish religious individuals from non-religious ones) and the emerging science of cultural evolution have recently entered into a reciprocal relationship, each having something to offer the other. The theory of cultural evolution offers the field of religious evolution a powerful set of models and concepts for explaining important traits and facts that are not explained by genetic evolution. But theories of cultural evolution face their own important challenges, and theorists within the field do not agree about how cultural evolution itself should be explained, and focusing on religion makes some abstract and difficult questions in this domain more concrete and tractable. Thus, the field of religious evolution also offers the theory of cultural evolution a way of clarifying its commitments, and of demonstrating its ability to respond to important challenges.

This dissertation addresses both sides of this reciprocal relationship, taking advantage of the opportunity to develop at the same time both a better understanding of the nature of religion and a better understanding of the nature of cultural phenomena in general. One goal, then, is to address philosophical, foundational questions about what religion is from within the scientific worldview. I address this general goal in two independent articles, which comprise Chapters 2 and 3. In the fourth chapter, however, I pursue a different goal, extracting from the study of religion a methodological lesson that applies for the study of cultural phenomena in general. Twenty years ago adaptationist theories in psychology appealed almost exclusively to genetic selection at the individual level, but developments since then have caused a growing number of scientists to suspect that this is too narrow a view of human evolution. I argue that the study of religion confirms these suspicions. Thus, by examining religion from the evolutionary
perspective, we learn not just about the nature of religion, but also about the nature of the evolutionary perspective itself.
Preface

The body of this dissertation consists of three independent research articles (Chapters 2-4), along with an introduction (Chapter 1) and conclusion (Chapter 5). The articles comprising Chapters 2 and 3 have been submitted for publication, but are still under review. Because the primary articles address three independent topics within the same focused subject area, some minor redundancies in content occur across chapters. So too do some minor shifts in terminology. However, since each of these chapters has been written as a self-contained essay, the meanings of the relevant terms should remain clear within their particular contexts of use.

All content is the original work of the author.
# Table of Contents

Abstract........................................................................................................................................... ii

Preface................................................................................................................................................ iv

Table of Contents ............................................................................................................................... v

List of Figures .................................................................................................................................... vii

Acknowledgements .......................................................................................................................... viii

1. Introduction .................................................................................................................................... 1
   1.1 Overview................................................................................................................................... 1
   1.2 Basic Motivations, and a Note on Reductionism ................................................................. 4
   1.3 Defining Religion and Culture .............................................................................................. 8
      1.3.1 What is Cultural Evolution? ............................................................................................. 10
      1.3.2 What is Religious Evolution? .......................................................................................... 17
   1.4 The Metaphysics of Cultural Selection .................................................................................... 26
      1.4.1 Making Memes Come True ............................................................................................. 28
      1.4.2 Social Learning Without Imitation ............................................................................... 36
   1.5 Challenges to the Dual-Inheritance Approach ........................................................................ 44
      1.5.1 The Challenge from Evoked Culture Theorists ........................................................... 45
      1.5.2 Is Cultural Selection "Superfluous"? ................................................................................ 53
   1.6 Conclusion .................................................................................................................................. 65

2. Group Selection in the Evolution of Religion: Genetic Evolution or Cultural Evolution? ......... 68
   2.1 Introduction ............................................................................................................................... 68
   2.2 What Is Religious Altruism? ..................................................................................................... 71
   2.3 Haidt's Appeals to Genetic Group Selection .......................................................................... 76
   2.4 Why Group Selection in Humans is Unique ........................................................................... 80
      2.4.1 Misinterpreting Gene-Culture Coevolution .................................................................. 81
      2.4.2 Against Genetic Group Selection in Humans ................................................................. 84
   2.5 Comparing Genetic and Cultural Group Selection in Context .............................................. 88
   2.6 Group Selection in the Evolution of Religion ......................................................................... 92
   2.7 Conclusion ................................................................................................................................ 95

3. The Goldberg Exaptation Model: Unifying Adaptationist and By-product Theories of Religion .... 98
   3.1 Introduction ............................................................................................................................... 98
   3.2 By-product Theories of Religion ............................................................................................ 100
3.3 Adaptationist Theories of Religion .................................................................105
  3.3.1 The General Form of the Adaptationist Approach ................................105
  3.3.2 Where Adaptationist Theories Differ .......................................................107
3.4 Two Different Meanings of ‘Adaptation’ ......................................................111
3.5 Two Different Meanings of ‘Religion’ ..........................................................116
3.6 The Goldberg Exaptation Model ....................................................................119
  3.6.1 Sosis’s Mereological Analysis .................................................................119
  3.6.2 McCauley’s Metaphor and the Role of Exaptation .................................122
3.7 Unifying By-product and Adaptationist Accounts ..........................................126
3.8 Conclusion .......................................................................................................129

4.1 Introduction ......................................................................................................132
4.2 The Crux of the Issue: Genes vs. Culture .....................................................135
  4.2.1 Setting Group Selection Aside..................................................................137
  4.2.2 Evolving Traits and Evolving Environments ..........................................140
4.3 Broad Adaptationist Theories of Morality and Religion ...............................146
  4.3.1 ‘Altruism’ and ‘Prosociality’ ....................................................................149
  4.3.2 The Tribal Instincts Hypothesis ...............................................................153
    4.3.2.1 Cultural Group Selection and the Evolution of the Social Environment ...154
    4.3.2.2 The Culturally Evolved Environment and Moral and Religious Psychology ....159
4.4 Narrow Adaptationist Theories of Morality and Religion ............................164
  4.4.1 Evoked Culture: How Genetic Similarities Produce Cultural Differences ....166
  4.4.2 Proportionality Theories of Religion and Morality ..................................170
4.5 How Are Broad and Narrow Theories Related? ............................................177
4.6 What Needs to be Explained about Morality and Religion? ........................181
  4.6.1 Mapping the Moral Domain ....................................................................183
  4.6.2 Cross-Cultural Variability: Explanation vs. Accommodation ...............187
4.7 Conclusion .......................................................................................................195

5. Conclusion .........................................................................................................202

Bibliography ..........................................................................................................208
List of Figures

Figure 3.1 Narrow vs. Broad Forms of Adaptationist Framework.............................................112
Figure 3.2 The Goldberg Exaptation Model................................................................................123
Figure 4.1 Broad vs. Narrow Forms of Adaptationist Framework.............................................133
Figure 4.2 The Broad Adaptationist Framework *in Practice* ......................................................139
Acknowledgments

I'd like to thank all of the members of my advisory committee for their very helpful advice and comments—not just about the dissertation, but about the world of professional philosophy in general.

I’d also like to thank Prof. Stephen Stich, who first encouraged me to pursue my interest in religion, and who helped to put me in a very good position to do so.

Special thanks go to my advisor, Prof. Eric Margolis, for truly excellent guidance throughout the entire process of writing this dissertation (among many other things). His planning, organization and practical advice—even about philosophical issues—made a difficult and complicated process quite manageable, even under time constraints. I could not ask for a better professional mentor, and I’ll be striving to follow his example in many ways for many years.

Finally, I must thank my wife, Rebecca Trainor, who may have sacrificed more to make this dissertation happen than anyone else, myself included.
Chapter 1

Introduction

1.1 Overview

Over the last two decades, two distinct fields of scientific inquiry have blossomed into mature, thriving disciplines: the study of religious evolution, and the study of cultural evolution. In just the most recent decade, however, these two disciplines have become tightly interwoven, as social scientists from across the fields of psychology, anthropology, biology, sociology, economics and philosophy have begun to see that each has important things to offer the other. This dissertation is about the growing body of theory located at the intersection between these fields, and about the implications this relationship holds for each discipline independently.

The theory of cultural evolution offers the field of religious evolution a powerful set of conceptual tools for explaining important facts not explained by genetic evolution, including facts about the psychological traits of religious individuals, the evolution of particular religious systems, and the relationship obtaining between religion and morality. But theorists disagree about the role cultural evolution plays in explanations of human psychology, and about how cultural evolution itself should be explained, so focusing the lens of cultural evolutionary theory specifically upon the phenomena of religion also helps to make more concrete and tractable some abstract and difficult questions about cultural evolution itself. Thus, the field of religious evolution also offers the theory of cultural evolution a way of clarifying its commitments, responding to important challenges, and demonstrating its explanatory potential.

The dissertation addresses both sides of this reciprocal relationship, taking advantage of the opportunity to develop at the same time both a better understanding of the nature of religion and
a better understanding of the nature of cultural phenomena in general. The body of it—Chapters 2, 3 and 4—is comprised of three independent articles. The first two of these articles address one side of the relationship, focusing primarily on questions about religious psychology. The third article then addresses the other side, focusing on implications of the psychology of religion for methodological questions about the study of cultural evolution. This introductory chapter is primarily devoted to describing the basic motivations of the project, and to laying out the relevant background. In Section 5, however, I consider and reject two significant challenges to the views I defend, one of which concerns the nature of cultural evolution in general, while the other concerns the cultural evolution of religion, specifically.

In Chapter 2 I discuss a debate about the role group selection plays in explaining one trait that is of special importance for understanding the social aspects of religion, including the relationship between religion and morality: altruism. Accounts of altruism based on group selection have received a great deal of attention in recent years, and one of the most prominent defenders of this theory, biologist David Sloan Wilson, has argued that religiously motivated altruism reveals the influence of group selection on human populations. Still more recently, the influential psychologist Jonathan Haidt has adopted Wilson’s account, developing it further and incorporating it into his own theory of moral psychology. But Haidt and Wilson appeal to cultural evolution and genetic evolution together in the same theory, and this means their account faces a crucial question: is group selection supposed to explain culturally inherited religious traits, or genetically inherited religious traits? They answer “both,” but this answer is based on a misunderstanding of the relationship obtaining between genetic evolution and cultural evolution. I argue that when this misunderstanding is addressed, it becomes clear that group selection is significant only in cultural evolution, and not in genetic evolution.
In Chapter 3, I address a more general question, which some have taken to be the fundamental question in the field: is religion an adaptation? The earliest theories answered “no,” arguing that religious beliefs are adaptive by-products, or products of cognitive systems whose adaptive value is based on other, non-religious functions. More recently, however, evidence has accumulated for the view that religion is an adaptation, because of its role in promoting cooperation. This created the now venerable debate between adaptationists and by-product theorists, which every major theorist in the field has eventually been forced to address. I argue, however, that adaptationist and by-product theories are not only consistent, but actually complimentary parts of a single, unified theory. I extract from the existing literature a model of the religious phenotype that has been offered independently (and in very different terms) by theorists from both sides of the adaptation/by-product debate. I then argue that shared assumptions embodied in this model are already sufficient for unifying adaptationist and by-product accounts of religion, and for moving beyond the debate.

In the fourth chapter, I turn to consider what the study of religion in particular can teach us about the role of cultural evolution in psychology more generally. In each of the preceding chapters, I identify a methodological distinction between “broad” and “narrow” forms of adaptationist framework—a division that underlies the debate between adaptationists and by-product theorists. The narrow framework is defined by rejection of appeals to cultural selection and group selection, and all by-product theories of religion are based on this approach. In contrast, the broad framework, on which virtually all adaptation theories of religion are based, includes appeals to cultural evolution and group selection, in addition to appeals to individual-level genetic selection. In Chapter 4 I argue against the narrow framework in general, by identifying important traits and facts about morality and religion that cannot be explained by
genetic evolution, but that can be explained within a broader adaptationist framework. Once these facts have been explicitly identified, it becomes apparent that there is no motivation for imposing the constraints on selection-based theorizing that the narrow framework imposes.

The theory of cultural evolution explains a lot about the origin and function of religion, and by examining how it does so, we stand to learn almost as much about cultural evolution as we do about religion. During the same recent decades in which the evolutionary perspective has become widely adopted as the foundation for an interdisciplinary science of religion, the evolutionary perspective itself has begun to change, at least as it is applied to human psychology. Twenty years ago adaptationist theories in psychology appealed almost exclusively to genetic selection at the individual level, but developments since then have caused a growing number of scientists to suspect that this is too narrow a view of human evolution. I argue here that the study of religion confirms these suspicions.

1.2 Basic Motivations, and a Note on Reductionism

At the most basic, fundamental level, this project is an exploration of the potential for pursuing social science in a manner that is fully continuous with the rest of science. Even when it does not challenge or conflict with the integrated body of theory accumulating in other scientific disciplines, much of social science has proceeded in a manner that is methodologically and theoretically disconnected from the rest of science. Sometimes, as in the case of Weber (1991), for example, this is because researchers’ aims and methods are explicitly “antipositivist,” and are more concerned with normative or phenomenological issues than with empirical issues. At other times, as in the case of Durkheim (1995), the aims and methods may be straightforwardly
scientific, but the ontological and theoretical commitments nevertheless function to isolate the claims made from work in other disciplines. Still other studies in social science, of course, are in no way inconsistent with the aims, methods or theoretical commitments of other scientific disciplines. Nevertheless, they remain entirely disconnected from them. The theories offered are grounded in concepts and constructs that either do not relate at all to theories and concepts from the so-called “hard” sciences, or that leave any potential relationships obscure and unidentified. They are disconnected in practice, if not in principle.

The work contained here is an attempt to bring the full weight of scientific knowledge to bear upon some important phenomena of human social life, and I see evolutionary theory as the most powerful means currently available for doing this. The articles compiled here focus primarily on religion, but they also address questions about other cultural phenomena, such as morality, and about culture itself, as a natural phenomenon. They do so by using evolutionary theory to integrate facts and theories gathered from across a wide range of scientific disciplines. Accordingly, from the point of view of many social scientists my motivations will probably appear reductionist in the most pejorative sense of that term. From my own point of view, however, there is no reason why this should be a pejorative term, and I have three points to make in an effort to bring these two points of view into accord.

First, I do not claim that the materialist ontology, quantitative methods, and objective perspective of science are capable of explaining everything there is to explain about religion, culture, or other aspects of human social life. I take scientific activity in general to be guided by the aim of predicting observable phenomena, and I take scientific theories to be devoted to the aim of providing *predictive* explanations. Many theoretical questions about religion and culture simply aren’t questions of prediction, and I see nothing wrong with such questions. If one’s
theoretical aims involve questions of phenomenology, ethical justification, or epistemic rationality, then I am willing to grant that reducing religion to causal processes and physical facts may not help with those aims. And since I see no general reason to think such questions are unimportant or problematic, I also see no general reason to think that the posits and methods necessary for answering them are unimportant or problematic. These are simply different kinds of questions, which delineate different domains of theoretical activity, and answering them may well require adopting different methods and ontological commitments. Of course, certain reductive explanations may interact with—and conflict with—certain non-reductive explanations, so I don’t mean to suggest that these issues can or should be pursued in isolation from one another. But such questions of interaction and conflict can arise only after the relevant reductive and non-reductive explanations have been provided. My claim is only that, whatever else religion and culture may be, they are also natural phenomena, and reducing them to the terms of scientific theories provides the best predictive theories of those phenomena.

Second, for those theories in social science that do address questions of prediction, it may be that some explain facts that never will be explicable in any terms that have anything to do with evolutionary theory. For other predictive theories, it may be that the data they explain are ultimately explicable in such terms, but that we have yet to develop the theoretical resources for actually doing so. It remains an open question exactly which facts and phenomena can actually be explained—and explained best—in those terms. Unless and until an evolutionary framework can actually be translated into concepts that provide better answers to the very same questions that other theories address, there is no reason why a reductive approach should seem threatening.

Third, in claiming that evolutionary theory provides the best means available for predicting social phenomena, it is crucial to note that by “evolutionary theory” I do not mean the theory of
**genetic** evolution, specifically. Biology in general is dominated by evolutionary theorizing, and this is largely because of the fantastic success of Darwin’s theory of genetic selection, which explains an incredible range of phenomena throughout the living world. Clearly, humans, too, are part of the living world, and their properties and activities can be explained by the same principles that explain the properties and activities of any other species. Yet in a laudable effort to bring these continuities to the fore, traditional applications of evolutionary theory to human psychology (e.g., Wilson 1978; Dawkins 1976; Barkow, Cosmides, and Tooby 1992; Pinker 2002) have emphasized “nature” over “nurture,” overlooking or playing down the role of social learning in accounts of social behavior. In other words, evolutionary theories have tended to reduce cultural phenomena in general to only those features of individuals that are inherited genetically, and explained by genetic selection. My own work is very much a reaction against this approach. Even for strictly predictive purposes, I agree that to reduce the phenomena of society and culture to only what may be explained by genetic selection is to reduce them to something too thin, or narrow. One of my primary motivations is to show that much more can be explained from within an evolutionary framework than this traditional evolutionary approach has recognized.

I agree that genetic selection is important for explaining human social behavior, for the same reasons that it is important for explaining the social behavior of other species. But I disagree that this provides any reason to leave facts about social learning out of an evolutionary account. Learning, too, is a biological phenomenon, and social learning, in particular, is even more important for explaining human behavior than it is for explaining behavior in other species. The theory of cultural evolution I defend here provides a way of organizing and explaining these
facts about learning, a way that can be fully integrated with the standard theory of genetic evolution.

The fact that human evolution can be explained in accordance with the same principles that explain evolution in other species does not imply that human evolution is similar in every respect to the evolution of other species. Even if the same general principles apply universally across species, each particular lineage possesses its own unique evolutionary trajectory, and, as the philosopher Kim Sterelny has often emphasized (2003; 2012), this is especially the case in our own lineage. Cultural evolution represents a significant departure from the sorts of evolutionary processes responsible for explaining social behavior in most other species, and the theory of cultural evolution I discuss and defend here is an explicit attempt to show that an evolutionary account of human social life, especially, must place a very heavy emphasis on cultural learning.

The questions I pursue here, then, are questions about how religion and culture should be understood from within the scientific perspective. I assume that the answers to these questions will give rise to many additional questions about what wider implications of this scientific picture might be, but those questions extend beyond the scope of this project.

1.3 Defining Religion and Culture

From within the scientific perspective, then, what are religion and culture? Naturally, more detailed answers to these questions will emerge gradually, over the course of the discussions collected here. But it will be helpful, before going further, to provide a general description of the objects of study and shared assumptions that characterize the two disciplines in question.
To this end, it will also be helpful to consider briefly how scientific theories of religion relate to the more broadly naturalistic theories of religion, such as those offered by earlier philosophers like Hume (1992; 1993), Nietzsche (2007) and James (2004). While naturalistic theories and scientific theories share many basic metaphysical and epistemological commitments, these naturalistic philosophers were not primarily interested in developing and enriching our empirical knowledge about religion. Rather, they were concerned with various philosophical questions about religion, and they provided naturalistic sketches of religious phenomena in order to defend their philosophical claims. It is important that in doing so they adopted a materialist ontology, but to do this is not to actually do the work of generating detailed predictive theories for the observable phenomena of religion. Even James did not pursue the study of religion by collecting empirical data, or by running experiments designed in accordance with the scientific method, although he did study other natural phenomena in this way. And conversely, scientists studying the evolution of religion today are not attempting to draw larger philosophical conclusions from their descriptive theories, at least not in the course of their scientific work. They are concerned with predictive explanations, and their aim is to increase the breadth of the religious phenomena explained while continuing to improve the depth and robustness of the explanations provided.

So these, too, are distinct but complimentary fields of theoretical inquiry. I will be focusing on the scientific theories themselves, rather than on their philosophical implications, and it isn't obvious that a more detailed empirical account of religious psychology will have a substantial impact upon the broadly naturalistic assumptions on which earlier philosophical theories have been based. But it also isn't obvious that it won't. These are early days yet for the science of
religion, and it remains to be seen what new philosophical implications may emerge as the naturalistic conception of religion develops into a fully fledged body of scientific theory.

1.3.1 What is Cultural Evolution?

Three different theoretical approaches have been applied to the study of cultural evolution, which I will refer to as the dual-inheritance approach (cf. Boyd and Richerson 2005), the meme approach (cf. Dennett 1995, Ch. 12), and the evoked culture approach (also known as the “epidemiology of representations” approach, cf. Sperber 1996). All three purport to explain the relative frequencies of cultural ideas, beliefs, norms, skills, practices and attitudes within populations, and in this sense all three take cultural evolution to be analogous to genetic evolution. All three also agree that the difference between cultural evolution and genetic evolution depends on how individuals come to acquire their cultural beliefs and behaviors. What they do not agree about is how the acquisition of cultural information by individuals should be characterized.

According to the dual-inheritance approach, individuals inherit cultural traits in a non-genetic way, through social learning. According to the meme approach, by contrast, the transmission of cultural information is characterized as replication, not inheritance. The meme approach concurs with the dual-inheritance approach in assuming that individuals acquire cultural information through social learning, but it does not see social learning as a process in which an already-existing individual (an organism) acquires a new trait. Instead, it treats an instance of cultural transmission as the birth of an altogether new individual, a new living entity with its own goals and interests, which bear no logical connection to the goals and interests of the organism doing the learning. Importantly, however, both of these approaches are based on
appeals to a specifically cultural form of selection process, which is distinct from genetic
selection. And because both take this selection process to be dependent upon transmission by
social learning, both define culture specifically in terms of what individuals learn from each
other.

The evoked culture approach, by contrast, explains cultural evolution without appealing to
cultural selection. On this view, culture in general is not defined in terms of transmission by
social learning, but rather in terms of patterns of distribution for cultural traits observed across
distinct groups. Cultural traits are, roughly, traits that are shared by individuals within the same
group, while differing from one group to the next. According to the evoked culture approach,
cultural beliefs and practices are neither replicated nor inherited, but are merely “evoked” by
particular environments, as the specific manifestations of universal, genetically inherited traits.
Only genetic traits are literally inherited on this view, and individuals are assumed to be
genetically similar across cultures. Nevertheless, the same genes may produce different
phenotypic manifestations in different environments. People from cultures located at high
latitudes tend to wear more clothing than people from cultures located near the equator, but this
is not because individuals learn from the other people in their group how much clothing they
should wear. Rather, different environments simply evoke different specific responses from
individuals who are genetically similar, and who possess the same genetic adaptations for
thermoregulation.

The evoked culture approach and the two selection-based approaches thus use the word
“culture” in slightly different ways. Selectionist theories define culture in terms of social
learning, while evoked culture theories specifically do not define culture in terms of social
learning. As a result, when both types of theory are used to explain “culture,” they are not
actually applied to the same empirical phenomena. Genetic selection is but one causal factor among others (including migration, mutation, genetic drift, etc.), which all work together to explain genetic evolution. Likewise, cultural selection is just one evolutionary factor among others, and the evoked culture approach represents a different, non-selectionist evolutionary factor. As a result, there need be no empirical conflict between the evoked culture approach and the two selection-based approaches. These two types of approach may be used to explain different types of traits—those that spread through social learning, and those that spread by other means. Accordingly, the evoked culture approach is compatible with the other two approaches in the same way that selection-based causes of genetic evolution are compatible with other evolutionary factors, such as migration or drift.

The dual-inheritance approach and the meme approach, however, are not similarly compatible with each other, and the relationship between them is subtle and complex. In the rest of this sub-section, I will provide an overview of each approach, in part by comparing and contrasting them. In Section 1.4.1, I argue that the meme approach should be rejected in favor of the dual-inheritance approach.

Appeals to cultural selection in general are based on modeling techniques originally designed to explain genetic evolution, rather than cultural evolution. Empirical models of cultural selection are based on the methodological innovation of applying the same “old” models of genetic selection to a “new” form of heritable variation—variation produced by learning, rather than by genes. In this sense, then, the only difference between genetic selection and cultural selection concerns the manner in which individuals acquire their cultural phenotypes. As a result, there are many strong similarities between theories of genetic selection and theories of cultural selection. In another sense, however, “merely” changing the mechanism of acquisition
from genes to learning changes everything. Magnitudes of selection pressure, rates of reproduction, sources of variation, and patterns of inheritance all change dramatically, and this is not meant to be an exhaustive list. So despite the general, fundamental similarities between these two distinct types of selection, the evolutionary dynamics and outcomes they produce are radically different in many ways. It is thus important to note that the explanatory power of cultural selection theory in no way depends upon whether, or to what degree, cultural selection is similar to genetic selection. The relevant similarities are methodological, not theoretical, so the question is only whether significant explanatory benefits can be attained by using selectionist models to explain how cultural traits spread.

On this much, the dual-inheritance approach and the meme approach are in agreement. They disagree, however, about how acquisition by social learning itself should be understood. According to the meme approach, this is best understood as a process of replication. Biologist Richard Dawkins (1976) first coined the term ‘meme’ as a way of drawing a close analogy between the units of genetic selection—genes—and those of cultural selection. Genes are replicators, he argues, and it is an inevitable side effect of their tendency to make copies of themselves that some genes end up making more copies of themselves than others do. Those that do succeed in making more copies of themselves, of course, are the ones that possess greater fitness than the others. Noting that socially transmitted ideas, beliefs, behaviors, norms, skills, practices, attitudes, and so on also seem to make copies of themselves as they spread from one individual to another, the meme approach conceives of cultural selection, too, as the inevitable consequence of meme replication. Memes are replicators by definition, and those that succeed in making more copies of themselves than others are the ones that possess greater cultural fitness.
The memetic approach seeks to explain cultural evolution by appealing to the fitness values of these cultural replicators.

By contrast, the dual-inheritance approach does not define social learning as a process of replication (Richerson and Boyd 2005, p. 82; Henrich, Boyd, and Richerson 2008). Instead, social learning is treated as *cultural inheritance*. Social learning takes many forms, and it occurs in many different species (Heyes and Galef 1996). In its simplest form, it is nothing more than “stimulus enhancement,” wherein one animal simply alerts another to the existence of some stimulus in the environment, such as a food source or a predator, typically without intent or awareness. For example, Magurran and Higham (1988) report that minnows increase the frequency of their flight responses after observing the behavior of other minnows who have been threatened by a predator. On the other end of the spectrum, however, the most complex and sophisticated form of social learning is subtle, precise imitation, which in humans is often achieved through deliberate linguistic communication and explicit instruction. But whether or not social learning is intentional, and whether or not language or pedagogy are involved, humans are voracious social learners, and they are able to imitate the activities of others with extremely high fidelity.

As advertising companies are well aware, many people are simply not satisfied with imitating the subtleties of Michael Jordan’s behavior on the basketball court. They also copy his preferences in underwear and footwear, even when those have nothing to do with basketball. According to the dual-inheritance approach, the success of Nike’s Air Jordan shoe line is best explained as the spread of a psychological trait—a preference for wearing a certain type of shoe—which individuals inherit culturally, not genetically. On this approach, then, individuals may inherit any given psychological trait in one of two ways: by genetic relatedness, or by social
learning. And each of these two forms of inheritance is said to support its own distinct form of selection process.

So far, the difference between the meme approach and the dual-inheritance approach may seem rather small. Both define culture in terms of information transmitted by social learning, and both take imitation to be a form of social learning that supports a distinctly cultural, non-genetic form of selection process. So what difference does it make whether imitation is viewed as replication or as cultural inheritance? To imitate someone else’s behavior is to just to copy it, so when one person imitates another it would appear that some cultural belief or behavior has been replicated. This, indeed, is exactly what Dawkins (1976, p. 192) and Dennett (2006b, p. 136) say. They take on board everything that dual-inheritance theorists say about imitation, but they also go on to say, in addition, that imitation counts as a form of replication. This additional ontological commitment, asserting that imitation is a form of replication, is where the two theories fundamentally diverge. And this difference makes a difference.

As Dennett puts it, the meme approach is “a still more radically Darwinian theory of cultural evolution” than the dual-inheritance approach (Dennett 2006b, p. 136). What he means by this is that, when imitation is understood as replication, the packets of information that are copied may be understood as biological entities in their own right, rather than as traits possessed by organisms. If memes are replicators, like genes, then they possess their own fitness values, like genes. And just as selection for some genes over others results in adaptations that further those genes’ “interests,” selection for some memes over others also results in memetic adaptations that further those memes’ “interests.” But the interests of memes are logically independent from the interests of the organisms whose minds they inhabit, and they are also independent from the interests of those organisms’ genes. Accordingly, Dennett describes
memes as symbionts, who live with us—or within us—in the same way that bacteria, viruses and parasites do. In particular, this is the view he takes of religion:

Now it is not our fitness (as reproducing members of the species Homo sapiens) that is presumed to be enhanced by religion, but its fitness (as a reproducing—self-replicating—member of the symbiont genus Cultus religiosis). It may thrive as a mutualist because it benefits its hosts quite directly, or it may thrive as a parasite even though it oppresses its hosts with a virulent affliction that leaves them worse off but too weak to combat its spread. (2006a, p. 84-85)

Since memes have their own adaptive interests, Dennett claims, selection will favor memetic adaptations that increase the relative frequencies of some memes over others. When one person imitates another, the imitator becomes infected, or invaded, by an altogether different living thing, which possesses its own selection history, its own adaptive traits, and its own interests and biological fitness.

Dual-inheritance theorists do not pursue the more radically Darwinian approach of treating memes as biological entities in their own right. One important consequence of this is that, on this approach, imitation is just one form of cultural inheritance among others. Forms of social learning other than high-fidelity imitation may also be considered cultural inheritance, but they can’t be considered replication. Indeed, Henrich, Boyd and Richerson (Henrich, Boyd, and Richerson 2008; Henrich and Boyd 2002) demonstrate that models of cultural selection can play the same role in explaining cultural evolutionary outcomes whether or not social learning takes the specific form of high-fidelity imitation. For example, they note that one way of learning from
others is by taking an average across a range of models, rather than by imitating any particular model, and this form of cultural inheritance results in the same long-term, population-level effects on the spread of ideas. More fundamentally, the difference here is that dual-inheritance theories view social learning in terms of the inheritance of traits, which are possessed by organisms, rather than in terms of the replication of new biological entities, or new individuals that are not organisms at all.

As I argue in detail in the next section, there are both methodological and theoretical reasons to reject the meme approach to cultural selection in favor of the dual-inheritance approach. And as I explain much later on, in Chapter 4, the theory of cultural evolution in general cannot live on evoked culture alone, but the evoked culture approach nevertheless plays an essential role in explaining certain cultural evolutionary facts. I thus conclude that cultural evolution is best explained by a combination of the evoked culture approach and the dual-inheritance approach, although that is not to rule out the possibility that other approaches and other evolutionary factors might yet be developed in the future.

1.3.2 What is Religious Evolution?

Considered as a natural phenomenon, religion is highly complex. And considered as a concept, religion is highly abstract. For any given religious person, a wide variety of psychological traits will play some role or other in their religious activities. And across individuals, a bewildering variety of distinct forms of behavior and belief can easily and intuitively be recognized as religious. Yet on the other hand, intuitions also vary considerably regarding which forms of belief and behavior count as the religious ones. Does Scientology count? Does Zen Buddhism? What is the difference between specifically religious forms of supernatural belief and
supernatural belief in general? Is supernatural belief necessary, or essential, for religion? As considerations such as these illustrate, if one wants to study religion from a scientific perspective, it isn’t clear what, exactly, one’s target explananda ought to be.

There is no standard, widely accepted definition of religion in naturalistic terms, and there is no general theory of religion to which scientists from different disciplines may appeal when designing their studies and research programs. So on the assumption that to explain the evolution of religion is to explain a certain set of behavioral and psychological traits, the important question is: which traits are the religious ones? And science currently has no clear answer to that question. I see no reason to doubt that a widely accepted, general theory of religion and religious evolution can eventually be provided. Indeed, I hope to have developed here some of the building blocks for such a theory, and I am optimistic that the cultural evolutionary account of religion most fully articulated by psychologist Ara Norenzayan (2013) is at least an early prototype for such a theory. But these building blocks and this particular theory are new, and their acceptance is by no means universal. So the best that even an optimist can say is that a general, widely accepted theory could, possibly, be on the horizon.

Thus far then, when identifying which traits count as the religious ones, scientists have had nothing to appeal to but folk intuition. As a result, what we find in the literature is a variety of different theories, which explain a variety of different traits. Each of these traits, considered on its own, may plausibly be considered religious. Yet little theoretical work has been devoted to identifying the theoretical relationships between these accounts, or the functional relationships between the traits explained by them—recent synthetic accounts offered by dual-inheritance theorists are an exception (Norenzayan 2013; Atran and Henrich 2010). One of my primary aims has thus been to connect these dots, and to consider in detail how independent theories and traits
are related to one another. In doing so, I have identified some basic assumptions about religion that are widely shared among scientists working in this area, and that help to make the study of religious evolution a coherent scientific project.

Most generally, the target explanandum of the discipline is the religious phenotype as a whole, a more or less coherent system of functional traits that distinguishes (according to common-sense intuition) religious individuals from non-religious individuals. Given the complexity of religious behavior, and the abstract, messy nature of the folk concept religion, no single religious trait can settle the question of whether a given individual is religious. Belief in supernatural agents has been one of the most popular target explananda among theories in the field (cf. Boyer 2001, Atran 2002, Barrett 2004) and this may indeed be one of the strongest candidates for a necessary or essential feature of the religious phenotype. But even if belief in supernatural agents does constitute a necessary condition for being religious, it does not constitute a sufficient condition as well. A Christian individual might also believe in ghosts, for example, and her ghost beliefs may play no role at all in her religious cognition and behavior. Thus, her supernatural-agent beliefs about ghosts would not be sufficient to make her religious, even if she is a religious person. Further, it isn’t clear whether belief in supernatural agents is even a necessary or essential feature of the religious phenotype. Some Buddhists follow the Buddha himself in believing that supernatural agents play no role in Buddhist practice, but it isn’t intuitively obvious that such individuals are not religious. Belief in supernatural agents may count as a religious trait, but whether it does or not depends upon the functional role these beliefs play in relation to other traits of the religious phenotype.

As a result, inquiry breaks down into two distinct but related questions: (1) how did the various component traits of the religious phenotype evolve, and (2) how did the functional
relationships between these components evolve? The first question focuses on “parts-level” explanations, addressing the parts of the religious phenotype independently, while the second focuses on “system-level” explanations about how these parts come together to form a whole, and about what each part’s relationship is to the other parts. In Chapter 3 I defend a model of the religious phenotype designed to keep the answers to these two types of question clearly distinguished from each other, while also explaining how they are related. According to this model, many important features of the religious phenotype are traits that originally evolved prior to the evolution of the phenotype itself, and that originally evolved to perform other, non-religious functions. More recently, however, these otherwise disparate traits have come together to form a system, and in the process they have evolved additional, secondary functional roles. That is, they have acquired new functional relationships with each other on top of, or in addition to, the original, non-religious functions they already possessed. Like capacities for representing supernatural agents, which also enable people to form beliefs about ghosts, many constituent traits of the religious phenotype also perform other non-religious psychological roles. Distinguishing between system-level facts and parts-level facts provides a way of determining when each part is, and is not, functioning in a specifically religious way.

For any given trait, then, one of the first questions that must be addressed is how individuals acquire it—through social learning, through genes, or through some other mechanism? From a psychological point of view, I assume, this is already an intrinsically important question. For example, it is important to have an explanation for the fact that children born to Christian parents will only become Christians themselves if they undergo the right learning experiences. But this question is also important for a different, more fundamental reason, which concerns the basic motivations mentioned above for studying religion from an
To study religion as a natural phenomenon is to study a certain form of human behavior and cognition, a certain set of psychological facts. And, like other psychological facts, facts about religious psychology are primarily, if not exclusively, explained by identifying the functions of various psychological traits and capacities. Evolutionary theory has something important to contribute to this enterprise: an empirically grounded method for identifying and individuating functions. While traditional, non-evolutionary theories of religion have always sought to identify the function(s) of religious belief and behavior, they have done so without access to an independent set of empirical facts against which they could test their functional hypotheses. Selection-based theories of religious evolution can change this.

As the philosopher Robert Cummins (2000; 2002) and the linguist Noam Chomsky (Chomsky 1959; 1976) have independently argued, psychological explanations are not like explanations in physics. Psychology proceeds not by subsuming observations under law-like regularities, but by attributing functional capacities, or abilities, which are realized by internal systems and processes. According to Cummins, the basic explanatory strategy of psychology is functional analysis, according to which the explanation of “bigger,” more complex capacities is achieved by decomposing them into “smaller,” less complex capacities. If vision, for example, is a “big” capacity, then explaining it involves analyzing it into “smaller” sub-capacities, such as color constancy, depth perception, edge detection, motion detection, and so on. But how are all these functions and sub-functions to be identified and individuated?

One method is to appeal to one’s teleological intuitions. Imagine that while crossing a heath, you pitched your foot upon, not a watch, but some obscure tool whose function was unclear. It would be natural to wonder what the tool is for, and to generate some hypotheses in an effort to reverse-engineer it. But in the absence of any further information about where the tool
came from, intuitive hypotheses about its function are all one has to work with. There is no way to test these intuitions against independent empirical facts.

Now, however, imagine encountering this same tool on a shelf in the hardware store. It would still be natural to wonder what it is for, and to generate some intuitive hypotheses in an effort to reverse-engineer it. But now these hypotheses can also be tested after they have been generated. One can simply ask a salesperson what the tool was designed for. The tool’s etiology—its history of design—can be used to confirm or disconfirm intuitive hypotheses about what it is for. Of course, in psychology what we are attempting to reverse-engineer are not artifacts designed by people, but naturally occurring functions, which have not been designed according to the goals of any agent at all. But this just shows that the relevant functional attributions may be made in the same way that other natural functions are attributed in the biological world. In biology in general, what a trait is for is whatever it has been selected for.

This etiological method of identifying and individuating functions has been defended explicitly by evolution-minded philosophers of mind, such as Ruth Millikan (1984), Daniel Dennett (1987; 1991) and William Lycan (1981). But while etiological functionalism may be controversial in philosophical debates about how to naturalize intentionality, or about the metaphysics of phenomenal consciousness, there is nothing at all controversial about it in biology in general, or in the field of religious evolution in particular. All theories in this field are adaptationist theories, which explain the evolution of the religious phenotype by appeal to some form of selection process. One of the most important implications of this is that it provides an objective, empirical method for attributing functions, a clear set of facts against which functional claims and hypotheses may be tested.
It must be noted that this does not address the practical problem of actually accessing the relevant historical facts. Nervous system activity leaves no traces in the fossil record, and the fact that certain capacities are currently adaptive provides no guarantee that they have actually evolved to perform those functions. So care must be taken not to draw conclusions that extend beyond the historical evidence base. But this practical difficulty is not an insurmountable problem. Philosopher Edouard Machery (2007) describes no less than four potential sources of evidence for these etiological facts, noting that combining them in various ways can be quite powerful. So while etiological functionalism may incur the difficult practical burden of facing up to etiological facts, it also provides a powerful method for determining, in a principled manner, where one functional capacity ends and the next one begins. This is particularly important in the study of religion, which has been plagued by a problem I will call “functionalism run amok.”

Consider some intuitive, common-sense explanations for religious belief and behavior. When I first described my dissertation topic to my uncle, he volunteered a functionalist theory of his own: religion provides existential comfort, since most people need a reason to believe that life has some meaning or purpose. A different view of the function of religion that one commonly encounters holds that it provides explanations for otherwise puzzling experiences and phenomena, such as the origins of the universe, the meaning of dreams, or apparent violations of natural order. Zeus’s thunderbolts, for example, explained lightning at a time when no better account was available. Yet another view of the function of religion holds that it maintains social order by giving people a reason to think that genuine moral truths exist, and that someone very powerful is willing to enforce them. These explanations are all quite different from one another, but they are also all intuitively plausible. Some evidence can be mustered to support each of them, but this leaves it entirely open how these accounts relate to one another. Are they
conflicting accounts, or do they each tell just part of the story? If they are conflicting, how can we adjudicate between them? And if they aren’t, how can we explain why these three distinct functional roles come packaged together in the way we observe?

Anthropologist Pascal Boyer identifies this problem while criticizing what he calls “classical functionalism” in social science, an approach that that seeks to explain social interactions by identifying some social function they performed, some benefit they provide for the society. This approach fell out of favor with anthropologists decades ago, Boyer claims, largely because there were no empirical facts to check intuition-based theories against.

“Functionalism,” he says, “was accused of peddling ad hoc stories. Anyone with enough ingenuity could find some sort of social function for any cultural institution” (p. 25). One problem with this is analogous to the famous critique of adaptationist theorizing presented by Gould and Lewontin (1979), which Machery (2007, p. 4) neatly summarizes thus: “unconstrained speculations about adaptive traits are so easy to come by as to be of little epistemic value.” And just as unconstrained speculation about the adaptive functions of genetic traits amounts to little more than the telling of “just-so” stories, so too does unconstrained speculation about the social and psychological functions of complex phenomena like religion.

But the intuitionist method of functional attribution also creates another problem. When theories of this kind come into conflict with each other, debate can devolve into mere table pounding, with no way to adjudicate the stalemates that arise from conflicting intuitions. In philosophy this may be no cause for embarrassment, but that is because many lines of philosophical inquiry aren’t pursuing empirical questions in the first place, and intuition is all anyone can appeal to. That is not the case here. Boyer’s point about ad hoc stories is that, if functionalism is to be
useful for the purposes of predicting behavior, then it must be constrained by empirical facts. Otherwise, it is functionalism run amok.

Adaptationist approaches to the study of religion solve the problem of functionalism run amok by identifying the empirical facts about selection history that ground functional attributions. This is especially useful for addressing questions about the highly complex, highly abstract phenomena of human social life, such as religion and culture.

As I mentioned above, however, the particular adaptationist approach I defend is based on the dual-inheritance approach to cultural evolution, and on this approach there are two distinct types of selection process that could, in principle, be responsible for determining the function of any given trait. To say that a trait is an adaptation, or a product of selection, is to introduce a new kind of evolutionary question: what kind of adaptation is it? Which selection process gave rise to it? This, ultimately, is why the question of acquisition is so important for any given religious trait. If it is inherited genetically, then the etiological facts relevant to identifying its function(s) will be facts about genetic selection. If it is inherited culturally instead, then an entirely different selection process will be responsible for determining its function, and just how this alternative selection process should be understood is the topic of the next section. For while the meme approach holds that selection among memes is what determines the relevant etiological functions, the dual-inheritance approach holds that it is selection among traits possessed by organisms that accomplishes this task.
1.4 The Metaphysics of Cultural Selection: Replication vs. Inheritance

Laland and Brown (2011, Ch. 6) provide a brief but informative historical overview of the field of cultural evolution, and in it they dismiss the meme approach as having already failed, at least for scientific purposes. They first note that, “As scientific concepts go, the ‘meme’ had the best possible start—it was launched in one of the most popular scientific books of the twentieth century and given further attention through the highly successful writings of philosopher Daniel Dennett” (p. 141). They also point out that by the late 1990s several other authors had also published popular books on memetics, and conferences were being organized on the topic. In 1997, a new journal was founded: the *Journal of Memetics: Evolutionary Models of Information Transmission*. But despite these auspicious beginnings, they argue, the meme approach never managed to produce serious scientific research: “Although computer geeks ran away with the idea, generating a popular subculture of meme followers, in academic circles the meme fell on fallow ground… The *Journal of Memetics* closed down in 2005, the popular writings started to dry up, and interest in memes, at least in the scientific forum, began to wane” (p. 141). I know of no evidence to refute this historical analysis. Dennett’s meme-based theory of religion (2006a), for example, has had no discernible impact upon scientific theories of religious evolution. And meanwhile, as Laland and Brown emphasize, the dual-inheritance approach is still gathering steam, as scientific studies based on that approach continue to make new and important findings. Thus, it appears that while meme theory offers an intriguing perspective on cultural evolution, it is not a perspective that scientists have actually adopted and employed. It is the dual-inheritance approach, not the meme approach, which has done the heavy lifting involved in demonstrating the scientific value of cultural selection theory.
From a methodological point of view, I agree with Laland and Brown that this is already sufficient for rejecting the meme approach, and for adopting the dual-inheritance approach instead. But from a philosophical point of view, this introduces a new problem: why haven’t scientists adopted the meme approach? That it has failed to make a scientific contribution seems to suggest that there is something wrong with it, but it doesn’t show where, specifically, the problem lies. In the next sub-section, I attempt to diagnose the meme approach from the philosophical perspective, arguing that there is indeed a theoretical flaw in it. The problem is not that the meme approach is false, or that theories based on this approach make false predictions about how forms of cognition and behavior spread through imitation. Rather, the problem is that the meme approach is metaphysically extravagant; it requires additional ontological commitments, over and above those made by dual-inheritance theories, which add no explanatory or predictive value. My diagnosis is thus intended more in the spirit of rational reconstruction than causal explanation. The claim is that, to the extent that scientists may have avoided the meme approach because of suspicions about the metaphysical commitments involved, they have been right to do so. Puzzling through this issue will prove useful, not just because of what it shows us about the meme approach, but also because doing so reveals some important commitments of the dual-inheritance approach.

In the sub-section that follows, I address a different problem with the meme approach, which also provides a useful point of comparison with the dual-inheritance approach. Because meme theory is founded on the concept of replication, it views cultural transmission strictly in terms of high-fidelity imitation. But even if this explains many important cultural phenomena, it nevertheless leaves out other important cultural phenomena that the dual-inheritance approach
does not. Forms of social learning other than high-fidelity imitation also make important contributions to the accumulation of cultural information.

1.4.1 Making Memes Come True

Recall the basic difference between the meme approach and the dual-inheritance approach, which both purport to explain the events of a specifically cultural selection process. In Dennett’s terms, meme theory is “more radically Darwinian,” because it posits an altogether new type of replicator, a new kind of biological entity, agent or individual. A meme is not an organism, but is nevertheless its own kind of living being, possessing its own interests, its own selection history, and its own adaptations for making copies of itself. Citing Dawkins, Dennett emphasizes that “in our explanations we tend to overlook the fundamental fact that ‘a cultural trait may have evolved in the way it has simply because it is advantageous to itself’” (1995, p. 362). It is these advantages, then, which are registered from the meme’s-eye view, that are said to explain the cultural behavior and cognition of human individuals.

In an instance of imitation, a meme replicates itself. A new token meme is created, whether it appears in the mind of another person, or on a new piece of paper, or on a different computer. By definition, the new meme token must be functionally isomorphic to the “parent” token that produced it, while the material substrate responsible for realizing the relevant functional roles is irrelevant. Unlike genes and viruses, which are physical replicators, memes are functional replicators. That is, a new meme token must be functionally similar to its parent token, but that does not require that it also be physical similar. It does require, however, that the new meme token is numerically distinct from the parent token. When a new token meme is born, a new individual has been born into the world, bearing its own capacities for, and interests in, self-
replication. In this sense, replication among memes is analogous to reproduction among organisms.

By contrast, consider what an instance of imitation looks like from the dual-inheritance perspective (keeping in mind that imitation is not the only form of cultural inheritance). When imitation is conceived as inheritance, rather than as replication, the new token born into the world is a trait, not an altogether new individual. It is a new trait of an individual who already exists, and who already possesses her own capacities for imitating others. The concept of replication implies that one individual gives rise to another whole individual, which is similar in all (functional) respects to its “parent.” The concept of inheritance, by contrast, implies only that an individual has acquired something in particular from another individual. When it is the inheritance of wealth or property that is in question, it may be that only one individual at a time can be in possession of what is inherited, so when one person acquires it, another person loses it. This is not the case for cultural inheritance. When someone learns something from a cultural model, the model’s traits do not change, but the learner’s traits do. In other words, what happens in an instance of cultural inheritance is that one individual becomes similar to another individual in some particular respect—not in all respects. No new individuals are brought into existence. Rather, an already-existing individual simply changes, acquiring a new property.

Thus, while the meme approach takes a lineage of token individuals to be what persists over the course of a particular cultural selection history, the dual-inheritance approach envisions, instead, a heritable trait persisting throughout a given selection history. That is, where dual-inheritance theories describe how cultural traits evolve, as they are transmitted from one individual to another, meme theories describe how a lineage evolves, as one individual gives rise
to another. These individuals are not organisms, but they are nevertheless living beings in their own right, which are not to be confused with the organisms they inhabit, invade, or infect.

Importantly, this does not imply that the dual-inheritance approach views the functions of cultural traits in terms of the fitness advantages they confer on the organisms who possess them. If cultural traits were inherited genetically, then for them to possess any function at all would be for them to have made some kind of positive contribution to the genetic fitness of human organisms. But cultural traits are not inherited genetically, and this has important consequences for the role cultural selection plays in determining etiological functions. In the next sub-section I examine this issue at length. For now, however, the point is that there is a big difference between attributing functions to cultural traits, on one hand, and attributing interests to them, on the other. Because traits are not replicating individuals, they do not have their own reproductive interests, distinct from those of the individuals who possess them. And because cultural traits are not inherited genetically, there is no necessary connection between their functional properties and the genetic interests of the individuals who possess them. Cultural traits may have all kinds of effects on the genetic fitness of individuals: positive, negative and neutral. Indeed, in their *Not By Genes Alone* (2005), which lays out a detailed account of the fundamental concepts and commitments of the dual-inheritance approach, Richerson and Boyd devote one whole chapter to illustrating how cultural traits can increase genetic fitness, and another whole chapter to illustrating how they can reduce genetic fitness, instead.

Of course, if cultural traits had maladaptive effects more often than adaptive effects, then they would never have evolved in the first place. That is, our capacities for imitation (and other forms of social learning) would never have evolved. Everyone studying cultural evolution recognizes this, so everyone agrees that imitation capacities are adaptive overall. But this is fully
compatible with the fact that, once these capacities evolved, they introduced the possibility for some forms of cultural information to have the opposite effect. Richerson and Boyd argue that the adaptive value of cultural inheritance lies primarily in the adaptive flexibility provided by the accumulation, over many generations, of cultural knowledge (Boyd and Richerson 1995). Equipped with the accumulated knowledge of their cultural ancestors, they argue, humans possess powerful skills, beliefs, norms and practices that enable them to behave adaptively in even the most difficult environments. Accordingly, their view is that, “Acquiring information from others allows people to adapt to a wide range of environments, but it also opens a portal into people’s brains through which maladaptive ideas can enter—ideas whose content makes them more likely to spread, but do not increase the genetic fitness of their bearers” (2005, p. 150). They provide the handy example of Pentecostalist snake handlers, whose culturally inherited traits lead them to believe that God will protect them, while also leading them, often enough, to die by snakebite.

From the meme’s-eye view, such costly behavior is the exploitation of human individuals by insidious memes. If the beliefs of Pentecostalist snake handlers are viewed as replicating individuals in their own right, then it is easy enough to see why Dennett places such a heavy emphasis on the analogy between memes, on one hand, and viruses and parasites, on the other. For while he fully recognizes that memes have positive, adaptive effects more often than they have negative, maladaptive effects, Dennett also claims that the significance of the meme approach lies in the perspective it affords on harmful, maladaptive memes, in particular. In a section of *Darwin’s Dangerous Idea* (1995) entitled “The Philosophical Importance of Memes,” Dennett relies on a comparison with genetic selection, arguing that the significance of adopting the gene’s-eye view—the view of selection that is the topic of Dawkins’ *The Selfish Gene*
lies in its ability to show how the interests of genes may diverge from the interests of the organisms they inhabit: “Before Williams and Dawkins pointed to the alternative gene's-eye perspective, evolutionary theorists tended to think that it was just obvious that adaptations existed because they were good for the organisms. Now we know better. The gene-centered perspective is valuable precisely because it handles the "exceptional" cases in which the good of the organism counts for nothing…” (1995, p. 364). For the same reason, he argues, the significance of the meme approach also depends on its ability to handle cases in which the perspective of the meme’s interests diverges from the interests of its “host”: “The prospects for meme theory become interesting only when we look at the exceptions, the circumstances under which there is a pulling apart of the two perspectives. Only if meme theory permits us better to understand the deviations from the normal scheme will it have any warrant for being accepted” (1995, p. 364).

Dennett must be applauded for having the courage and honesty to state clearly and boldly what the conditions of failure would be for the theory he himself defends. But more importantly, it is striking that the conditions for failure Dennett sketches here are the same kind with which Laland and Brown are also concerned: a failure of significance, rather than of truth, or empirical accuracy. Laland and Brown do not attack the meme approach by pointing to cases in which the meme theory and the dual-inheritance theory provide competing predictions, in an effort to show that the dual-inheritance view is true, while the meme theory is false. And Dennett does not pursue this tactic either, in order to argue in the other direction. This is because, as Dennett recognizes explicitly (1995, p. 358), facts about memes must ultimately be reduced to ordinary psychological facts—facts about the cognitive and motivational traits of human individuals. The meme theory and the dual-inheritance theory are empirically equivalent, so Dennett’s claim is
not that we may eliminate the underlying facts of human psychology on which facts about memes supervene. Rather, he suggests that it will ultimately prove useful, in certain exceptional cases, to abstract away from the nitty-gritty details of psychology, and to focus instead on higher-order facts about memes.

As we might put it, Dennett’s defends the philosophical importance of memes, rather than their scientific importance, because he recognizes that the meme approach is empirically equivalent to the dual-inheritance approach. The mathematical models that have enabled theories of cultural selection to make new and important empirical findings are based on the dual-inheritance approach, not on the meme approach. They were borrowed from population geneticists who were also attempting to explain the spread of traits, rather than of some altogether new kind of biological individual. Dennett is fully aware of this, so his point seems to be that even though the two approaches are empirically equivalent, it is nevertheless useful to adopt the meme’s-eye view in certain cases. The question then becomes: why would this be useful? For the dual-inheritance theory explains all of the same exceptional cases, the cases in which the interests of the organism count for nothing, but it does so without adopting any of the additional ontological commitments that make the meme approach, in Dennett’s own terms, “a still more radically Darwinian theory.” People learn things from each other, and while most of the things they learn are beneficial to them, some of the things they learn are bad for them, instead. Why complicate this picture?

Consider again the analogy between memes and viruses. Even though they must commandeer the machinery of host organisms in order to replicate, viruses, too, are replicators in their own right, possessing interests and adaptations of their own. The functions of the virus’s adaptive traits, like those of any other biological individual, are defined or identified in
etiological terms, by appeal to the history of selection that has shaped the virus’s genome. But that selection history is entirely different from the selection history that has shaped the genome of the virus’s host, and this is supposed to be the whole point of comparing viruses with memes. When a virus infects a host, and causes the host to behave in some particular way, it would be a mistake to seek an explanation for this behavior among the host’s own traits and capacities. Those are the wrong traits to appeal to, because they belong to the wrong replicating individual. And for the same reason, Dennett says, when a meme infects a person, and causes her to behave in some particular way, it would be a mistake to seek an explanation for this behavior among her own psychological traits. Again, those are the wrong traits to appeal to, because they belong to the wrong replicating individual. The individual whose traits are relevant is the meme, and the functional features of these traits are determined by the meme’s selection history, not the selection history of the host’s genome.

But this brings out an important disanalogy between memes and viruses. When a virus infects a person, we really can appeal to the traits of a distinct and independent replicating individual. We can explain the host’s behavior by appealing to the traits of the virus instead of the traits of the host. This is not possible with memes, because the meme’s traits are the traits of its host. Qua replicator, a meme is ontologically dependent upon the functional, psychological activities of its host. Because memes are functional replicators, rather than physical replicators, they cannot even be identified and individuated without reference to the metaphysically prior functional activities of their hosts’ nervous systems. Once the meme for wearing Air Jordans, say, has been identified, it is possible to reframe questions about its spread as questions about how it replicates. But this will not change the fact that what is being said to replicate is already a psychological, functional trait of an organism. To say that the meme for wearing Air Jordans has
replicated is just to say that some person has acquired, from some other person, a motivation to buy a certain pair of shoes. As Dennett and Dawkins fully recognize, any instance of meme infection must ultimately be explained as an instance of imitation. But the capacity for imitation is a functional trait of the host, not of the meme, and the function of this capacity is determined by a history of selection acting on the host’s genes, not on the relevant meme. In order to even identify what the meme’s selection history has been, we must first presuppose the existence of functional traits belonging to the meme’s hosts.

We may, if we choose, declare that psychological facts about imitation also count as memetic facts about replication, and proceed to translate claims about social learning into claims about meme replication. But this will not make the underlying psychological facts disappear, and it will not allow us to explain or predict anything that we couldn’t already explain or predict. Accordingly, I see no reason to make the additional ontological commitments that the meme approach requires, involving an entirely new form of non-organismic biological individual, complete with its own interests, selection history and functional traits. It is possible to appeal to the meme as a distinct sort of replicating individual, but there is no way to replace explanations based on the traits of that individual with explanations based on the traits of ordinary organisms, as we do in the case of viruses. Facts about memes are just facts about the cognition and behavior of organisms, and, in the end, they are most perspicuously and parsimoniously explained by appeal to the psychological traits of those organisms themselves. In other words, while the meme theory is not false, it simply does not, as Dennett asserts, “permit us better to understand the deviations from the normal scheme.” Adopting the additional ontological commitments of meme theory affords an intriguing new way of looking at such cases, which
explains the popular success of the meme meme. But it doesn’t afford any better understanding, at least from the point of view of predictive explanation.

1.4.2 Social Learning Without Imitation: Explaining the Evolution of Environments

The argument just provided is, in my view, the primary reason for rejecting the meme approach in favor of the dual-inheritance approach. Even when cultural learning does take the form of high-fidelity imitation, and thus even when there is a relatively clear sense in which something might be said to have replicated, it is better to view the “units” of cultural selection as traits belonging to individuals, rather than as individuals in their own right. But there is also another reason: not all cultural inheritance does take the form of high-fidelity imitation. This point has been developed in detail by philosopher Kim Sterelny (2012), who argues that much of cultural evolution in humans is driven by what he calls “hybrid” learning, involving a combination of imitation and socially structured individual learning.

Consider the difference between learning a new word in English and learning the English language in general. A person who already knows how to speak English is in a very good position to learn a new word through imitation, because her language capacities make it possible to define even abstract terms explicitly, and to communicate their meaning clearly. The learner can thereby culturally inherit a new functional capacity, a new psychological trait: the ability to use this word. But that very specific capacity is different from the general English-speaking capacities that made it possible to communicate the definition of the word. Both types of trait are, of course, culturally inherited; language learning of all forms is clearly (perhaps even necessarily) social learning, and this is why linguistic traits in general follow clear patterns of cross-cultural variation. But while the learning of a new word through explicit definition is a
textbook example of high-fidelity imitation, learning how to speak a language in the first place is not. Certainly, a lot of imitation is also involved in learning any language, but language acquisition is not just a matter of observing the behavior of others and then copying it. Richerson and Boyd (2005, p. 82-83) point out, for example, that grammatical rules may be inferred from across multiple utterances, none of which is actually imitated. Moreover, for more abstract concepts, individual learning may be necessary for mastering the subtleties of application, so there may be much more to learn about how to use a new term than one can convey through direct, explicit definition. Learning the meaning of “apple” may take only a second, and may take place through straightforward imitation. But mastering the meaning of words like “ontological” or “Hegelian” may take years, and may require a good bit of individual learning about ontology or Hegel.

Sterelny argues convincingly that much of our social learning is actually a complicated mixture of individual, asocial learning and direct social learning. However, he emphasizes that much of this individual learning is also socially influenced in indirect ways. “Most social learning,” he claims, “is hybrid learning: agents acquire skills through socially guided trial and error and socially guided practice. Children do get advice, instruction, and other informational head starts from others, but they get this support while engaged in exploratory learning in their environment” (2012, p. 28). The key point is that, “Learning mediated by the environmental effects of one’s social group is not solitary learning, even if the internal cognitive mechanisms deployed are the same as those used in exploration learning” (p. 31). By sending children to school, or by keeping the girls in the kitchen while sending the boys out to manage the herds, we mediate the learning environments of individuals in ways that affect what they learn on their own, in addition to what they learn through imitation.
Accordingly, Sterelny suggests that a good model for understanding many important forms of cultural inheritance is provided by the educational process of a traditional craft apprentice. Learning how to track and hunt, knit a sweater, knap flint into an arrowhead, or speak a language is more like learning how to craft a fine piece of furniture or a sword than it is like learning a new word, or a new scientific fact. Sterelny describes what he calls “the apprentice learning model” as follows:

Apprentice learning is hybrid learning: combining information from the social world with information from the physical-biological environment. It is learning by doing. But it is learning by doing in an environment seeded with informational resources. These include raw materials, in both their raw and processed forms. In addition, full and partial templates of the final product are available to guide action. So too are tools. Moreover, there are many opportunities to learn by observing highly skilled practitioners. Advice is often available from both experts and peers, for learning is often social and collaborative. Indeed, the learning trajectory of an apprentice is often at least partially organized by experts. The expert organize the trial-and-error learning of the less expert by a combination of (i) task decomposition and (ii) ordering skill acquisition, so that each step prepares the next. (p. 35)

Consider how this applies in the academic context, where the apprentice is a graduate student. While imitation may explain how a student learns particular historical or scientific facts, for example, learning particular facts is very different from learning how to do research. Graduate
training is learning by doing, and while many of the skills and methods involved in conducting research are learned through imitation, imitation is clearly not all there is to the acquisition of these skills and methods. Professors structure the trial-and-error learning of students, and advice from fellow students may also alter the subsequent course of one’s individual learning—about ontology or Hegel, for example. Many of the skills and concepts that the apprentice researcher ends up acquiring are the products of a complex, dynamic interplay between individual learning and social learning. This shows that much of what is social about social learning in humans involves factors other than direct imitation alone.

The fact that the scope of social learning extends beyond imitation has important consequences for the role that cultural selection and cultural inheritance play in explaining people’s behavior. Culturally inherited traits are psychological traits, which explain cognition and behavior in functional terms, by ascribing functional capacities. And as I’ve mentioned, it is widely accepted among scientists studying cultural evolution that psychological functions should be ascribed in etiological terms: what a functional capacity is for is whatever it has been selected for. Accordingly, the question arises how histories of cultural selection contribute to the process of determining the functions of cultural traits. And the answer is different, I argue, depending on what kind of learning is involved in cultural inheritance.

For traits acquired through straightforward imitation, the relevant selection history stretches back through the cultural generations. To the extent that someone simply adopts the same beliefs that her cultural ancestors possessed, the selective events that determine those beliefs’ psychological, functional properties will have occurred prior to the agent’s own adoption of the belief. When a child today acquires the belief that Jesus is the Son of God, she is culturally inheriting a belief whose content evolved centuries before she was born. Accordingly, the
etiological facts responsible for determining this belief’s functional properties are facts about the traits of other individuals: beliefs belonging to the child’s cultural ancestors. In this sense, cultural inheritance explains psychology in the same way that genetic inheritance does. The functional properties of the trait in question are determined prior to the event of inheritance itself, and the agent simply takes them on board in a coherent package when she inherits the trait. Inheritance amounts to simply “plugging in” a pre-formed functional component, like adding an additional speaker to a stereo system.

By contrast, when cultural inheritance takes the form of hybrid learning, the psychological traits explained cannot all be explained by appeal to prior histories of cultural selection. The relevant etiology involves a critical admixture of individual learning, in which the acquisition of functional skills and beliefs is due to facts about the agent’s own traits and activities, rather than to facts about the traits and activities of her cultural ancestors. In these cases, cultural selection makes a different kind of contribution to explanations of the agent’s psychology. In addition to explaining the functional properties of cultural traits themselves, cultural selection also explains the evolution of the environments that structure ordinary individual learning.

Consider the learning of social norms. One way a person might learn to follow a norm is by copying what other people do when they are following the norm. When visiting a new town, you notice that everyone else politely says “Thank you!” to the bus driver as they get off the bus, leading you to feel that if you don’t copy them, you’ll be violating a norm of gratitude. A different way of learning a norm, however, is to violate it first, then suffer the consequences of punishment, and then learn to follow the norm in order to avoid further punishment. The first scenario relies directly on social learning capacities—functional capacities that are specifically for imitating others’ behavior. The second, by contrast, describes ordinary individual learning,
proceeding by trial and error. Since only one of these learning processes appeals to capacities for
imitating the cognition and behavior of others, they are importantly different from a
psychological point of view. But in either case, a history of cultural selection for the relevant
normative beliefs and behaviors will figure essentially in an explanation for the norm-learner’s
behavior.

When a norm is adopted through imitation, what selection explains is the function of the
new trait the learner has inherited, in virtue of etiological facts about the agent’s cultural
ancestors. This is not the case when the norm is learned individually, by trial and error.
Nevertheless, one cannot explain the relevant trial-and-error learning without appealing to the
antecedent presence of this norm within the individual’s social environment. Were she not
present within a society in which that particular norm had previously been selected, there would
have been no punishment for her to learn from. Sterelny’s appeal to the apprentice learning
model illustrates the fact that, despite the psychological difference between social learning
capacities and individual learning capacities, individual learning itself can still be quite social, in
another sense: it is learning that shaped by the behavior of other people.

In addition to determining the functional properties of traits inherited by social learning,
then, cultural selection also plays an important role in structuring the environments in which both
individual and social learning take place. A child born into a culture of horse nomads, such as
those that have traditionally occupied the Eurasian steppe, will grow up to learn a lot about
horses. Much of this learning will take the form of direct cultural inheritance, as the child copies
the other people around her when learning to ride or care for horses. But much of what the child
learns about horses will be due to individual, trial-and-error learning, in an environment that
happens to be filled with horses. Cultural selection explains the functions of the cultural traits
that the child inherits by copying others, but it also explains other things that the child learns
because of others, if not from them. For example, selection for the beliefs and motivations that
led the child’s cultural ancestors to keep horses in the first place explains why her own individual
learning takes place in an environment filled with horses. Thus, cultural selection not only
explains the functional features of culturally evolved traits, but it also explains important
features of culturally evolved environments. And the evolution of both is important for
explaining many features of human psychology.

Cultural selection impacts both the social environment and the physical (non-social)
environment. Its effects on the social environment include, most notably, the evolution of certain
culturally inherited traits themselves, including those comprising the norms, practices, beliefs
and institutions of religion. But these social phenomena also have downstream effects upon the
physical environment, including the cities we build and the artifacts we design and manufacture.
Humans dramatically modify their physical environments in all kinds of ways, and they do so
largely on the basis of culturally inherited knowledge. So in the same manner in which a history
of cultural selection explains the current features of the social environment, it also explains many
current features of the physical environment. This point is worth developing in a bit more detail.

Without learning anything about building from anyone else, a lone individual might, by the
end of his life, be able to build a pretty nice house out of stone or wood. But building a
skyscraper would be out of the question, even if the person had access to an army of workers,
and all the tools and raw materials out of which modern skyscrapers are constructed. Actual
skyscrapers are built using techniques and concepts that have gradually increased in functional
complexity and sophistication over many cultural generations. Indeed, one important sense in
which cultural selection and genetic selection are similar is that both explain the gradual
increases in functional complexity that particular traits undergo, through a process of descent with modification.

The more widespread a cultural trait is, and the longer it is retained within a population, the more opportunity there is for someone to improve upon it. This is the cultural evolutionary analog of a lucky, beneficial mutation. If the innovation really is an improvement, then it too is likely to spread, and to be retained over generations, and eventually improved upon. As this cycle of improvement, spread and retention continues, cultural knowledge becomes more useful, efficient and functionally elaborate. We move from huts made of thatch and wood to buildings made from stone and brick, and eventually to skyscrapers made of concrete and steel. The same pattern applies for scientific knowledge, as our culturally inherited beliefs evolve from Ptolemy’s model of the solar system to Copernicus’s, and then on to Kepler’s. Tomasello (1999) calls this the ratchet effect of cultural selection, while Richerson and Boyd (2005) call it cumulative cultural evolution, and it is one of the most fundamental principles of the theory of cultural selection. Among other things, it is fundamental to explaining the rapid and radical changes that human environments—both social and physical—have undergone during the evolution of civilization.

In Chapters 2 and 3, I focus primarily on the role cultural selection plays in explaining the evolution of cultural traits—specifically, traits of the religious phenotype. In Chapter 4, however, I turn to discuss the role cultural evolution plays in explaining the evolution of human environments. This issue is important, because it provides a way for dual-inheritance theorists to respond to some of their most important critics. While everyone studying cultural evolution agrees that evoked culture accounts are an essential part of the story, controversy remains about whether there is also another part of the story. As I explain in the next section, Steven Pinker
argues that appeals to cultural selection are “superfluous,” while prominent advocates of the evoked culture approach present their own theories as competing, conflicting accounts, implying that the two approaches offer explanations of the same target facts, or phenomena. The role that cultural selection plays in explaining the evolution of the social environment shows both that it explains facts that the evoked culture approach does not, and also that it is not superfluous.

1.5 Challenges to the Dual-Inheritance Approach

While appeals to cultural selection are becoming increasingly common, some prominent theorists of cultural evolution and evolutionary psychology continue to reject appeals to the dual-inheritance approach in general, and to its application in the study of religion, specifically. Explicit arguments against the approach are rare, but I discuss below the only one I am aware of, due to evolutionary psychologist Steven Pinker. Most scientists who reject the approach signal their rejection not by arguing directly against it, as a general framework, but rather by focusing on particular empirical theories derived from that framework. They present their own theories as competing alternatives, which provide better explanations for the relevant target phenomena.

One such challenge is directed specifically at the theory of religion provided by the dual-inheritance approach, which has been most fully articulated and defended in psychologist Ara Norenzayan’s recent book *Big Gods* (2013). In an article entitled “Explaining Moral Religions,” anthropologists Pascal Boyer and Nicholas Baumard (2013) provide an alternative account of the cultural evolution of religion, which is based on the evoked culture approach, rather than on the dual-inheritance approach. Though they never say explicitly what they take the relationship to be between these two cultural evolutionary theories of religion, the positive account they provide is
accompanied by a separate text box objecting to Norenzayan’s account, implying that they see the relationship as one of competition. As I’ve already mentioned, however, and as I argue in detail in Chapter 4, evoked culture theories in general do not, in fact, conflict with dual-inheritance theories. While evoked culture accounts explain behavior by identifying the functional properties of genetic traits, dual-inheritance accounts explain, instead, psychological capacities acquired through learning. Accordingly, I take issue here with the assumption that this theory provides a competing alternative to Norenzayan’s account, and with the implication that it provides a better explanation for the cultural evolution of religion.

1.5.1 The Challenge from Evoked Culture Theorists

When the meme approach was at the height of its popularity, anthropologist Dan Sperber (1996), along with his students Pascal Boyer (2001) and Scott Atran (2002), published objections to cultural selection theory in general, on grounds that imitation isn’t literally replication. In response, dual-inheritance theorists Joseph Henrich and Robert Boyd presented modeling work showing that these objections apply only to the meme approach, and not to the dual-inheritance approach (Henrich and Boyd 2002; Henrich, Boyd, and Richerson 2008). This appears to have swayed Atran, since he has subsequently teamed up with Henrich to publish a defense of the dual-inheritance theory of religion (Atran and Henrich 2010). Sperber and Boyer, however, continue to generate evoked culture theories, and continue to reject theories appealing to cultural selection. While they do not argue explicitly against the dual-inheritance approach itself, they continue to treat dual-inheritance theories as competing, conflicting accounts.

---

1 This, then, is a completely different way of arguing to the same conclusion I drew in Section 1.4.1, when comparing the dual-inheritance view to the meme view: what the dual-inheritance approach depends on is a notion of imitation (for traits), rather than a notion of replication (for individuals).
Both Sperber and Boyer have recently teamed up with yet another of Sperber’s former students, Nicholas Baumard, whose research focuses on the evolution of the psychology of fairness. Along with collaborator Jean-Baptiste Andre, Sperber and Baumard present a theory of the evolution of morality (2013), which Boyer and Baumard then build upon to present a theory of the relationship between morality and religion (2013). Both theories are based on the adaptive value of mutualism, a form of cooperative behavior in which participants enjoy mutual benefits to genetic fitness. This is not to say that participants enjoy equal benefits, however, since individuals engaged in mutualistic interactions may not invest equal costs in the interaction. What is important is that each person’s benefits be proportional to the costs she herself pays, rather than to the benefits others receive, so that, in the end, each participant really does end up with a net benefit from the interaction. Accordingly, these theories maintain that humans universally possess an innate sense of fairness, based on intuitions about proportionality. These intuitions allow us to determine whether, and in what ways, purportedly mutualistic interactions have actually turned out to be mutually beneficial for all parties involved. I thus refer to these theories as “proportionality theories,” and while Baumard, Andre and Sperber offer a proportionality theory for morality in general, Baumard and Boyer offer a proportionality theory of moralizing religions, specifically.

The proportionality theory of morality is not presented primarily as a theory of cultural evolution, but as Baumard, Andre and Sperber point out (p. 70-71), the account does lend itself naturally to the evoked culture approach. On this approach, recall, different specific behaviors are evoked from the same universal, genetically inherited traits, because these traits respond differently in different environments. Just as the same genetic adaptations for thermoregulation cause people in cold environments to wear more clothing than people in warm environments, the
same genetic adaptations for monitoring proportionality are said to lead people in different social environments to make different judgments about fair distribution, fair punishment, and other questions about fairness. Baumard and Boyer then exploit this feature of the proportionality theory of morality to develop a corresponding theory of the cultural evolution of religion. Noting that most religious systems throughout human history have been based on gods who were believed to take no interest in the moral behavior of humans, they attempt to explain how specifically moralistic religious systems have now become so widespread among humans in general.

I discuss the details of both proportionality theories in Chapter 4, where I argue that both provide partial explanations for the cultural evolution of morality and religion, respectively, but that these partial explanations must be complemented by additional explanations based on cultural selection. In their article on morality, Baumard, Andre and Sperber disagree with this conclusion, and they provide an explicit argument to show that their evoked-culture account conflicts and competes with the account of moral evolution provided by dual-inheritance theories. I consider this argument in detail in Chapter 4, and reject it. But Baumard and Boyer do not discuss the nature of the relationship between these two theories of the cultural evolution of religion. It is clear that they view the dual-inheritance account of religion as a competing theory, because they frame their own account as an explanation of the same general phenomenon that the dual-inheritance theory also explains, and they explicitly object to the dual-inheritance theory itself. What is not clear is why they take it to be a competing account.

For many years, dual-inheritance theorists have motivated their appeals to cultural selection by showing that it explains a very important feature of human cooperation: its size and scale (Richerson and Boyd 1998; Boyd and Richerson 2006; Chudek, Zhou, and Henrich 2013).
They observe that the evolutionary mechanisms of kin selection and reciprocal altruism, or mutualism, explain why most species cooperate only in small groups, while a few species (e.g., eusocial insects) cooperate in large groups composed entirely of genetic relatives. But they also point out that no other species cooperates in enormous networks composed of millions of genetically unrelated individuals, as the human inhabitants of modern nations do. This unique form of cooperative behavior is what Boyd and Richerson call “the puzzle of human cooperation,” and they argue that cultural selection, and specifically cultural group selection, provide a partial solution to this puzzle. Groups possessing prosocial beliefs and norms, they claim, which cause individuals to invest more heavily in cooperation and collective action, have an advantage in competition with other groups, whose beliefs and norms place less emphasis on working together. And one of the most important mechanisms by which prosocial behavior increases group fitness—or, more precisely, the cultural fitness of a group’s beliefs and norms—is by increasing the group’s size. All else being equal, big groups outcompete small groups, and prosocial norms and beliefs sustain individual-level cooperative interactions that help groups grow to enormous sizes.

The dual-inheritance account of religion then builds from this foundation, claiming that beliefs about moralizing gods, who monitor, reward and punish the moral behavior of humans, increase prosocial behavior in general. So on this view, religious commitment increases prosocial behavior, and this, in turn, increases group size and group fitness. Consequently, the dual-inheritance theory of religion, too, has been motivated largely by appeal to the puzzle of human cooperation. Norenzayan summarizes the account as one in which two evolutionary puzzles—the emergence of large-scale cooperation and the emergence of moralizing religions—provide the solutions to each other: “Prosocial religions, with their Big Gods who watch, intervene, and
demand hard-to-fake loyalty displays, facilitated the rise of cooperation in large groups of anonymous strangers. In turn, these expanding groups took their prosocial beliefs and practices with them, further ratcheting up large-scale cooperation in a runaway process of cultural evolution” (2013, p. 8).

By contrast, none of the theorists from Sperber’s evoked-culture camp ever mentioned the evolution of large-scale cooperation until last year, even though they—and Boyer in particular—have been developing evoked culture theories of religion for twenty years. Now, however, they offer an evoked culture account designed specifically to explain what the dual-inheritance theory of religion explains: the relationship between religion and large-scale cooperation. And they suggest that the relationship is not what the dual-inheritance camp claims. They agree that moralizing religions and large-scale societies have evolved together, but they deny that religion has actually contributed to the rise of large groups. Instead, they claim, large groups evolved first, and independently, bringing with them material prosperity. And this, they suggest, “allows people to detach themselves from material needs (food, protection, affiliation)… and turn their attention to other domains of evolved preferences, such as maximizing personal wellbeing, enjoying friendship, and cultivating aesthetics, the good life that is portrayed as the goal of many moral movements” (Baumard and Boyer 2013, p. 277, Box 3). Moralizing religions, they propose, arose from the ranks of the affluent beneficiaries of large-scale cooperation, who had the spare time and spare resources necessary to turn toward moralizing, and away from the worldly concerns of biological need. And once these belief systems arose, they happened to resonate with or evoke intuitions about proportionality and fairness, which caused them to spread. Thus, on this view, the rise of large-scale societies is a pre-condition for the emergence of moralizing religions, but moralizing religions are not, in addition, a cause of the rise of large-
scale groups. Where Norenzayan describes a reciprocal, two-way causal relationship between moralizing religions and large-scale cooperation, Baumard and Boyer describe a one-way relationship, in which moralizing religions are a consequence of the rise of large groups, but they are not also a factor contributing to the rise of large groups.

This positive account that Baumard and Boyer provide is underdeveloped, at best. What is it about merely having free time that would cause people to devote that time to moralizing in a specifically religious way? In a time before moralizing movements had already arisen, what explains why moralistic motivations appear on a list of evolved preferences that otherwise includes maximizing personal wellbeing, enjoying friendships, and cultivating aesthetics? Moreover, as I argue in Chapter 4, in addition to concerns of fairness, moralizing religions address many other domains of moral concern that have little to do with fairness, and that are not explained by intuitions about proportionality. Finally, Norenzayan cites ethnographic evidence showing that moralizing gods are more prevalent in societies characterized by forms of material insecurity, such as chronic water shortages, rather than by material prosperity (2013, p. 129).

I am also unconvinced by the critical claims Baumard and Boyer level against the dual-inheritance account. They make a straw man of their opponent. The theory they object to holds that, all else being equal, religious beliefs and practices are one factor contributing to the rise of large groups. The claim that Baumard and Boyer refute, however, is the much stronger claim that moralizing religions are “the magic bullet that made large-scale societies possible” (p. 277, Box 3). This claim is easily refuted in the manner they pursue, by identifying six ancient empires that were large, but that did not have moralizing religions. But while these counterexamples show that religion is not necessary for the emergence of large-scale cooperation, dual-inheritance theorists have never suggested otherwise. The evolutionary trend they identify characterizes
intergroup conflict taking place across thousands of groups over thousands of years. Citing a handful of counterexamples hardly amounts to an adequate consideration of the relevant evidence.

Nevertheless, my primary concern here is not with the empirical details of the relationship between religion and large-scale cooperation. It is, rather, with the assumption that these two accounts of that relationship conflict, or compete. Baumard and Boyer clearly present their theory as the better of two competing explanations, on the assumption that both theories explain the same target phenomenon. But there are, in fact, two different ways in which they do not explain the same target phenomenon.

First, while their account offers a way of explaining, in evoked-culture terms, how the emergence of large groups contributed to the moralization of religion, this provides no reason to deny the existence of the other side of the reciprocal relationship that Norenzayan describes. That is, to say that the evolution of large-scale society contributed to the moralization of religion is perfectly consistent with the additional claim that, in turn, the moralization of religion also contributed to the evolution of large-scale society. When Norenzayan refers to a “runaway process of cultural evolution,” he is positing a mutually reinforcing positive feedback loop, like the evolutionary “arms race” that caused both cheetahs and gazelles to become fast runners. But to say that the speed of the cheetah contributed to the speed of the gazelle is not to deny that the speed of the gazelle also contributed, in turn, to the speed of the cheetah. Thus, even if it is true that the material security afforded by large groups contributed to the spread of moralizing religions, that would be no reason to deny that, once a moralizing religious system had spread, it could further contribute to the growth of the group as well.
The other reason why the two accounts do not conflict is the one I discuss in Chapter 4. Evoked culture accounts use genetic selection to identify the functional properties of genetically inherited traits. Dual-inheritance accounts recognize this method of psychological explanation, and assume that it is fundamentally important. But when they appeal to cultural selection to explain various forms of belief and behavior, they do not do so by appealing to genetic selection, or genetically inherited traits. Evoked culture accounts and cultural selection accounts systematically explain different target phenomena, because one approach explains genetically inherited psychological capacities, while the other explains culturally inherited, socially learned psychological capacities. Accordingly, even when we follow Baumard and Boyer, and focus on the contribution that large-scale society makes to the moralization of religion, we are only focusing on one contributing factor. Even if fairness intuitions are evoked by moralizing religions, and even if that facilitates the spread of moralizing religions over other religions, that is no reason to deny that certain culturally inherited beliefs and norms also facilitate the spread of moralizing religions. One reason for this is, as I’ve noted, that there is more to religion than fairness. Socially learned norms and beliefs about keeping the Sabbath or avoiding pork do not spread because they evoke intuitions about proportionality, because they have nothing to do in the first place with negotiating interactions that are supposed to be mutually beneficial. Another reason, however, is that even within the domain of fairness there are questions about which socially learned beliefs and practices do a better job of evoking intuitions about proportionality. The account Baumard and Boyer present explains what the Muslim faith and the Catholic faith have in common that makes both successful. It does not explain what happens when both belief systems appear in the same population, or when groups possessing these two different moralizing religions come into competition with one another.
In light of all this, I see no reason to believe that Baumard and Boyer’s evoked culture account provides a *better* explanation for the cultural evolution of religion than Norenzayan’s account does, for the simple reason that it simply explains different facts. Like any other evoked-culture account, it provides a partial explanation for the facts of cultural evolution. Cultural selection provides a different partial explanation, which is complimentary. Thus, even if Baumard and Boyer do attempt to provide a stronger, more developed version of the positive account they sketch here, that will in no way threaten the dual-inheritance account of the relationship between moralizing religions and large-scale societies.

### 1.5.2 Is Cultural Selection “Superfluous”? 

Among those who reject theories based on dual-inheritance, only Pinker (2010; 2012) has, to my knowledge, offered an argument that applies to the dual-inheritance approach in general, rather than to some specific theory derived from it. Interestingly, this argument bears a striking resemblance to my own argument against the meme approach (Section 1.4.1), in that it is an argument from parsimony. Pinker’s complaint is that appeals to cultural selection are “superfluous” (2010, p. 8996). They involve a “radical revision of evolutionary theory” (p. 8996) based on “exotic evolutionary mechanisms” (p. 8993), but these additional ontological commitments don’t pay for themselves—they don’t help us to explain and predict anything we couldn’t already explain and predict without them.

The important difference between Pinker’s argument and mine, however, is that while I can provide a meme-independent way of explaining all of the empirical facts explained by meme theories, Pinker rejects the dual-inheritance approach without providing any alternative explanation for many of the empirical facts explained by dual-inheritance theories. He evidently
does not recognize the existence of these facts, but it’s unclear whether this lack of recognition takes the form of ignorance or of outright denial. Either way, however, since he provides no argument to show that the relevant facts don’t exist, he provides no reason for anyone else to deny or reject them. Accordingly, I conclude that Pinker provides no reason to think that the metaphysical commitments of the dual-inheritance approach are superfluous.

In a 2010 paper, Pinker sets out to solve an evolutionary puzzle. Given that our genetic ancestors had no reason to pursue “abstract intellectual feats such as science, mathematics, philosophy and law” (p. 8993), why do we, today, possess the psychological capacities to engage in such activities? Developing an idea first presented by Tooby and Devore (1987), he suggests that the solution to the puzzle is that humans occupy a “cognitive niche,” an evolutionary trajectory in which selection has driven our lineage into an ever-increasing investment in cognition and information as a means for meeting our biological needs. As he puts it, “The cognitive niche differs from many examples of niches discussed in biology in being defined not as a particular envelope of environmental variables (temperature, altitude, habitat type, and so on), nor as a particular combination of other organisms, but rather the opportunity that any environment provides for exploitation via internal modeling of its causal contingencies” (p. 8995). But while Pinker includes “reasoning” as a basic adaptation to the cognitive niche, he does not think it is the only such basic adaptation. He takes the cognitive niche to depend on a “triad of adaptations” (p. 8998), which also includes capacities for grammatical language and for cooperating with others. Our abstract intellectual feats, then, are accomplished by re-deploying these capacities in new domains. While there has been no genetic selection for activities like science and philosophy themselves, the linguistic capacity for metaphorical abstraction allows us
to apply more concrete and practical forms of cognition to more abstract and intellectual
questions and aims.

Pinker considers it a primary advantage of this theory that it provides a general explanation of the evolution of human psychology “without resorting to exotic evolutionary mechanisms” (p. 8993). “Given the undeniable practical advantages of reasoning, cooperation, and communication,” he claims, “it seems superfluous, when explaining the evolution of human mental mechanisms, to assign a primary role to macromutations, exaptation, runaway sexual selection, group selection, memetics, complexity theory, cultural evolution (other than what we call “history”), or gene–culture coevolution (other than the commonplace that the products of an organism’s behavior are part of its selective environment)” (p. 8996). Among the eight “exotic” evolutionary mechanisms listed here are three to which the dual-inheritance framework is committed: group selection (but only for cultural traits), cultural evolution (by which Pinker means cultural selection, as dual-inheritance theorists see it, not as memeticists see it), and gene-culture coevolution (which I discuss at length in Chapter 2, and again in Chapter 4). So not only does Pinker seem to think that the theory of the cognitive niche obviates the entire dual-inheritance framework, but he also takes it to obviate the need to appeal to a number of other evolutionary theories, which have been proposed to explain other features of human cognition and behavior.

While my concern here is specifically with Pinker’s argument against the dual-inheritance approach, it’s important to see that this argument is embedded within a larger argument that is incredibly—indeed, absurdly—ambitious. In order for this argument to go through, Pinker would ultimately need to show that the three general psychological adaptations he appeals to, “reasoning, cooperation, and communication,” are the only capacities necessary for explaining
all of the data explained by all of these other theories. If any of these theories explains facts about human cognition and behavior that aren’t explained by the three general capacities Pinker identifies, then the evolutionary mechanisms involved aren’t superfluous after all. But Pinker provides no general reason or a priori principle showing that the data explained by these theories actually can be explained by the capacities he appeals to, and also he makes no other attempt to do so. Regarding memetics, of course, I agree with Pinker; those evolutionary mechanisms really are superfluous. But in defending that claim, I have provided principled, a priori reasons why any empirical fact explained by the meme approach can also be explained by the dual-inheritance approach. Pinker offers no similar reasons for thinking that something similar can be done for each of these other theories.

Accordingly, while I am in no position to defend theories of macromutation, complexity theory or runaway sexual selection, I would be genuinely surprised if all of the observations these theories explain can, in fact, be explained by some combination of reasoning, language and cooperation. I can speak in defense of exaptation, since this will be the topic of Chapter 3. But my own appeals to exaptation depend on prior commitments to cultural selection, gene-culture coevolution, and cultural group selection, so I will put exaptation aside here in order to focus on defending the dual-inheritance approach.

Before I proceed, I must add that when Pinker notes parenthetically that cultural selection is something “other than what we call ‘history,’” he is alluding to a different but closely related argument, which he develops elsewhere (2012). In an online article devoted primarily to arguing against genetic group selection, Pinker also takes the time to lodge an objection directed specifically at cultural group selection, again on grounds that it is superfluous. This time,
however, his claim is that it is conventional history, rather than adaptation to the cognitive niche, that makes cultural selection superfluous.

It won't work to switch levels [to the cultural level] and say that group selection is really acting on the norms and institutions of successful states. The problem is that this adds nothing to the conventional historian's account in which societies with large tax bases, strong governments, seductive ideologies, and effective military forces expanded at the expense of their neighbors. That's just ordinary causation, enabled by the fruits of human ingenuity, experience, and communication. The truly Darwinian mechanisms of high-fidelity replication, blind mutation, differential contribution of descendants to a population, and iteration over multiple generations have no convincing analogue. (Section 3, final paragraph)

In light of these claims, I interpret Pinker’s overall argument as follows: the posits of cultural selection theory (which include gene-culture coevolution and cultural group selection) are rendered superfluous by some combination of the adaptations of the cognitive niche and the facts and methods of conventional history. I’ll refer to these as the *psychological objection* and the *historical objection*, respectively. I take the psychological objection to be the claim that to explain the evolution of reasoning, language, and cooperation is to explain everything we need to know about the evolution of human psychology. I take the historical objection to be the claim that conventional history explains everything we need to know about the evolution of particular
cultural groups, or societies. In responding to this argument, I’ll address the historical objection first.

First, there are a number of confusions embedded in Pinker’s views about what it means for cultural selection to be a “truly Darwinian mechanism” capable of explaining things in a way that is distinct from the “ordinary causation” of conventional history. Evidently, he assumes that this requires that cultural selection be a convincing analogue of genetic selection, since that is a selection process that involves high-fidelity replication, blind mutation, differential contribution of descendants to a population, and iteration over multiple generations. But as I’ve already explained, this assumption is simply false. The explanatory value of cultural selection in no way depends upon its being analogous to genetic selection. The relevant question is simply whether treating the spread of cultural traits as a process of selection allows us to explain and predict things we couldn’t otherwise explain and predict. In addition, however, Pinker demonstrates here that he isn’t distinguishing between Darwinian selection in general and genetic selection in particular. For while genetic selection is based on high-fidelity replication, we’ve already seen that replication in general is not essential for selection-based explanation. Inheritance of traits is sufficient. Moreover, while genetic selection is based on blind mutation, and while that is a philosophically important fact, blind mutation is not even essential to the role selection plays in explaining genetic evolution. There is no reason why an adaptive mutation created by God could not subsequently be selected. One of the many ways in which cultural selection and genetic selection happen to differ is in virtue of what Richerson and Boyd call “guided variation” (2005), in which cultural innovations are deliberately designed in ways that help them to spread through populations; the cultural equivalent of a mutation created by God to be adaptive. It is essential to a Darwinian selection process that traits be heritable once they have arisen, but inheritance need
not depend on replication, and how traits arise in the first place is no part of the process of selection itself. This is why population geneticists view mutation as one of the other causal factors—besides selection—that is responsible for explaining evolution.

In addition, there are two other problems with the claim that cultural selection adds nothing to conventional history. First, it misses the point. Second, it is false.

The theory of cultural selection is, first and foremost, a psychological theory, not a historical theory. The point is to use the facts of history to explain the traits of psychology, just as evolutionary biologists use the facts of history to explain traits of morphology or physiology. The dual-inheritance approach has primarily been employed to explain human cognition and behavior, so in evaluating whether its metaphysical commitments earn their keep, these are the explanations that need to be evaluated (e.g., House et al. 2013; Chudek et al. 2011; Bauer et al. 2014; Henrich et al. 2006). Pinker’s claim is that cultural selection adds nothing to the historian’s account of history, but that simply misses the point of the theory. Models of cultural selection provide a way of organizing the facts of history for psychological purposes, by focusing the scope of historical examination specifically on facts about what people have learned from other people, and then using patterns among these facts to explain the psychology of individuals in the present. As I argued in Section 1.4.2, these models do this both by explaining the functional properties of cultural traits themselves, and by explaining the evolution of particular social environments, which then explain, in turn, both the social learning and the individual learning that people inhabiting those environments undergo.

Moreover, while the primary applications of cultural selection theory have been psychological, rather than historical, this framework nevertheless does add something to conventional history. For example, Slingerland and Chudek (2011) use patterns of cultural
selection as evidence for the claim that folk intuitions supporting mind-body dualism were prevalent in ancient China. If this is true, then it would appear to be a new and interesting historical fact, which would not have been identified without considering selection among culturally inherited traits. It may be possible, in principle, to demonstrate the same fact using the methods of conventional history, but that is just to say that conventional history could be used to provide additional evidence for the same claim. It does not change the fact that patterns of selection provide evidence for this claim themselves. And in the same way, cultural selection has been employed to explain the relative frequency of religious beliefs involving “big gods,” such as the God of the Abrahamic traditions, who monitor, reward and punish the moral behavior of human individuals (cf. Norenzayan 2013). While this theory is aimed primarily at explaining functional features of religious psychology, it does so, in part, on the basis of a prior historical claim: that the growth of large-scale societies has to some degree been facilitated by psychological commitments to a certain type of religious belief. Again, this historical claim could also be supported by conventional historical narratives. But the cultural selection framework allows Norenzayan to integrate traditional history with independent facts about human psychology, providing a different source of evidence for the claim in question, which conventional history does not provide. And if it were possible to identify such large-scale, population-level patterns without the cultural selection framework, it is difficult to see why conventional historians never managed to do so.

Now to the psychological objection, the claim that the triad of adaptations to the cognitive niche explains all there is to explain about human psychology, such that other capacities explained by other evolutionary mechanisms are superfluous. I will simply focus on the two most obvious and important capacities explained by the dual-inheritance approach, which
Pinker’s theory of the cognitive niche does not address. The first is the capacity for imitation itself. The second is the psychology of norms: the set of capacities enabling individuals to identify, follow and enforce the norms of their social group.

Some very important facts that Pinker attempts to attribute to language capacities are not explained by language at all, but by capacities for social learning. It’s true, as he says, that “Grammatical language has clear advantages in the transmission of information. Because it allows messages to be composed out of elements, rather than drawn from a finite repertoire, it confers the ability to express an unlimited number of novel messages” (p. 8994). But when “transmission of information” refers to the communication of information, it is very different from the transmission of cultural information to which theorists of cultural selection often refer. To communicate information is one thing, and to actually learn from, adopt or imitate on the basis of the information communicated is something else entirely. One person can learn from another by simply observing her actions, without any linguistic communication occurring at all. This explains why social learning also occurs in chimpanzees, crows, rats, guppies and other species that have no capacity for language (Galef and Laland 2005). And not only is language not necessary for social learning, it also isn’t sufficient. If you tell me that you’re tired, then I learn something about you. But I don’t learn anything from you, and it is only by learning from you that I become psychologically similar to you in a way that can plausibly and usefully be treated as a form of inheritance. It is cultural inheritance, based on social learning, that makes the psychological activities of distinct individuals instances of the same psychological trait, and these similarities are also essential to the manner in which cultural knowledge accumulates over time, increasing in functional complexity and sophistication in the manner of the ratchet effect described in Section 1.4.2. It is only by preserving the same information across cultural
generations that this information is built upon, developed and improved upon over time. Accordingly, while Pinker aptly identifies the importance of this process of accumulation, he woefully mischaracterizes it, attributing it to a capacity for language, rather than a capacity for social learning:

An obvious interdependency connects language and know-how. The end product of learning survival skills is information stored in one’s brain. Language is a means of transmitting that information to another brain. The ability to share information via language leverages the value of acquiring new knowledge and skills. One does not have to recapitulate the trial-and-error, lucky accidents, or strokes of genius of other individuals but can build on their discoveries, avoiding the proverbial waste of reinventing the wheel. (p. 8995)

The adaptive benefits of social learning are *amplified* by language, to be sure. But efficient communication will not lead to the sharing or accumulation of new knowledge and skills without an additional capacity to learn from what is communicated. Because he conflates language and social learning, Pinker offers no way of showing that social learning can be explained without appeal to exotic evolutionary mechanisms.

Also absent from Pinker’s account is any mention of the psychological capacities underlying normative cognition and behavior. While he devotes a brief section specifically to the evolution of cooperation, he makes no mention of norms within it. Instead, he attempts to explain the evolution of human cooperation in general by appealing to mutualism, or reciprocity—forms of social interaction that are mutually beneficial for all participants involved. This allows Pinker
to explain “sociality” and “cooperation” in humans in a way that simply overlooks many important forms of social interaction and cooperation. In particular, as I argue in Chapter 4, it overlooks most of the domain of moral psychology, and it explains nothing about the role religious belief plays in generating social cohesion, facilitating norm learning, or motivating people to follow and enforce norms.

Pinker is right to note that mutualism can explain the evolution of “a suite of moral emotions such as sympathy, gratitude, anger, guilt, and trust, which impel an individual to initiate cooperation, reward reciprocators, and punish cheaters” (p. 8994). But this list does not include other important moral emotions, such as sanctity and shame, which play important roles in normative behavior. And moreover, while the emotions Pinker lists here do play functional roles in mutualistic interaction, they also play other roles as well. Sympathy may indeed cause someone to behave in a way that is more sensitive to the interests of her partners in a mutually beneficial interaction, but it will also cause her to respond to the sight of someone being injured in a car accident, and that has nothing to do with mutually beneficial interaction. Similarly, if someone feels that she has been betrayed by a member of her own ethnic group, she is likely to experience both anger and distrust, but not because the disloyal person has cheated on a mutually beneficial arrangement. Norms of loyalty, harm and sanctity, along with a wide range of other social norms, help members of the same group to “stick together,” and to cooperate even when doing so is personally costly. But these forms of cooperation, and the psychological capacities that make them possible, cannot be explained by selection for mutualism.

Both of these general capacities—capacities for social learning and for normative behavior—are fundamental features of human psychology, and crucial factors explaining the uniquely human evolutionary niche. But the triad of adaptations Pinker appeals to does not
include them, and he provides no reason to think that they can be explained without resorting to exotic evolutionary mechanisms. By contrast, these are the psychological capacities that dual-inheritance theorists have been focusing on explaining for years, using exotic mechanisms. One of the cornerstones of the theory of cultural selection is what Boyd and Richerson call the *tribal instincts hypothesis* (2005, p. 214), which holds that a unique combination of cultural group selection and gene-culture coevolution explains why humans possess genetic adaptations—‘tribal instincts’—for social learning in general, and for norm acquisition in particular. I summarize the details of this account in Chapter 4.

I conclude that Pinker’s argument fails for a number of reasons. Not only does it rest on false assumptions about what it would mean for the ontological commitments of cultural selection theory to provide worthwhile explanations and predictions, but it also rests on false assumptions about (1) what needs to be explained about human psychology, (2) what is explained by the theory of the cognitive niche, and (3) what is explained by theories based on the dual-inheritance approach. Pinker’s rejection of appeals to cultural selection appears to be motivated by an attitude of dogmatic conservatism regarding appeals to Darwinian selection in general. He has evidently decided that genetic selection at the individual level is the only selection process of any scientific value, but the reasons he provides for this decision amount to little more than a methodological prejudice. However, Pinker is not the only prominent theorist harboring this prejudice, and Chapter 4 is devoted to grappling with it directly, and in a more general manner.
1.6 Conclusion

Applying theories and methods from the science of cultural evolution to the target phenomena of the science of religious evolution not only helps us to explain those phenomena, but also helps us to identify how those theories and methods should be applied to cultural phenomena in general. While some specific theories based on the evoked culture approach are more convincing than others, it is likely that patterns of cross-cultural distribution for many religious traits can be explained in this way. But it is also clear that the evoked culture approach alone is insufficient to explain the evolution of religion, and this has important consequences for the study of culture in general.

The particular beliefs, norms and practices of particular religious systems are not themselves genetically inherited traits, and nor are they universal forms of psychological response that are simply evoked by the right environmental conditions. The belief that Muhammad ascended to heaven from a site in East Jerusalem, for example, is not just evoked in certain environments, in the way that a mountainous physical environment might uniformly evoke the belief that herding will be more profitable than farming. The particular traits of particular religions are acquired by imitation, a special capacity for social learning that only certain species of animal possess, and that no species possesses to an extent comparable to humans. The theory of cultural selection is based on the methodology of treating beliefs and practices acquired in this way as heritable traits, and then using models of selection borrowed from population genetics to track and explain how such traits spread through populations. This methodology has yielded a body of empirical theory explaining both why humans in general possess capacities for copying others in the first place, and why the cultural models available to particular individuals possess the traits that end up getting imitated. That is, cultural selection
explains the psychology of individuals not only directly, by explaining the functional features of their culturally inherited traits, but also indirectly, by explaining important features of the environments in which individuals do their learning by trial and error. As a result, cultural selection is useful not just for explaining traits that people acquire by social learning, but also for explaining traits they acquire through individual learning. Thus, to explain the psychology of religious individuals, and to explain how many religious traits evolved, appeals to both evoked culture and cultural selection will ultimately be necessary.

Two ways of understanding cultural selection itself have been put forward, but it is the dual-inheritance approach, not the meme approach, that has actually been responsible for demonstrating the explanatory power of cultural selection theory. We could, in principle, translate facts about human individuals into facts about memes, abstracting away from the social learning capacities of organisms to focus instead on the replication capacities of memes. But since the selection history responsible for identifying a meme’s capacities for replicating must ultimately be explained, in turn, by prior facts about its hosts’ capacities for social learning, adopting the meme’s-eye view will not enable us to explain or predict anything we could not already explain and predict. Thus, cultural selection is best understood as selection among heritable traits, which belong to individual organisms, rather than as selection among replicating entities viewed as individuals in their own right.

I conclude that, pace meme theorists like Dennett and Dawkins, the meme approach is superfluous, at least for scientific purposes. But, pace evoked culture theorists like Sperber and Boyer, and other skeptics of cultural selection, like Pinker, the dual-inheritance approach is not superfluous. To explain the cultural evolution of religion, we will ultimately have to combine theories of evoked culture with theories of dual-inheritance. And as I argue in Chapters 2 and 4,
this is the case for important non-religious traits as well, including many important features of moral psychology. And if religion and morality are best explained by a combination of evoked culture and dual-inheritance, then I see little reason to doubt that this will be the case for many other cultural phenomena as well.

Of the articles that follow, two—Chapters 2 and 3—focus primarily on what the theory of cultural evolution explains about the empirical phenomena of religion. The remaining chapter, conversely, focuses primarily on what the empirical phenomena of religion demonstrate about the theory of cultural evolution. Thus, each chapter focuses on identifying some basic way in which work from one of these disciplines makes a contribution to work in the other. But these primary conclusions do not exhaust the important ways identified in which the disciplines mutually inform and mutually reinforce one another. I hope it will gradually become clear that the intersection between them is a rich source of theoretical resources from which each field benefits in a wide variety of ways. My aim here has been to identify some of the ways that are most important.
Chapter 2

Group Selection in the Evolution of Religion: Genetic Evolution or Cultural Evolution?

2.1 Introduction

In the blooming scientific literature on the evolution of religion, all theories explain the traits of religion by appealing to natural selection. Theorists differ, however, with regard to the types of selection to which they appeal. Some adhere closely to the biological orthodoxy of inclusive fitness theory, appealing only to genetic selection occurring at (or below) the level of the individual (e.g., Boyer 2001; Barrett 2004; Bering 2006; Bloom 2007). Others adopt a broader adaptationist perspective, appealing to cultural evolution in addition to genetic evolution, and to multilevel selection instead of selection at only the individual level (e.g., Wilson 2003; Haidt 2012; Atran and Henrich 2010; Norenzayan and Shariff 2008; Bulbulia and Mahoney 2008).

And since both genetic evolution and cultural evolution may occur at either the individual level or the group level, this broader approach recognizes no less than four distinct forms of natural selection that could potentially explain the evolution of a given trait. I thus distinguish between the theoretical framework of broad adaptationism, which recognizes four distinct forms of selection, and narrow adaptationism, which recognizes only one.

Naturally enough, narrow adaptationists tend to focus specifically on religious traits that are amenable to explanation by genetic selection at the individual level, while broad adaptationists focus on other kinds of traits. Here I consider a religious trait that narrow adaptationists have not attempted to pursue (perhaps for principled reasons, as I explain below), but for which two competing accounts have been offered by broad adaptationists. The trait in
question is religious altruism—altruistic behavior that is caused specifically by religious beliefs and practices. Both theories attempt to explain this trait by appealing to group selection, but in the broad adaptationist framework there are two different ways of doing this. One is to claim that group selection explains the genetic evolution of religious altruism. The other is to claim, instead, that group selection explains the cultural evolution of this trait.

I argue here against the gene-based approach: even in its most sophisticated form, the account of religious altruism based on genetic group selection is conceptually flawed. Moreover, diagnosing this flaw is important, because its implications extend beyond the study of religion to the study of human evolution more generally.

Biologist David Sloan Wilson (2003) was the first to appeal to genetic group selection in a theory of religion, but his account is no longer the most fully developed version of this approach. Psychologist Jonathan Haidt (2012) has recently adopted Wilson’s account, and integrated it into his own theory of moral psychology. This brings to Wilson’s approach a rich, detailed description of the psychological traits being posited as targets of group selection, allowing Haidt to say much more about precisely what it is that is said to be selected at the group level. More importantly, however, Haidt also says more about the details, as he sees them, of the relationship between genetic evolution and cultural evolution. It’s in these details that the confusion arises. Haidt makes a serious attempt to integrate his appeals to genetic group selection with his appeals to cultural evolution, but he nevertheless becomes tangled in the complex web of evolutionary dynamics that arises within the broad adaptationist framework. And when the source of the confusion is identified, it reveals general reasons to remain skeptical about genetic group selection in humans—reasons that do not apply to cultural group selection as well. As a result, the failure of Haidt’s account is significant for two reasons.
First, with regard to religious altruism, it renders the account based on cultural group selection the only viable theory available. Since there are only two accounts on offer, the failure of one of them leaves only the other one to explain religious altruism. Of course, to show that a theory is “the last man standing” is not exactly to show that it is true, and the task of marshalling positive evidence in its favor is not one I take up here. Fortunately, however, this evidence exists, and a growing number of scientists has been marshalling it elsewhere (Norenzayan and Shariff 2008; Norenzayan 2013; Atran and Henrich 2010; Henrich and Gervais 2010; Bulbulia and Mahoney 2008; Atkinson and Whitehouse 2011).

Second, the failure of Haidt’s appeals to genetic group selection also carries important implications for the wider application of the theory of group selection in social science. For while it shows that appeals to genetic group selection should be rejected, it also shows that this is no reason to reject appeals to group selection altogether. This is important, because it illustrates how social scientists may safely appeal to group selection without making questionable biological commitments.

Though he adopts D. S. Wilson’s theory of religion as the basis of his account, Haidt also follows Edward O. Wilson, founder of the field of sociobiology, in emphasizing evolutionary similarities between humans and other species, especially eusocial insects (E. O. Wilson 1978, 2012). All three of these theorists take group selection in humans to be an instance of a more general biological pattern, and all three take group selection to explain altruism in humans in the same way that it explains altruism in other species. But in biology appeals to group selection are extremely controversial, even for eusocial insects (cf. Wilson and Sober 1994; Okasha 2006, Ch. 6). When E. O. Wilson and his collaborators published their group-selectionist account of eusociality (Nowak, Tamita, and Wilson 2010), the number of authors included on a response
paper objecting to it made the response look as much like a petition as a theoretical objection. So to explain human altruism by appeal to this more general biological pattern is to make the account of altruism in humans dependent upon a much larger, more general issue, and one plagued by long-standing and deeply entrenched debate.

Haidt accepts this burden, devoting a full chapter of his book to a “retrial” of the theory of group selection, which he claims has been “falsely convicted and unfairly banished” in biology (p. 191). But if it is true that genetic group selection does not occur in humans, then it isn’t true that human altruism is explained by this more general biological pattern, and appeals to group selection in humans do not depend on the outcome of this more general biological debate. Analyzing the problems in Haidt’s account thus illustrate how psychologists, anthropologists, sociologists and economists may employ the theory of group selection without making biological claims that are highly controversial.

2.2 What is Religious Altruism?

In an article from 2008, psychologists Ara Norenzayan and Azim Shariff review the behavioral and experimental evidence for religious altruism. And in his recent book-length defense of the cultural group selection account of religion, Norenzayan (2013) provides a detailed discussion of the proximate psychological mechanisms that connect beliefs about morally concerned supernatural agents with forms of behavior that he refers to interchangeably as “altruism,” “prosociality” and “cooperation.” Haidt disagrees about how such psychological traits evolved, but he does not disagree about what the relevant target traits are. The studies he cites (2012, p. 256) are the same ones Norenzayan and Shariff mention in the 2008 article, which include some
of their own studies. Indeed, Haidt simply adopts the hypothesis of supernatural monitoring that forms the core of Norenzayan’s account. In Haidt’s terms, “Creating gods who can see everything, and who hate cheaters and oath breakers, turns out to be a good way to reduce cheating and oath breaking” (2012, p. 256).

Here it becomes significant that both of these theories employ the broad adaptationist approach, because the meaning of the concept of altruism changes in the broad framework in a subtle but crucial way. To appeal to cultural evolution and group selection in addition to individual-level genetic selection is to weaken the conditions necessary for a behavior to count as an instance of altruism. This may partially explain both why narrow adaptationists have not attempted to explain religious altruism at all, and why broad adaptationists are often tempted to use other terms, such as ‘prosociality’ and ‘cooperation,’ to stand in for the term ‘altruism.’

Narrow adaptationists have famously insisted that to explain how altruism evolved is actually to explain altruism away. As Trivers put it, “Models that attempt to explain altruistic behavior in terms of natural selection are models designed to take the altruism out of altruism” (Trivers 1971, p. 35). And as Ghiselin put it, “Scratch an ‘altruist’ and watch a ‘hypocrite’ bleed” (1974, p. 247). The theory of inclusive fitness has succeeded in explaining many forms of behavior that are intuitively considered altruistic according to the ordinary, common-sense concept of altruism, but narrow adaptationists have taken this explanation to show that, in fact, the relevant behaviors don’t count as forms of altruism after all. The narrow framework explains so-called ‘altruism’ by showing that the behaviors in question aren’t actually costly, and thus don’t count as altruistic, because their overall contributions to fitness are positive, not negative. This shows that the definition of altruism being employed is based on overall costs; a behavior only counts as altruistic if it yields net costs in the currency of genetic fitness.
Broad adaptationists agree in defining altruism as behavior that is costly for the agent and beneficial for some other individual, and they agree that the relevant costs and benefits are to be measured in the currency of genetic fitness. What they do not agree about is what it means for altruism to be costly. According to broad adaptationists, to say that altruism is costly is not to imply that it reduces genetic fitness overall. In the broad framework, altruism implies costs, but it does not imply net costs. This difference is crucial, because if ‘religious altruism’ is taken to refer to behavior that confers net costs, as narrow adaptationists typically assume, then it is no longer clear that religious altruism even exists. Accordingly, it is easy to see why narrow adaptationists have not bothered to develop an account of religious altruism.

Altruism once posed an important challenge to adaptationist explanations of behavior in general, and narrow adaptationists met this challenge precisely by drawing a distinction between costs and net costs. At that time, narrow adaptationists had the choice of adopting either the weaker definition of altruism, based on simple costs, or the stronger definition, based on net costs. They chose the stronger definition, apparently for rhetorical reasons: claiming that altruism doesn’t even exist is a striking way of emphasizing the difference between costs and net costs that forms the basis of their account. However, it is important to see that there is no principled theoretical reason why narrow adaptationists could not have chosen the weaker definition instead. In the terms of inclusive fitness theory it makes perfect sense to say that foraging behavior, for example, is costly. In addition to metabolic costs for locomotion and cognition (if food is hard to find), foraging exposes many animals to increased risk of predation. These are all costs, and there is no principled reason why altruism could not be defined as costly in exactly the same sense in which foraging is costly. The difference is only that altruism possesses these costs by definition, while it is no part of the definition of foraging that it must be costly. This would
rob narrow adaptationists of the rhetorical effect of claiming that the behaviors they explain don’t count as altruistic, but it wouldn’t change anything about the substance of their explanation: just like foraging, these costly behaviors evolved because the costs involved were not net costs.

However, since narrow adaptationists did choose the stronger definition, they are not likely to regard as altruistic the behaviors that are explained by broad adaptationist theories of religious altruism. In the studies that Norenzayan and Haidt both appeal to, participants do such things as refraining from cheating even when their cheating can’t be detected, or giving money to charity or to strangers in economic games. These behaviors are obviously costly, and this suggests a clear sense in which prosocial behavior is something over and above ordinary social behavior. But nothing in the design of these studies shows that these forms of behavior are fitness-reducing overall. The relevant costs might easily be outweighed by other fitness benefits. For example, perhaps individuals who don’t cheat and give money to strangers acquire good reputations, and perhaps this ultimately increases their fitness by increasing their access to the benefits of cooperation.

Norenzayan and Haidt are both fully aware of this, as are broad adaptationists in general. They do not assume that these behaviors reduce fitness overall. Rather, when referring to such behaviors as ‘altruism,’ they take advantage of the fact that these behaviors are obviously costly, and thus obviously altruistic in the weaker sense. The fact that they are costly at all is already sufficient to justify use of the term ‘altruism’ according to the common-sense concept of altruism. However, it does not justify use of this term according to the technical definition adopted by narrow adaptationists a few decades ago. This may be one reason why, while broad adaptationists do use the term ‘altruism,’ they do not always do so consistently.
More importantly, unlike narrow adaptationists, broad adaptationists have no motivation to adopt the stronger definition of altruism, because they do not presuppose, a priori, that costs must be balanced against corresponding benefits. In the narrow framework, once some prevalent form of costly behavior has been observed, the question of overall fitness arises automatically, because observed costs must be explained by corresponding benefits. But cultural evolution and group selection each provide separate ways of explaining how a costly form of behavior can evolve even if it is costly, overall, from the point of view of individual genetic fitness (cf. Wilson and Wilson 2007, Richerson and Boyd 2005, Ch. 5). As a result, the existence of benefits to individual genetic fitness does not simply follow from the existence of costs to individual genetic fitness. It is an empirical question whether there are any such benefits, a question that will depend upon which form of selection explains how the trait evolved. In the broad framework, the question of a balance between costs and benefits in the currency of individual fitness arises only after it has been determined empirically (1) how individuals inherit the trait in question, and (2) what level of selection is responsible. For many of the traits of religion, and for many other forms of cooperative or prosocial behavior pursued by broad adaptationists, these questions do not have obvious answers.

In light of all this, it is necessary to say explicitly that in the broad adaptationist theories I will be discussing, ‘altruism’ refers to traits that are costly only in the weaker, short-term sense, not in the overall sense. This definition preserves the meaning of ‘altruism’ in ordinary, common usage, according to which giving money to a stranger is already considered altruistic regardless of its overall effect on the giver’s fitness. But this definition also conflicts with the stronger technical definition adopted by narrow adaptationists, according to which it is not clear whether giving money to strangers counts as altruism until the overall fitness value of the act has been
determined. Thus, according to the definition of ‘religious altruism’ that I employ here, to say that some behavior is altruistic is to say that it is costly only in the same sense in which foraging is costly.

2.3 Haidt’s Appeals to Genetic Group Selection

The basic, general form of the group-selectionist theory of altruism holds that altruistic traits spread by increasing the fitness of the altruist’s group relative to other groups, even if those traits simultaneously reduce the fitness of the altruist herself, relative to other individuals. If group-level benefits outweigh individual-level costs, then altruism can evolve in spite of its costs. In a collaborative paper, David Sloan Wilson and Edward O. Wilson (2007) offer the following summary of what they hope will become “sociobiology’s new theoretical foundation”:

“Selfishness beats altruism within groups. Altruistic groups beat selfish groups. Everything else is commentary.” And in other work, both Wilsons take this principle to explain altruism in humans in the same way that it explains altruism in other species, most notably ants, bees and wasps (E. O. Wilson 1998, 2012; D. S. Wilson 2003).

D. S. Wilson (2003) exploits this reasoning to develop a theory of religion. He gathers a wealth of evidence supporting the claim that successful religions increase altruistic behavior. For example, he describes the difference between the responses of Christians and pagans to a plague in Rome around 260 CE. While pagans fled, and pushed the sick out onto the streets before they were dead, many Christians fulfilled their religious duty to care for the sick, making martyrs of themselves in the process (2003, p. 153). This example does not show that these beliefs actually increased the fitness of the Christian group, because it does not show that Christians only

---

2 For a detailed survey of the general debate over levels of selection, see Okasha (2006).
provided nursing care to other Christians. But it does show that Christian beliefs succeeded in motivating an exceedingly costly form of altruism. The theory of group selection is then offered as an explanation for why such costly traits have not only persisted in spite of their costs, but have actually become far more frequent in the overall population.

It is D. S. Wilson that Haidt follows in appealing to religion, but it is E. O. Wilson that Haidt follows in placing heavy emphasis on the similarity between humans and eusocial insects. Haidt adopts the social behavior of bees as an emblem of the power of group selection, and he offers a slogan describing the role of this process in human evolution: “We are 90% chimp and 10% bee.” This claim is puzzling, however, because of one crucial dissimilarity between humans and bees. What group selection would explain in bees would be the evolution of altruistic genotypes; bees inherit altruistic traits through their genes. But even when humans inherit the specific religious beliefs and practices of their own parents, they do not do so by inheriting genes for those specific beliefs and practices. In the same way, genetic inheritance does not explain why people speak the same language as their parents. Haidt’s appeals to cultural evolution show that he takes traits like these to be transmitted culturally, rather than genetically, but this difference between humans and bees introduces important questions about what it means for us to be 10% bee.

In the theory of cultural evolution subscribed to by those who study religion, cultural transmission is defined as a specific form of social learning known as cultural learning. Cultural learning is high-fidelity copying, or imitation, and the key insight of cultural evolutionary theory is that imitation may be treated as a non-genetic form of inheritance—an alternative mechanism by which psychological traits may spread through human populations. Mathematical models based on this notion of inheritance rigorously describe a form of natural selection that is wholly
distinct from genetic selection (cf. Richerson and Boyd 2005). It is thus crucial to recognize that the basic, general form of the group-selectionist account of altruism remains neutral with regard to the mechanism of inheritance for altruistic traits. Altruistic behavior may be selected for at the group level whether individuals acquire their altruistic traits genetically or through cultural learning.

Accordingly, when Haidt claims that human evolution is similar to bee evolution, one might hope what he means is just that the logic of group-level selection is the same regardless of how the relevant traits are inherited. The claim that group selection explains altruism in both bees and humans is entirely compatible with the claim that bees inherit altruistic traits genetically, while humans have to learn to be altruistic. But Haidt thinks the similarity runs deeper, encompassing not just the general logic of group selection, but the mechanism of inheritance as well. That is, like D. S. Wilson (2003, p. 155, 175), he appeals not just to group selection, but to genetic group selection.

Haidt describes two general features of moral psychology that he thinks have been influenced by genetic group selection. First, he posits a mechanism called the hive switch, a “group-related adaptation” (2012, p. 223) to which he devotes a full chapter of his book. The hive switch is meant to be a capacity to “transcend self-interest and lose ourselves (temporarily and ecstatically) in something larger than ourselves” (p. 223). And according to Haidt, it “cannot be explained by selection at the individual level,” because it is “for making groups more cohesive, and therefore more successful in competition with other groups” (p. 223). Clearly, the hive switch is meant to be part of the 10% of our psychology that is bee-like, having evolved by the same process of genetic group selection that explains altruism in bees.
The other important effect of genetic group selection on moral psychology, according to Haidt, is quantitative rather than qualitative. At the core of his theory of moral psychology are six basic capacities that he calls *moral foundations*. Concerns about fairness, for example, are explained by appeal to a fairness foundation: a package of cognitive and motivational capacities that allows people to track the social distribution of goods and respond in functionally appropriate ways. The other foundations he posits are *care*, *loyalty*, *authority*, *sanctity* and *liberty*. And because Haidt is laudably concerned that this model should explain more than just the judgments of American university students, the moral foundations are posited as universal traits that originally evolved by individual-level genetic selection (the 90% that is part chimp).

So while Haidt supposes that these traits were already in existence before group selection began to take effect, he also argues that, in recent evolutionary history, group selection has modified their innate settings, or their relative levels of activity in various contexts (p. 216).

But while he makes these claims about genetic group selection, Haidt also assumes that the moral foundations of individuals may be affected by cultural learning, and are thus subject to cultural evolution. He does not view genetic group selection as a process in which group selection occurs in genetic evolution *instead of* cultural evolution. Rather, he sees it as a process in which genetic evolution occurs *through* cultural evolution. Here he takes himself to be appealing to the theory of *gene-culture coevolution* developed by Robert Boyd and Peter Richerson, according to which cultural evolution can exert a causal influence upon genetic evolution. However, as I explain in what follows, Haidt misunderstands the theory of gene-culture coevolution. For cultural evolution to causally influence genetic evolution is not for genetic evolution to occur *through* cultural evolution, in the way Haidt imagines. As a result, he mistakenly sees gene-culture coevolution as a way of integrating appeals to cultural evolution.
with appeals to genetic group selection. In fact, however, the theory of gene-culture coevolution provides general reasons to remain skeptical of appeals to genetic group selection in humans, as Boyd and Richerson themselves argue.

2.4 Why Group Selection in Humans is Unique

Whereas Haidt follows Wilson and Wilson in treating group selection in humans as an instance of a more general biological pattern, Boyd and Richerson take precisely the opposite approach. They too explain human altruism by appealing to group selection, but they motivate this appeal specifically by claiming that it explains what is unique about altruism in humans. They first note that altruism in other species occurs only in close-knit groups composed of genetically related individuals, although a small number of trusted partners in reciprocal exchange are sometimes included as well. They then point out (2006) that altruism in human cooperation stands out starkly against this background:

In contrast to the societies of other animals, virtually all human societies are based on the cooperation of large numbers of unrelated people. This is obviously true of modern societies in which complex tasks are managed by enormous bureaucracies like the military, political parties, churches, and corporations… Even contemporary hunter-gathers who are limited to the least productive parts of the globe have extensive exchange networks and regularly share food and other important goods outside the family. Other animals do none of these things. (p. 108)
Not all of this cooperation is altruistic, of course, but much of it is. For example, much to the chagrin of social conservatives, governmental systems of social welfare often benefit “have-nots” at the expense of “haves” who are not their genetic relatives, and whom they will never even meet. Further, people often donate anonymously to charities, tip at restaurants they don’t expect to return to, and give up seats to pregnant women on buses.\(^3\) None of these forms of altruism can be explained by genetic relatedness or reciprocity, and none of them occurs in other species.

Such uniquely human forms of altruism constitute what Boyd and Richerson call “the puzzle of human cooperation,” and they offer as a solution to the puzzle what they call the *tribal instincts hypothesis*. Haidt takes himself to be adopting this hypothesis, but the view he actually adopts is distorted by his misinterpretation of the theory of gene-culture coevolution.

2.4.1 Misinterpreting Gene-Culture Coevolution

That genetic evolution causally influences cultural evolution is a mundane point. It’s no surprise, for example, that cultural traditions for preparing sweet foods have been impacted causally by the events of genetic selection that explain why humans like sugar in the first place. What is striking about gene-culture coevolution, however, is that this causal relation can also run in the other direction: prior events in cultural evolution can also affect the subsequent course of genetic evolution. Since culturally inherited traditions can sustain stable forms of behavior that persist for centuries and even millennia, and since some of these forms of behavior have serious consequences for individuals’ genetic fitness, cultural evolution can create a social environment

\(^3\) In case it isn’t obvious where the fitness costs lie in giving up one’s seat on a bus, note that standing is more metabolically expensive than sitting down. This is why we become “tired” —motivated to *stop* standing—after standing for long periods, and why pregnant women become tired sooner than others.
containing selection pressures that are strong enough, and that last long enough, to influence genetic evolution.

Perhaps the most well documented example concerns the evolution of lactose tolerance among adults (Richerson and Boyd 2005, p. 192; Holden and Mace 2009; Laland, Odling-Smee, and Myles 2010). Before humans began keeping livestock, they were unable to digest milk beyond the first few years of age. But when people in some groups established cultural traditions for domesticating milk-producing livestock, this created an environment in which a new source of nutrition was readily available—if only it could be digested. In other words, the cultural evolution of livestock domestication created a selection pressure for genes enabling the digestion of lactose. Eventually, the genes responded, and today adults in these cultures (but not all cultures) have genes for digesting dairy products.

Since the distinction between cultural evolution and genetic evolution is based upon a prior distinction between genetic inheritance and cultural inheritance, it follows by definition that any genetically inherited trait that evolves by group selection is a product of genetic group selection, not cultural group selection. Haidt, however, interprets the thesis of gene-culture coevolution in such a way that cultural group selection produces genetic adaptations (2012, p. 210):

If cultural innovations (such as keeping cattle) can lead to genetic responses (such as adult lactose tolerance), then might cultural innovations related to morality have led to genetic responses as well? Yes. Richerson and Boyd argue that gene-culture co-evolution helped to move humanity up from the small-group sociability of other primates to the tribal ultrasociality that is found in all human societies. According to their “tribal instincts hypothesis,” human groups have always been in competition to
some degree with neighboring groups. The groups that figured out (or stumbled upon) cultural innovations that helped them cooperate and cohere in groups larger than the family tended to win these competitions (just as Darwin said).

Haidt is correct to say that, according to the tribal instincts hypothesis, cultural traits ("innovations") can be selected for at the group level. But this would be cultural group selection, not genetic group selection, and Haidt does not acknowledge this. Instead, he focuses on the "genetic responses" to group-selected cultural traits, drawing the conclusion that gene-culture coevolution makes it possible for genetically inherited traits (e.g., the hive switch and the innate settings of moral foundations) to be selected for at the group level.

The problem here is that while both genetic selection and group selection are occurring in this coevolutionary process, the traits selected for at the group level are not the traits that are inherited genetically. Like any other coevolutionary process, gene-culture coevolution involves two distinct processes of selection, each of which explains the evolution of different traits inherited in different ways. The tribal instincts hypothesis holds that cultural evolution for one trait (a culturally inherited one) may subsequently influence genetic evolution for a different trait (a genetically inherited one). But even if selection for the relevant cultural trait takes place at the group level, this would imply nothing at all about which level of selection is responsible for the relevant genetic trait.

In the same way, many traits of flowers have coevolved with traits of the bees that pollinate them. But if the relevant bee traits were to evolve by group selection, this would obviously be no reason to suppose that the traits of the flowers also evolved by group selection. Whether they did or not would be an entirely separate empirical question. In such a case, all of
the relevant traits are inherited genetically, yet distinct processes of selection explain bee traits and flower traits, and these distinct processes may have occurred at different levels. In gene-culture coevolution the difference between the two coevolving traits is in one sense even more pronounced, since they aren’t even inherited in the same way. Traits acquired by learning are distinct from traits acquired genetically, so distinct processes of selection are needed to explain the evolution of each. Just as in the coevolution of bee traits and flower traits, the fact that one trait evolves by group selection implies nothing at about the level of selection responsible for the other trait.

According to the tribal instincts hypothesis, group selection occurs only for traits inherited by cultural learning. While cultural group selection is said to influence the subsequent course of genetic evolution, the resultant process of genetic evolution is said to take place at the individual-level, not at the group level. Indeed, Richerson and Boyd explicitly argue against appeals to genetic group selection in humans, as I explain in the next section. Haidt clearly disagrees with the conclusion of this argument, but he makes no attempt to respond to it.

2.4.2 Against Genetic Group Selection in Humans

Richerson and Boyd’s argument against genetic group selection in humans is based on the role of variation in selectionist explanations, and it is a fairly technical argument. Accordingly, a brief digression on variation is necessary.

In order for selection to be what explains the frequency of a trait in a population, there must have been distinct variants of that trait to select from. Put metaphorically, if Mother Nature doesn’t have distinct options to select from, then her role as Selector doesn’t do any explanatory work. This is a matter of degree; the greater the variation, the more significant the role of
selection. If the giraffes in a population have wildly different neck lengths, then even weak selection can be an important part of the explanation for changes in average neck length. But if all the giraffes have necks of almost exactly the same length, then only very strong selection pressures will be capable of making such minor variations significant in an explanation for changes in average neck length.

Moreover, group-level selection requires group-level variation. If variation exists within a group—such that members of the same group possess different variants—then selection at the group level will not be selection for some variants over others; when the whole group gets selected, all of these variants are selected together, in a package deal. Variation within the group is variation among individuals, and it can only support individual-level selection. For selection to take place at the group level, members of the same group must possess the same variant, while members of different groups possess different variants. As it is often put, variation must be “concentrated” at the group level.

For the same reasons, genetic selection requires genetic variation. If the traits that vary are inherited culturally, rather than genetically, then by definition group selection acting on those traits would be cultural group selection. Of course, for any given psychological trait it is an open empirical question how individuals inherit it, and for some traits this question is quite difficult. But to explain a trait by appeal to selection is to assume that this empirical question has already been closed. To say a trait is a product of genetic group selection already entails that variants of this trait are inherited genetically, while to say it is a product of cultural group selection already entails that it is inherited by cultural learning. Since appeals to gene-culture coevolution always refer to two different forms of selection acting on two distinct traits, they also require reference
to two distinct forms of variation: one due to genetic differences, another due to differences in learning.

Given all of this, a major reason why Boyd and Richerson appeal to cultural group selection is that cultural evolution concentrates variation at the group level. Cultural inheritance causes individuals to become more similar to each other with regard to a given psychological trait, thereby reducing the variation between those two individuals. When the individuals involved are members of the same group, this reduces variation within the group. And when, as is often the case, the relevant trait is not also present in competing groups, this decrease in variation within the group is not accompanied by a corresponding decrease in variation between groups. That is what it means to concentrate variation at the group level.

By itself, of course, this is no evidence against genetic group selection. It shows, however, that by appealing to genetic group selection Haidt is making a substantial empirical commitment. Since genetic group selection presupposes the existence of group-level genetic variation, a crucial empirical question arises: how much group-level genetic variation is there among humans? Haidt never considers this question, and he provides no evidence for the required form of variation.

Boyd and Richerson do consider the question, and they argue that group-level genetic variation is virtually non-existent in humans. For any kind of genetic selection to occur, the relevant genetic variation must be continuously maintained throughout the course of the selection process. But genetic selection takes a long time, and during that time individuals tend to migrate, bringing their genes with them across group boundaries. This reduces genetic variation at the group level, concentrating it at the individual level instead. The question then becomes a matter of degree: if rates of migration are relatively low, or if group-level selection is relatively strong,
then genetic group selection may occur, in principle. In fact, however, Richerson and Boyd cite modeling work showing that this is not what happens: “Even very small amounts of migration are sufficient to reduce genetic variation between groups to such a low level that group selection is not important” (2005, p. 203).

In addition, a recent empirical study found that group-level genetic variation in humans is no higher than group-level genetic variation in chimpanzees (Langergraber et al. 2011). Since chimpanzees aren’t especially altruistic, this bodes ill for the hypothesis that genetic group selection explains why humans are especially altruistic. Chimpanzee groups also face conditions of significant intergroup competition, but chimpanzees don’t engage in widespread altruism toward strangers. If these psychological differences between humans and chimps were the result of genetic group selection in humans, as Haidt suggests, then we should observe higher levels of group-level genetic variation among humans than we do among chimps. According to this study, we don’t.

Once these errors in Haidt’s evolutionary reasoning are identified, we are left with no good reason to suppose that genetic group selection explains human altruism. Indeed, what we are left with are positive reasons to think otherwise. And because this argument against genetic group selection is based on general concerns about the amount of group-level genetic variation in humans, the problem with Haidt’s appeal to genetic group selection is not specific to religiously motivated forms of altruism, or even to altruistic traits in general. Any theory that purports to explain any trait in humans by appealing to genetic group selection will have to overcome this obstacle. Appeals to cultural group selection do not incur this burden.
2.5 Comparing Genetic and Cultural Group Selection in Context

Though Haidt doesn’t use the same terms as Boyd and Richerson, he too appeals to the fact that cultural evolution increases the power of group selection. And he does so in a passage that is crucial for understanding his evolutionary account, because it contains an example that provides a concrete statement of the details, as he understands them, of genetic group selection. But the passage also exemplifies the errors I’ve been describing, and identifying the effect of these errors in the context of a concrete example is instructive, because it clearly illustrates the difference between the correct version of the tribal instincts hypothesis and the incorrect version that Haidt relies upon. In Haidt’s version, there is no difference between genetic group selection, on one hand, and cultural group selection that leads to gene-culture coevolution, on the other. In the correct version, when cultural group selection leads to gene-culture coevolution, the resulting process of genetic evolution takes place at the individual level, not at the group level.

The passage in question focuses on altruism in the specific context of soldiers’ temptation to defect or desert. When faced with a choice between risking one’s own life for the group and deserting to preserve one’s own health, individuals face a temptation to free ride on the altruism of others, selfishly leaving it to them to pay the costs of collective defense. As Haidt points out, however, real armies have ways of addressing this temptation:

In a real army, which sacralizes honor, loyalty, and country, the coward is not the most likely to make it home and father children. He's the most likely to get beaten up, left behind, or shot in the back for committing sacrilege. And if he does make it home alive, his reputation will repel women and potential employers. Real armies, like most other effective groups, have many ways of suppressing
selfishness. And any time a group finds a way to suppress selfishness, it changes the balance of forces in a multi-level analysis: individual-level selection becomes less important; group-level selection becomes more powerful. (2012, p. 194)

Though he does not mention variation, Haidt is correct in noting that whenever a group finds a way to suppress selfishness, this makes group selection more powerful. To say that a group has found a way to suppress selfishness is to say that individuals within that group have become more similar to each other with regard to some altruistic phenotype, and thus that variation on this trait has been concentrated at the group level. This would indeed make group selection for those traits more potent, just as Haidt says. But what kind of variation are we talking about here, cultural or genetic?

Haidt focuses explicitly on genetic variation. The passage above continues:

For example, if there is a genetic basis for the loyalty and sanctity foundations—which can support this sort of sacralized groupishness—then these genes can become more common in the next generation, because groups in which they are common replace groups in which they are rare, even if these genes impose a small cost on their bearers (relative to those that lack them within each group). (p. 194-95)

The appeal to a “genetic basis" for the moral foundations of loyalty and sanctity shows that Haidt takes himself to be describing genetic group selection. But this sort of genetic basis is not the sort capable of supporting genetic group selection. That traits of loyalty and sanctity are genetically
inherited is entirely plausible, but it does not show that the relevant genes vary at the group level. Groups with higher frequencies of genes for loyalty may indeed outcompete groups with fewer genes for loyalty, and they may pass those genes on to their younger generations. But if individuals migrate between groups of these two types at even a low rate over the course of the purported process of selection, then selection of the loyal group won’t be selection for loyal genes. It will be selection of a group that possesses both types of genes, loyal and disloyal alike. Cultural inheritance, by contrast, can maintain such group-level differences in spite of the flow of genes across group boundaries.

Further, the traits that actually provide the group-level variation in Haidt’s example here are not the same traits that he claims are inherited genetically. The advantages to group fitness in this example come from social norms that specify how deserters should be punished. Some groups’ norms will be more effective than others at keeping soldiers on the battlefield, and this will increase group-level fitness in military competition. But the psychological traits that constitute these group-level differences are not the traits of soldiers who face a decision about whether to fight or run. They are, instead, the traits of other people: the generals who make the rules, the officers who enforce them, and the women and employers back home who refuse to mate with and employ selfish cowards. This means that the traits that actually produce the relevant variations in group-level fitness involve doling out punishment to deserters. But the traits that Haidt claims have a genetic basis are the traits of the soldiers themselves, and they have to do with avoiding punishment, not doling it out.

While some groups really have shot deserters in the back, U. S. Military officers certainly do not do this today, because their group’s norms for punishment specify that individuals must be given a fair trial before execution. In Haidt’s account, differences between groups with regard
to such norms are the source of group-level variations in fitness. But when we look at the traits of individuals that constitute these group-level differences, it seems unlikely that they are inherited genetically. I doubt Haidt supposes that U. S. Military officers are genetically disposed to give deserters a fair trial before execution, while people in other groups are genetically disposed to shoot first and ask questions later. Rather, U. S. Military officers have simply adopted, by learning, the practices of their cultural predecessors. Differences of this kind among the norms and values of human groups are exactly the sort of differences that led Haidt to appeal to cultural evolution in the first place. He is right to say that they enable group-level selection to occur, but it follows that the variants that are selected for at the group level are cultural variants, not genetic variants.

That traits like loyalty and sanctity are inherited genetically would support the claim that they evolve by genetic selection, but it would not show that they evolve by genetic group selection. And when we explicitly consider whether selection for these traits would occur at the group level or at the individual level, the answer is clear. Haidt’s claim is that, because of the punishments doled out by other members of the group, loyal and sanctimonious individuals would have higher fitness than selfish and cowardly individuals in the same army. This means that the relevant variations in fitness appear within a single group, rather than between different groups. That is the definition of selection at the individual level.

Once we distinguish in this account between (1) the culturally inherited traits that lead people to dole out punishment, and (2) the genetically inherited traits that lead people to avoid such punishment, it becomes clear that the coevolutionary story Haidt means to tell is ultimately the one that Boyd and Richerson originally proposed. In this story, an antecedent process of cultural evolution for norms against desertion changes the social environment in which soldiers’
decisions about desertion take place. This changes the payoffs of those decisions, such that it is more costly overall to desert and face punishment than it is to fight and risk injury or death in battle. Culturally evolved norms create a genetic selection pressure by making desertion too costly from the point of view of genetic fitness, so after such norms have evolved soldiers find themselves in a social environment in which it is adaptive to stay and fight. But that is to say that individuals who stay and fight will have higher genetic fitness than other individuals within the same group. And selection among individuals in the same group is individual-level selection, not group selection.

This is a straightforward illustration of gene-culture coevolution leading to selection for genetically inherited traits such as loyalty and sanctity. But the role of group selection in this account is entirely confined to culturally inherited traits. There is genetic selection in response to a prior process of cultural selection, but the relevant genetic traits are selected for at the individual level, not at the group level. There is also a process of group selection occurring, but the traits that vary at the group level are inherited culturally, not genetically. There is no genetic group selection in this account, and this coevolutionary process is very different from the process of genetic group selection that explains altruism in bees.

2.6 Group Selection in the Evolution of Religion

Though Haidt never responds to Richerson and Boyd’s arguments against genetic group selection, his rejection of the distinction between cultural group selection and genetic group selection is deliberate, and not simply a matter of overlooking technical details. In the chapter he devotes specifically to religion, he argues against the theory offered by Atran and Henrich
(2010), who also offer a broad adaptationist theory of religion based on Richerson and Boyd’s the tribal instincts hypothesis. And Haidt criticizes them specifically because they appeal to “purely cultural group selection,” and “doubt that there has been any genetic selection for religion” (p. 258; Haidt’s emphasis). Haidt considers this implausible, he says, because it underestimates how quickly genetic evolution can occur in response to cultural evolution (p. 258). But because Haidt misunderstands the nature of gene-culture coevolution, this objection simply misunderstands their theory. Nothing they say is inconsistent with the possibility of very recent gene-culture coevolution, but, more importantly, the rate of gene-culture coevolution simply isn’t relevant to the question whether group selection for religion has been cultural group selection or genetic group selection.

As the mention of “purely” cultural group selection already hints, Haidt takes the difference between his own theory and that of Atran and Henrich to be a difference in degree. This follows naturally from the assumption that altruism in humans is not special among altruistic species. As Haidt sees it, group selection has been stronger in bees than in humans—and so we are merely 10% bee—but humans and bees have nonetheless both evolved by the same evolutionary process: genetic group selection. As he sees it, what’s at stake in his disagreement with Atran and Henrich is how much genetic group selection may have occurred in recent human evolution. In fact, however, the disagreement is based on a qualitative difference between two kinds of group selection.

Ironically, Haidt’s conceives of gene-culture coevolution in such a way that it has made human evolution more similar to the evolution of bees, rather than less similar—despite the fact that there is no cultural evolution and no gene-culture coevolution in bees. What he apparently envisions is not actually a coevolutionary process at all, but a single process of group selection
that begins as cultural group selection and is gradually transformed into genetic group selection, as gene-culture coevolution progresses. But that would be what we might call “gene-culture group selection,” rather than gene-culture coevolution. If this concept were coherent, it would indeed make human evolution more similar to bee evolution. But it overlooks precisely what is coevolutionary about gene-culture coevolution. For if cultural selection and genetic selection were not two distinct evolutionary processes, then neither one could act as a cause of the other. Haidt mistakes a relation of causation for a relation of constitution, inferring that it is possible for a single process of group selection to be both cultural and genetic selection at the same time, though in varying proportions.

No one thinks that selection for practices of keeping livestock gradually became selection for genes enabling the digestion of lactose in adulthood. Rather, one form of selection for one of these traits occurred first, and it subsequently caused another form of selection for a different trait. For the same reasons, I’ve argued, cultural selection for norms against desertion did not gradually become selection for genes for loyalty or sanctity. Rather, these are two qualitatively different types of trait, distinguished by two qualitatively different types of inheritance, and a prior process of selection for one trait acted as a cause for a subsequent process of selection for the other. What Haidt imagines, instead, is a single process of group selection that begins as “purely” cultural group selection, and then gradually becomes impure cultural group selection, as gene-culture coevolution transforms the process into genetic group selection. Thus, when he finds Atran and Henrich denying that genetic group selection plays any role in religious evolution, he infers that this must be because they disagree about the extent to which this transformation has taken place, given the timeframes involved.
However, when Atran and Henrich argue against genetic group selection, they say nothing at all about the rate of gene-culture coevolution. Instead, they cite Boyd and Richerson (2002), pointing out that there are good reasons for skepticism about genetic group selection that do not also apply to cultural group selection. And what Boyd and Richerson say in the paper they cite is this: “cultural variation is much more susceptible to group selection than genetic variation” (p. 294-95). Atran and Henrich do not deny that cultural group selection may have been a cause of very recent genetic evolution. What they deny is that this kind of genetic selection would be a form of group selection. Without group-level genetic variation for selection to act upon in the first place, the speed of the gene-culture coevolutionary process doesn't matter. Haidt simply misinterprets their reasons for rejecting genetic group selection, which are the same reasons that Richerson and Boyd provide.

2.7 Conclusion

All of this substantially deflates the content of Haidt’s claim that we are 10% bee. If group selection does explain altruism in bees, then humans would appear to have something important in common with bees: group selection would explain why both species display higher levels of altruism than most other species. If this were all it meant for humans to be 10% bee, then the metaphor would be apt. But the purpose of explaining how altruism evolved in humans is not merely to explain why humans are more altruistic than other species, and to explain similarities between the altruism of humans and the altruism of other species is not to explain many of the features of human altruism that we most want to understand.
For example, we learn nothing from bee evolution about the role of moral education in producing altruistic behavior. And, more important for present purposes, we learn nothing from bees about the role that religion plays in producing altruistic behavior. Yet the psychological facts that Haidt originally set out to explain were facts about moral and religious psychology. These were the target phenomena that led him to appeal to group selection in the first place. So, since altruism in bees has nothing to do with morality or religion, similarities between human evolution and bee evolution are simply beside the point.

Fortunately, an account of uniquely human forms of altruism is still available, even if the evolution of altruism in humans is not like the evolution of altruism in bees. The reasons I’ve given for remaining skeptical of genetic group selection in humans do not apply to cultural group selection as well, so the role of religion and morality in producing altruistic behavior may yet be explained by group selection. Among all other species in which high levels of altruism are observed, group-level variation in traits is due to an underlying pattern of group-level variation in genes. That does not appear to be the case for humans. Yet it may still be that humans inherit their religious beliefs and moral practices non-genetically, by imitating the beliefs and practices of their parents, siblings, friends, religious leaders and others. If so, then cultural group selection may explain the role that religious and moral forms of cognition play in causing people to do costly things for the sake of others, even when those others are people they will never meet.

If cultural group selection does explain these distinctly human traits, then it might also affect the subsequent course of genetic evolution. This would not be genetic group selection, however, because even if it were selection for genetically inherited traits, it would not be selection for genes that vary at the group level. A coevolutionary process of this kind would be nothing at all like the process of genetic group selection that Haidt, D. S. Wilson and E. O.
Wilson claim is the explanation for altruism in bees and other eusocial species. As a result, appeals to group selection in humans may stand (or fall) independently of the debate over group selection in biology. The distinction between cultural group selection and genetic group selection thus makes it possible for broad adaptationists in the social sciences to appeal to group selection without thereby committing themselves to more sweeping—and more tenuous—claims about the role of group selection in the rest of the biological world.
Chapter 3

The Goldberg Exaptation Model: Unifying Adaptationist and By-Product Theories of Religion

3.1 Introduction

A central debate in the scientific study of religion is that between adaptationists, who claim that religion is an evolutionary adaptation (Sosis and Alcorta 2003; Sosis 2009; Haidt 2012; Powell and Clarke 2012; Bering and Johnson 2005; Irons 2001; Wilson 2003; Bulbulia 2009), and by-product theorists, who deny that religion is an adaptation (Boyer 2001, 2003; Pyysiäinen and Hauser 2010; Bloom 2007; Barrett 2004; Pinker 2006; McCauley 2004; Lawson and McCauley 1990). According to the by-product account, religion per se has no evolutionary function, and religious belief is explained as a by-product of the activities of cognitive traits adapted for other, non-religious functions. I argue here, however, that this debate is empty; adaptationist and by-product theories are mutually consistent. As a result, the question immediately arises how these theories may be related, and whether they may in fact form parts of the same unified theory. I address these question by identifying a model of the religious phenotype that has already been suggested, independently, by two theorists who stand on opposite sides of the debate. Though the terms of the debate have obscured the commonalities, I argue that shared assumptions captured in what I call the Goldberg Exaptation Model are already sufficient for resolving the debate, and for unifying adaptationist and by-product theories.

The theorists involved take themselves to be in disagreement over the answer to an empirical question: is religion an adaptation? But while adaptationists answer “yes” and by-product theorists answer “no,” these two theories do not actually offer competing explanations
for the same empirical facts. Instead, they offer different explanations for different facts, as a result of the fact that the concepts religion and adaptation have different meanings in the two theories. After first describing the positions in the debate, I will argue that they are, in fact, mutually consistent.

I then proceed to extract the Goldberg Exaptation Model from the existing literature, and use it to show how by-product theories and adaptationist theories may be integrated. The model was first suggested in metaphorical terms by a by-product theorist, philosopher Robert McCauley (2004). He compares the religious phenotype to a Rube Goldberg device—a system composed of distinct components that already possess their own, independent functions. However, a substantially similar model has also been proposed by an adaptationist theorist, anthropologist Richard Sosis (2009). In more literal terms, Sosis discusses the role of exaptation in explaining how the religious phenotype evolved, where exaptation is the process by which traits that already possess one adaptive function come to acquire a new, additional function. After first identifying the commitments that lie at the intersection of these two conceptions of the religious phenotype, I show how they render adaptationist accounts and by-product accounts not merely consistent, but related. For while by-product accounts explain how the component parts of the Goldberg device evolved in the first place, adaptationist accounts explain how these parts subsequently came together to form a functionally integrated system—the Goldberg device itself.
3.2 By-product Theories of Religion

While philosophers such as Hume (1992, 1993), Nietzsche (2007) and James (2004) have offered accounts of religion in naturalistic terms based on a materialist ontology, their theories have not stimulated actual programs of empirical research on religious belief and behavior. And while there is a long tradition of empirical work on religion in social science (cf., Durkheim 1995; Weber 1991; Tylor 1871; Evans-Prichard 1956, 1965; Bellah 2011), these theories do not relate the social, phenomenological and philosophical phenomena they describe to theories of physical facts offered in the non-social sciences. However, in the last two decades researchers from across the disciplines of psychology, anthropology, economics and biology have begun to pursue the study of religion from an evolutionary perspective, and the adoption of this shared framework has led to the emergence of a new set of theories that are not merely consistent with what we already know about the living world in general, but that are capable of using this knowledge to develop richer and deeper explanations of religion. Yet conceptual differences within this new field hinder progress toward this goal.

The first of these two decades was dominated by an approach now commonly referred to as the cognitive science of religion, or CSR, and virtually all theories in this tradition are by-product theories. Adaptationists joined the discussion a decade later, but by the time of the first International Conference on the Evolution of Religion, in 2007, Richard Sosis (2009) reports that “One of the most heated topics of discussion concerned whether or not religion should be considered an adaptation or a byproduct” (p. 315). He mentions this in an article devoted entirely to defending the adaptationist theory against such an array of criticisms from by-product theorists that he is compelled to limit his responses to only five of the most common objections. Sosis takes as an epigraph for this article a quote from a member of the opposing camp,

4 An exception is the account offered by Jesse Bering (2006), discussed later on.
psychologist Lee Kirkpatrick, who at least agrees that “The first question to be addressed by any evolutionary approach to religion is whether religion is an adaptation or a byproduct of adaptations designed for other purposes” (p. 315). Whether this question really is the primary question in the field is arguable, but the number of researchers who have explicitly aligned themselves with one side or the other in this debate is evidence that it is at least a primary question.

The CSR tradition focuses specifically on the cognitive traits involved in forming beliefs about supernatural agents. It also focuses specifically on traits that are more or less universal among humans (cf. Boyer 2001; Atran 2002), and these traits are explained by appeal to genetic selection, in accord with the theory of inclusive fitness. It is thus important to note that the debate in question concerns whether religious traits are adaptations for religion, in particular, since everyone involved in the debate is an adaptationist in the more general sense that they explain their target traits as products of natural selection, rather than as products of other evolutionary forces. Further, CSR theorists are keen to point out that the universal adaptations to which they appeal are capable of explaining not only the universality of religious belief, but also the variation in specific beliefs observed across cultures. This variation is said to occur because the same cognitive systems produce different specific beliefs given different environmental inputs.

One of the most influential theories of this kind holds that people believe in supernatural agents because their capacities for detecting natural agents in the environment are hyperactive (Guthrie 1993). Psychologist Justin Barrett (2004) has dubbed this trait the hyperactive agency detection device, or HADD, and it is the hyperactivity itself that is said to be a genetic adaptation. This is because it is thought to be adaptive for agency detection systems to be biased
toward false positives over false negatives, meaning that it’s better to jump to the false conclusion that there is an agent in one’s vicinity than it is for there to actually be an agent in the vicinity that one hasn’t detected. As a result, it is claimed, we are genetically predisposed to believe in the existence of agents on the basis of scant evidence—a rustle in the bushes might suffice. Thus, the fact that supernatural agent beliefs show up in nearly all societies is explained by the fact that individuals in all these societies share the same history of genetic selection for the HADD. That these beliefs vary in their specific contents from one society to the next is explained by the fact that different environmental contexts may trigger all sorts of different false attributions of agency.

Of course, the HADD by itself cannot explain everything about the formation of religious beliefs. For one thing, religious beliefs are not simply the idiosyncratic beliefs of individuals, but are shared widely across individuals. Hyperactive agency detection does not explain this, but another influential CSR theory does. Anthropologist Pascal Boyer (2001, 2003) explains why some supernatural beliefs spread through social groups better than others do by appealing to innate and universal patterns of memory and inference favoring the formation of minimally counterintuitive (MCI) beliefs. Boyer notes that religious representations typically contradict our ordinary intuitions, and he argues that this makes them stand out as particularly interesting and memorable. Representations of a talking bush, for example, are more salient and compelling than representations of an ordinary bush, and this, Boyer suggests, explains why individuals are more prone to think about, remember, and pass on to others representations of talking bushes. On the other hand, if counterintuitive ideas are too counterintuitive, then they will simply be incoherent, difficult to process, and difficult to remember. Hence, the spread of successful religious ideas is explained by their being counterintuitive, but only minimally so.
The adaptive value of MCI beliefs is said to be due to the information-processing efficiency of categorization and conceptualization in general. If an object can be categorized as an apple, for example, then a wide range of particular inferences and predictions can be made on that basis, such as that it will taste sweet, that it must be eaten before it rots, and that other people may value it. In the same way, it is claimed, if an object is categorized as an agent possessing a mind, then a wide range of other inferences and predictions will follow from this classification, such as the prediction that it can perceive what happens in its vicinity and be pleased or displeased according to what it perceives. It is the cognitive efficiency of the capacity for categorization in general that is said to be adaptive on this view, but there is nothing particularly religious about that. The adaptive benefits of supernatural-agent concepts are no different from those provided by concepts in general, so supernatural agent concepts are simply a by-product of cognitive processes adapted for purposes much more general than religious belief itself.

The HADD theory and the MCI theory explain different cognitive traits using the same general form of explanation, according to which what is adaptive about the traits in question has nothing to do with what is religious about them. That general form of explanation is what I refer to here as the “by-product theory” or the “by-product approach.” But these two traits are not the only features of the religious phenotype for which by-product accounts have been offered. Lawson and McCauley (1990) explain the cognition involved in participating in religious rituals by appealing to universal mechanisms for representing intentional action in general. On their view, ritual involves an attempt to intervene upon the world by interacting with supernatural agents, but beliefs about supernatural agents are just by-products of cognitive systems the function of which is to negotiate interactions with ordinary, natural agents.
Bloom (2007) suggests yet another such trait, arguing that universal intuitions of mind-body dualism are “a natural by-product of the fact that we have two distinct cognitive systems, one for dealing with material objects, the other for social entities. These systems have incommensurable outputs. Hence dualism emerges as an evolutionary accident” (p. 159). If folk dualism is the result of a universally shared history of genetic selection for two distinct cognitive systems, then it is no surprise that we find dualistic notions of bodiless agents appearing across a wide variety of religious systems. But Bloom’s point is not that these distinct systems evolved for the purpose of forming religious beliefs. Rather, it is that these beliefs occur as a side effect of the selection pressures that explain why dualistic beliefs are possible in the first place.

By-product theorists do not assume that the by-product approach offers a complete account of religious belief all by itself. For example, while McCauley (2004) enthusiastically endorses both the HADD and MCI theories, he also notes that the joint activity of these traits would also explain beliefs about Santa Claus, leprechauns and fairies. “So,” he concludes, “this is only part of the story, but it is a very important part” (p. 49). Bloom notes more generally that “It is obvious that some religious beliefs are entirely learned—nobody is born with the idea that the birthplace of humanity was the Garden of Eden, or that the soul enters the body at the moment of conception, or that martyrs will be rewarded with sexual access to scores of virgins” (2007, p. 150). Accordingly, it is more fair to characterize defenders of the by-product approach as being committed to pursuing explanation of this general form for many religious traits, including many that vary across cultures.
3.3 Adaptationist Theories of Religion

3.3.1 The General Form of the Adaptationist Approach

On the other side of the debate, we do not find adaptationists denying that religious beliefs are minimally counterintuitive, or that it would be adaptive for agency-detection systems to be hyperactive. Indeed, adaptationists have had little to say about the cognitive systems and processes that serve as proximate causes for religious belief-formation. They have focused instead on explaining other religious traits, especially traits of social motivation. And while there are several different adaptationist theories on offer, which explain several different motivational traits, there is striking and consistent agreement among them all regarding the general source of adaptive value for those traits. Adaptationist theories all agree that what is adaptive about religion is the role it plays in motivating cooperation, especially when cooperation is costly.

As adaptationists have noted (Atran 2002; Henrich and Gervais 2010; Norenzayan 2013, p. 161), what a purely cognitivist perspective cannot explain is the role of commitment in religious belief. Anyone may merely entertain, hypothetically, the propositions and inferences of anyone else’s religious system, and this shows that one may engage in the cognition of a given religion without actually believing it. For someone to be a religious believer, the relevant forms of cognition must carry motivational force. Genuine belief actually influences or guides one’s actions and decisions, and adaptationists focus specifically on beliefs that influence or guide actions that are in some sense costly to the believer. Thus, by contrast with cognitivist focus of by-product theories, the role of motivation is central to adaptationist theories.

Adaptationists about religion all agree that what makes religion adaptive is its ability to promote cooperative behavior, but they also recognize that our reasons and motivations for cooperating often come into conflict with reasons and motivations for not cooperating, and for
pursuing self-interest instead. Accordingly, they emphasize the ability of religious commitment to motivate cooperation even in cases where the goals and interests of the agent diverge from those of her community. The Christian emphasis on striving against the temptation of sin provides a perfect example of the sense in which religious systems focus upon conflicts between self-interest and “higher purposes,” which typically involve concern for others.

On the adaptationist view, religious beliefs are not needed to motivate other forms of cooperation, such as ordinary economic exchanges. Since these exchanges are based on mutual self-interest, questions of selfish temptation do not arise, and no costly motivations need to be explained. Thus, while adaptationist theories do not deny that religion may also facilitate self-interested cooperation in various ways, they place special emphasis on the adaptive value of prosocial or altruistic cooperation. For example, Norenzayan and Shariff (2008) take as their target explanandum not religion in general, but “religious prosociality, or the idea that religions facilitate acts that benefit others at a personal cost” (p. 58). They begin by noting that “texts of all major religions explicitly encourage prosociality in their adherents” (p. 58).

Obviously, not all religious systems posit an omnipotent God who is concerned with human moral behavior, but the God of Christianity, Islam and Judaism has more true followers than any other, and these systems conform beautifully to the general form of the adaptationist account: selection for cooperation leads to selection for beliefs and practices that promote prosociality. The same adaptive logic also applies, however, to most other belief systems that are intuitively classified as religious systems. For example, cooperation could also be increased by beliefs about ancestor spirits who are far from omniscient, but who still know enough about norm violations in the village to dole out rewards and punishments in the form of “fortune” and

---

5 The focus here is on evolutionary altruism, rather than psychological altruism. That is, in adaptationist theories of religion, the costs and benefits of altruism are measured in terms of the consequences of behavior for selection, rather than in terms of the mental states that cause the behavior.
“misfortune.” The Hindu and Buddhist principle of karma even posits supernatural rewards and punishments without positing any supernatural agent at all (even if Hindus and Buddhists also possess other beliefs involving supernatural agents). In this case, a supernatural force is sufficient, as long as beliefs about that force function to motivate compassion for others.

So while by-product theories take as their target explananda the cognitive traits that act as proximate causes of religious belief, adaptationist theories focus instead on the motivational consequences of such beliefs. Put differently, while by-product theories explain how agents represent the cognitive contents of religious propositions, adaptationist theories explain why agents adopt the attitude of belief toward those propositions. It is one thing to entertain a myth from a foreign culture about an all-powerful being who knows what people do and think, and then doles out punishments and rewards accordingly. It is quite another thing to actually believe that such a being is monitoring one’s own thoughts and actions. Only if someone adopts the attitude of belief toward such propositions will she actually be motivated to avoid those punishments and pursue those rewards. And it is only because she is motivated in these ways, adaptationist theories claim, that her supernatural beliefs are adaptive.

3.3.2 Where Adaptationist Theories Differ

Like the by-product theory, then, the adaptationist theory is in fact a general form of explanation shared by a family of more specific theories, each of which explains different traits of the religious phenotype. Unlike by-product theories, however, when adaptationist theories go to specify what the relevant traits are for, they do not all appeal to the same kind of selection process. And while particular adaptationist theories all take as their target explananda traits that
somehow contribute to prosocial cooperation, they identify different sorts of traits depending on the kinds of selection to which they appeal.

In addition to his theoretical work defending the adaptationist approach in general, Sosis also advances a particular adaptationist account of the trait of *costly signaling* in religious rituals (Alcorta and Sosis 2005). This theory, originally put forward by Irons (2001), begins by noting that collective enterprises generate a temptation for individuals to “free ride” on the efforts of others. The more everyone else invests in a collective project, the greater the temptation becomes for each individual to selfishly leave the costs of cooperating to others. “If three hundred other guys are already going out to fight,” the free rider reasons, “what difference will it make if I just sit this one out?” This temptation poses a problem for adaptationist theories of cooperation, because it demonstrates a sense in which cooperation reduces fitness, rather than enhancing it. The theory of costly signaling then addresses this problem by suggesting a mechanism by which certain forms of religious ritual help genuine cooperators to identify each other as partners, so that they may exclude free riders from their cooperative activities and avoid exploitation.

A dramatic example comes from Hindu devotees in Malaysia, who carry “burdens” in an annual procession celebrating a mythical victory of good over evil. These burdens often involve piercing the body with thick metal pins, and some participants even suspend their whole bodies in mid-air by hooks pierced through the skin of their backs and legs. Anyone willing to do *that*, it seems, is capable of putting religious commitments above self interest. If those commitments happen to reflect normative principles discouraging the exploitation of others, then publicly displaying burdens in such a conspicuous manner may function as a way of credibly advertising to others that one is not a free rider. Genuine cooperators thus have a mechanism for identifying
each other as partners, so that they may reap the benefits of cooperation while avoiding the costs of exploitation.

Sosis emphasizes the benefits of this cooperation for the individuals who actually commit the costly acts, and thus his account is based on individual-level genetic selection, just as by-product theories are. While the signaling behavior is costly, he argues, these costs are offset by the benefits of avoiding free riders, which accrue to the signaler herself. Sosis also acknowledges, however, the possibility that selection for costly signaling may occur at the group level, rather than at the individual level (Sosis and Alcorta 2003). While free riding reduces the fitness of individuals, it also reduces the fitness of groups. The greater the proportion of free riders in an army, for example, the smaller the number of soldiers who actually show up on the battlefield, and the poorer the group’s chances of winning. The same logic also applies, however, for less direct forms of group-level competition. Many forms of foraging, farming and hunting require a great deal of cooperation, and in a group full of free riders there would be little food to go around.

Since free riding is a special form of selfishness, this group-selectionist account of costly signaling is just a special case of the more general principle on which all group-selectionist accounts of religion are founded. These theories begin by emphasizing the role of religion in motivating altruistic behavior, and then argue that the costs of altruism to the individual may be more than offset by benefits that accrue to the group as a whole. As a result, even when altruism reduces the fitness of the altruist relative to other individuals, it may nevertheless also increase the fitness of the altruist’s group relative to other groups. Consequently, altruistic behavior can be selected for in spite of its costs, and if religion promotes altruism, then it, too, may be selected for at the group level.
The group-selectionist theory of religion was originally put forward by biologist David Sloan Wilson (2003), but it has recently been developed further by psychologist Jonathan Haidt (2012), who integrates it into his own theory of moral psychology. And in addition to appealing to group selection, both theorists also rely heavily on the theory of cultural evolution as well. One reason for this is already hinted at in Bloom’s admission, above, that genetic inheritance does not explain how individuals acquire the particular beliefs and practices of their particular faith. Even when children inherit the religious beliefs of their own parents, they inherit them culturally, rather than genetically; children learn the beliefs and practices of their religion in the process of acculturation, just as they learn to say “please” and “thank you.” The fundamental insight of cultural evolutionary theory is that this kind of learning may be treated as a form of non-genetic inheritance, thus making it possible explain such traits in terms of a Darwinian selection process.

Another important reason for appealing to cultural evolution, however, is that the traits of particular religions spread through populations on timescales far too brief to be explained by genetic selection. Boyer’s appeals to the minimally counterintuitive nature of religious beliefs may constrain, to some degree, the forms of belief that are likely candidates for transmission from one person to another. But it does not explain why certain MCI beliefs have spread across the globe and persisted for thousands of years, while most others have not. Among all the possible MCI beliefs that may have captured people’s imaginations over the last two thousand years, what is so special, for example, about the belief that Jesus was resurrected three days after he was crucified? According to cultural evolutionary theory, the exceptional prevalence of this counterintuitive belief can be explained by other properties of it that cause people to adopt it from others.
To appeal to both cultural evolution and genetic evolution in the same group-selectionist theory, however, is to introduce a critical question: is group selection supposed to explain genetic evolution, or cultural evolution? That is, how are the group-selected traits transmitted—by learning, or by genes? To recognize two different forms of inheritance is to recognize two very different types of selection process, and thus two very different forms of adaptationist explanation. Wilson suggests that both forms of group selection may play some role, but he does not attempt to identify particular traits that are products of one form of group selection or the other. Haidt, however, appeals specifically to genetic group selection to explain certain traits, and this brings his theory into conflict with an alternative group-selectionist theory defended by anthropologists Atran and Henrich (2010) and psychologists Norenzayan and Shariff (2008; Norenzayan 2013), who argue that religious prosociality is a product of cultural group selection, not genetic group selection.

Thus, while we find by-product theories appealing only to genetic selection at the individual level, we find adaptationist theories appealing, in addition, to cultural evolution and group selection. As I argue in the next section, this difference leads to very different commitments regarding what counts as an adaptation, and this is one important source of the misunderstanding that has been mistaken for an empirical debate.

### 3.4 Two Different Meanings of ‘Adaptation’

Since traits inherited in either of two ways (genetic vs. cultural) may be selected for at either of two levels (group vs. individual), the theoretical framework employed by adaptationists about religion recognizes no less than four distinct types of selection that may be invoked to explain a
given trait. By contrast, the framework on which by-product theories are based recognizes only one of these four types: genetic selection at the individual level. I thus distinguish between broad and narrow forms of selectionist explanation, and this methodological distinction results in an important difference between the concepts of adaptation employed in adaptationist and by-product theories, respectively. (See Figure 3.1.)

**Figure 3.1 Narrow vs. Broad Forms of Adaptationist Framework**

Since broad adaptationists recognize four distinct forms of selection process, there is a big difference in these accounts between a trait’s being a product of selection *in general* and it’s being a product of *individual-level genetic* selection, in particular. On this view, any trait that is a product of selection will count as an adaptation, regardless of what kind of selection is involved. So, for example, for a broad adaptationist it would not follow from the fact that a trait is an adaptation that there has been any change in gene frequencies in the relevant population. For narrow adaptationists, by contrast, *all* selection is genetic selection, and to abstract away from
facts about gene frequencies would be to abstract away from facts essential for determining whether a trait is an adaptation or not.

If there is a genuine debate underlying the apparent debate between adaptationist and by-product theories of religion, I suspect it lies here, between those who are willing to recognize non-genetic traits as products of selection and those who insist that selectionist explanations are legitimate only for genetically inherited traits. But if this is indeed a genuine point of debate between broad and narrow adaptationists, it is important to see that it is not an empirical debate about how any particular trait evolved, and it is not a debate about religion. Rather, it is a methodological debate about how to explain the evolution of psychological traits in general.\(^6\)

In one sense, then, the broad adaptationist concept of adaptation is a dramatic departure from the ordinary use of ‘adaptation’ in evolutionary theorizing. The concept of adaptation is widely employed today because of its fantastic success in biology in general, and that success is due to theories that do depend on empirical facts about changes in gene frequencies for traits possessed by individuals. From the point of view of orthodox biological theory, the idea of a “cultural adaptation” or a “group-level adaptation” sounds outlandish at best, and incoherent at worst. Yet while all human activities are biological phenomena, some of our activities—including religious ones—are biological phenomena that don’t occur in any other species. As a result, adaptationists about religion are attempting to apply biological concepts in a new and

\(^6\) As a result, it should be noted that while these questions are more general than those concerning religion, they are nevertheless much more specific than another set questions about adaptationism that have been much discussed in the literature. Godfrey-Smith (2001) describes three different forms of adaptationism, and Lewens (2008) develops and expands this taxonomy to include no less than seven distinct forms of adaptationism. The questions addressed by this taxonomy, however, concern the relative importance of adaptationist vs. non-adaptationist explanations in biology in general, rather than the relative importance of two distinct forms of adaptationist explanation within the specific domain of human behavior and cognition.
different domain, and this appears to have led them to stray from the orthodoxy of biological terminology.

Nevertheless, except for the fact that such terminological shifts engender miscommunication, there is nothing unreasonable about the broad adaptationist definition of ‘adaptation.’ This is because it is reasonable to suggest that a trait counts as an adaptation as long as there is something it has been adapted for—as long as it possesses a function. On this view, traits’ functions are determined by selection, so any trait that is a product of any kind of selection can be said to have a function. Whether it is a trait of groups or of individuals, and whether it is inherited culturally or genetically, a trait counts as an adaptation as long as some form of selection explains its properties. And if it has been adapted by selection, broad adaptationists assume, then it counts as an adaptation. All that matters is the logic of selection in general, and its role in making teleological attributions.

Adaptationists about religion are sometimes quite explicit about this. David Sloan Wilson (2003) borrows the term “Darwin machine” from Plotkin (1994), and uses it as a way of referring to the general logic of selection processes. He then advocates an ultra-liberal application of this logic in the study of religion, arguing that not only cultural evolution but even ordinary rational decision-making (in which an agent selects among options) are examples of Darwin machines. He complains that “It is unfortunate that evolution is so often associated with genetic evolution, a slow process that gives the impression of an incapacity for change over the timescales that matter most to living people struggling with their problems. When we expand our view of evolution to include all Darwinian processes, we can begin to see how religions actually produce transformative change, even from a purely evolutionary perspective” (p. 35).
Philosophers Russell Powell and Steve Clarke, who have recently argued against the by-product theory, also insist that “the process of adaptation need not be tied to genetic transmission” (Powell and Clarke 2012, p. 16). Indeed, they offer a concise statement of the central assumptions of what I am calling broad adaptationism, appealing directly to the “cultural group” as a fitness-bearing unit of selection:

There are many clear-cut examples of non-genetic adaptations that increased the fitness of cultural groups, including the manufacture of fire (Wrangham 2009) and the construction of sea-faring vessels (Richerson and Boyd 2005), as well as the development of moral norms, military hierarchies, and sophisticated social exchange networks that were built gradually by ‘invisible hand’ mechanisms that are closely analogous to paradigmatic natural selection (Sterelny 2007). Religion may be a less functionally obvious case than these, but there is nothing inherently problematic about it being an adaptation of cultural groups. (p. 20)

Like Wilson, Powell and Clarke explicitly insist that it makes no difference whether the particular selection process involved is cultural evolution or genetic evolution, or whether it occurs at the level of groups or of individuals.

These are conceptual claims about what counts as an adaptation, not empirical claims about how any particular trait actually evolved. And while I agree that there is nothing “inherently problematic” about treating the traits of cultural groups as adaptations, it is nevertheless clear that this is not how the concept of adaptation is used in by-product theories. Because adaptationist theories of religion employ a broad adaptationist concept of adaptation, when they
assert that religion is an adaptation, what they are actually committed to is the assertion that religion is a product of *some kind of selection or other*. By contrast, when by-product theories assert that religion is *not* an adaptation, what they are actually committed to is the claim that religion is not a product of *individual-level genetic selection, in particular*. But this means that the by-product theory does not deny what the adaptationist theory asserts, because to deny that a trait is a product of a *particular* type of selection is obviously not to deny that it is a product of *some* type of selection. To insist that a trait is not a product of individual-level genetic selection is not to insist that it is not a product of some other kind of selection, such as cultural selection or group selection.

Once the empirical commitments of these theories are separated from their more general methodological commitments, it becomes clear that the by-product theory and the adaptationist theory are mutually consistent. Since they do not agree about which empirical facts are entailed by the claim that a certain trait is an adaptation, the accounts they offer are not competing accounts of the same empirical phenomena. They are simply different accounts of different facts.

### 3.5 Two Different Meanings of ‘Religion’

The differing concepts of adaptation I’ve been discussing provide one way of showing that adaptationist and by-product theories are mutually consistent, but there is also another reason why they do not conflict. Just as the concept of adaptation has different meanings in these two theories, so too does the concept of religion. When by-product theories explain “religion,” what they actually explain is something much more specific: cognitive traits that are inherited genetically and selected for at the individual level, and that are responsible for forming
representations involving supernatural agents. Likewise, when adaptationist theories explain how “religion” evolved, they too explain something rather more specific: motivational traits that may be inherited either by genes or by learning, and that may be selected for at either the group or individual level. But since these theories explain these different types of traits, they do no offer competing explanations for the same empirical facts.

This is not to suggest that adaptationist theories have nothing at all to say about cognitive traits, of course, since the contents of religious beliefs in these accounts remain essential to explaining cooperative behavior. The point is, rather, that cognitive traits are secondary in an important sense: they are posited only in virtue of their instrumental contributions to prosocial motivation. Beliefs about God’s moral attitudes may be one way to trigger the relevant motivations, but entirely different forms of cognition, such as atheistic beliefs about karma, may do the job as well. As long as a given cognitive trait generates the right kind of motivational responses, adaptationist theories have everything they need. As a result, while by-product theories and adaptationist theories both offer explanations for various cognitive traits, they do not offer explanations for the same kinds of cognitive traits. By-product theories identify the functions of cognitive systems and processes that act as the proximate, immediate causes of religious belief. By contrast, adaptationist theories explain the motivational consequences of religious belief. This opens up the possibility that the same belief might have both the causes attributed to it by the by-product theory and the consequences attributed to it by the adaptationist theory.

Suppose, for example, that the HADD causes someone to believe in the stories she’s heard about an invisible ancestor spirit who roams the village, and who will punish bad behavior if she sees it. Then suppose the person comes across an opportunity to steal something, and chooses not
to as a result of the belief that her theft might be detected by the ancestor spirit. In such a case, the potential thief pays an opportunity cost, passing up the gains to be made from stealing. At the same time, the potential victim benefits from the decision not to steal, making this exactly the sort of prosocial behavior that, according to adaptationists, makes religion adaptive. Thus, even if the HADD did not originally evolve to perform a specifically religious function, once it has already evolved, it might just happen to have motivational consequences that contribute positively to fitness, and it might thus begin to be selected for. In other words, a trait that originally evolved to perform some non-religious function might subsequently acquire a specifically religious function.

Here we see how systematic differences in the methodological commitments of adaptationist and by-product theories produce systematic differences in the types of facts that each purports to explain. Because the religious phenotype is a complex natural phenomenon, researchers must first analyze it into more manageable constituent traits, such as the HADD, or costly signaling behaviors. But instead of leading adaptationists and by-product theorists to pursue the same sorts of traits, these methodological differences lead them to offer accounts for different sorts of traits. So just as differences in the meaning of ‘adaptation’ render the theories logically consistent, so too do differences in the meaning of ‘religion,’ the natural phenomenon taken to be the target explanandum.

This concludes my argument for the consistency of the two theories, but there is more to be said about their relationship than merely that they are consistent. If they explain different constituent features of the same complex phenomenon, then how are these theories related? In what follows, I show how the Goldberg Exaptation Model addresses this question.
3.6 The Goldberg Exaptation Model

The Goldberg Exaptation Model (GEM) is a conception of the religious phenotype that I have extracted from the existing literature, because it has been proposed independently by theorists from both sides of the debate. It is not a synthesis or hybridization of two opposing views, but an account of the intersection that already exists between them, a set of shared assumptions regarding the target explananda for evolutionary theories of religion. These assumptions, I argue, are already sufficient to show how adaptationist and by-product theories may be integrated.

While McCauley was first to propose the GEM, he did so in metaphorical terms that can be more clearly understood against the background of a literal description. I thus begin with Sosis’s version. Both theorists present the model in the course of defending positions they take to be opposed, so their rhetorical aims are not just different but opposite. Nevertheless, the substance of their descriptive claims shows that there is nothing more at stake than a matter of emphasis. After defining the GEM itself in this section, I show in the following section how it may be applied.

3.6.1 Sosis’s Mereological Analysis

Sosis presents the central features of the GEM while accusing by-product theories of committing a mereological fallacy—explaining the parts when they should be explaining the whole. I disagree that Sosis’s mereological analysis shows the by-product theory to be in any kind of error, but the descriptive content of the analysis forms the core of the GEM.

Sosis notes that the religious phenotype is a complex natural phenomenon, an “adaptive complex” analyzable into many distinct constituent traits. He also observes, correctly, that what by-product theories explain are the functions of certain constituent traits, rather than the function
of the complex as a whole. He refers to the complex as “the religious system,” arguing that “It is the religious system, not the constituent parts, that produces functional effects and is the appropriate unit of an adaptationist analysis. A proper byproduct account of religion, which has yet to be offered, must explain why the religious system’s constituent parts recurrently coalesce across cultures” (2009, p. 323).

We must distinguish here between two different claims. First, there is the descriptive claim that the religious phenotype forms a complex functional whole composed of constituent functional parts, such that it makes sense to ask separate questions about the function of the whole system, on one hand, and about the functions of its constituent parts, on the other. Second, there is the normative claim that it is the whole, not the parts, that constitutes the proper or appropriate explanandum for evolutionary theories of religion.

In assessing the normative claim, it is important to understand Sosis’s rhetorical context. In the article in which this quote appears, his fundamental aims are defensive, not offensive. The purpose of the article is to defend the adaptationist approach against five common criticisms from by-product theorists. But, like his antagonists, Sosis assumes from the outset that the two theories offer competing accounts of the same target phenomenon. Consequently, his defensive strategy is to go on the offensive: he identifies what he takes to be the target explanandum for both theories, and argues that the by-product theory fails to explain that.

But given that this is his argumentative strategy, it is important to note the structure of the objection. The claim is that a proper by-product theory has yet to be offered, which means that no actual by-product account has even tried to explain the function of the religious system as a whole. If so, then the complaint cannot be that the by-product theory provides a bad explanation
for something it was never used to explain in the first place. Rather, it must be that the theory fails to explain what it should be explaining, or at least trying to explain.

But why should by-product theorists focus on explaining the whole, and why should they stop trying to explain the parts? If adaptationists are studying the whole, shouldn’t someone else be studying the parts? Sosis provides no argument against the value of explaining the constituent parts of the religious system, and it is hard to see why anyone would want to do so. Surely a complete understanding of the religious phenotype will ultimately require knowledge about both the specific functions of the parts and the general function of the whole, just as facts about the functioning of individual neurons are one important aspect of our knowledge of the functioning of the nervous system as a whole. If so, then there is nothing improper about theories that focus exclusively on constituent parts of system. To do so is to make a useful contribution to a larger theoretical enterprise.

If my arguments for the consistency of these theories succeed, then Sosis’s normative claim seems to be motivated by a false assumption, namely, that adaptationist and by-product theories offer competing accounts of the same target facts. In fact, Sosis’s own mereological analysis introduces yet a third sense in which the theories can be said to explain different things. On his account, the success of the adaptationist theory will depend on its ability to explain how and why the various parts have been assembled into a coherent functional whole, and it may do this whether or not the by-product theory succeeds in explaining the distinct functions of the parts themselves.

Fortunately, Sosis doesn’t actually need to defend his normative claim, because his mereological analysis is already sufficient to secure his defensive aims. These aims may be met simply by showing that adaptationist theories explain important facts that are not explained by
by-product theories. Sosis meets them by identifying a distinct domain of facts, concerning the functioning of the system as a whole, which only adaptationist accounts explain. This is entirely consistent with its also being the case that by-product theories explain other facts that only by-product theories explain.

### 3.6.2 McCauley’s Metaphor and the Role of Exaptation

McCauley describes the religious phenotype in very different terms, but the metaphor he uses aptly characterizes Sosis’s mereological analysis. According to McCauley (2004),

> The mind does not contain a specific department of religion. Instead, religion exploits a diverse collection of emotional and cognitive inclinations in human beings that enjoy neither logical nor psychological unity. The upshot of this analysis is that cognitively speaking religion is a Rube Goldberg device, which is to say that it is an exceedingly complicated contraption calling on all sorts of psychological propensities that are, otherwise, usually unlinked. (p. 48, emphasis in the original)

McCauley’s rhetorical tone here places heavy emphasis on the task of explaining the independent, non-religious functions of the constituent parts of the Goldberg device of religion, rather than on explaining their functional interactions as parts within a larger system. Nevertheless, there is nothing more than a difference in emphasis between McCauley’s view of the religious phenotype and Sosis’s. McCauley does not deny that the functional parts are linked
together as parts of a larger system, even if he does emphasize the fact that they are “otherwise usually unlinked.”

The metaphor of the Goldberg device maps beautifully onto Sosis’s mereological analysis. For while a Goldberg device may be cobbled together out of a number of distinct parts that already possess their own, independent functions, its charm lies in the fact that there is nevertheless something that these independent parts have all been cobbled together for. A Goldberg device possesses its own function, which is distinct from the functions possessed by its various parts. Indeed, the Goldberg metaphor is as amenable to Sosis’s rhetorical aims as it is to McCauley’s. Armed with this image, Sosis could say that by-product theories do not identify the function of the device itself, even if they do explain the functions of its parts. (See Figure 3.2)

Figure 3.2 The Goldberg Exaptation Model
By likening the religious phenotype to a Goldberg device, rather than to an ordinary device such as a car or a coffee maker, McCauley emphasizes the fact that the constituent traits already possess their own, independent functions, prior to their being assembled into a larger device. While this would indeed show that the constituent traits possess independent, non-religious functions, and thus that they are evolutionary by-products, it would also show that they are capable of acquiring new, additional functional roles as parts in a larger system. McCauley emphasizes the original functions over the new, additional functions they acquire, but it is nevertheless his own suggestion that there is a larger system—the Goldberg device itself. So while he emphasizes the importance of identifying non-religious functions of various constituent traits, he does not deny that these traits also play roles within a larger functional system. And that system constitutes his description of the religious phenotype.

Likewise, while Sosis emphasizes the importance of explaining the whole over the parts, he also acknowledges that the parts possess prior, independent functions. Indeed, he states in literal terms the same point that McCauley makes in metaphorical terms by comparing the religious phenotype to a Goldberg device instead of an ordinary device like a car or a coffee maker. If the religious system is like a Goldberg device, then understanding how the device performs its function will ultimately require us to understand how the parts perform their prior, independent functions. Sosis makes this point by appealing to the concept of *exaptation*:  

---

7 Powell and Clarke (2012) endorse this view as well, and they come closer than Sosis to seeing it as a reason for dissolving the debate, if not as a method for actually integrating the theories. They call their account a “pluralist” view, and they note that some traits may be by-products while others are adaptations. They even note that, “The evolution of religion debate is often cast as a forced choice between byproduct and selectionist theories of religion, but this is a false dichotomy” (p. 22). Nevertheless, instead of concluding that the debate is empty, they frame their argument as an objection to the by-product approach, and their pluralist view is ultimately indistinguishable from Sosis’s view. They conclude that “If religion is composed of a suite of co-opted byproducts that were gradually organized and improved over time through cumulative selection, then it looks less like an exaptation and more like an ordinary adaptation” (p. 23).
The most likely evolutionary scenario is that cognitive, emotional, and behavioral elements were exapted for use in a complex system of communication, cooperation, and coordination, namely the religious system. An exaptation is a preexisting trait that acquires a new role for which it was not originally designed by natural selection (Gould and Vrba 1982). Importantly, exaptations have functional effects but exapted traits are not modified when taking on their new role; if they are, adaptive modifications are known as secondary adaptations.

By explicitly describing constituent traits in the system as exaptations rather than secondary adaptations, Sosis grants that the original, non-religious functions of the parts are also part of the explanatory story.

So while Sosis and McCauley differ with regard to the importance they invest in explaining various aspects of the religious phenotype, both agree with the following three claims, which constitute what I am calling the Goldberg Exaptation Model:

(1) The religious phenotype is a complex phenomenon composed of more specific constituent traits.

(2) These constituent traits originally evolved to perform non-religious functions.

(3) After they had already evolved, the constituent traits were exapted by selection to play new, additional roles within a larger system that possesses its own function.

The crucial point, then, is that while (2) gives by-product theorists what they insist upon, (3) gives adaptationists what they insist upon. If important constituent traits of the religious complex
originally evolved to perform non-religious functions, then those traits are by-products, at least in the sense that by-product theorists insist upon. But this does not imply that they are only by-products. For if selection also explains why, after they initially evolved, the same traits were assembled into a larger system, then the system itself is an adaptation too, at least in the sense that adaptationists insist upon. And if, as seems clear, it is the system as a whole, rather than any particular constituent trait, that is responsible for the common-sense distinction between individuals who are religious and individuals who are not, then adaptationists seem to be within their rights in declaring the system an adaptation for religion.

3.7 Unifying By-product and Adaptationist Accounts

Having thus explained what I take the GEM to be, I now want to illustrate, by example, how it can be used to integrate by-product accounts of certain facts with adaptationist accounts of other facts. I approach this by considering a particular problem case posed by the trait of folk dualism. We’ve seen already that Bloom (2007) offers a by-product account of this trait, according to which representations of disembodied minds result from the “incommensurable outputs” of two distinct cognitive systems: one that evolved for representing the movements of ordinary physical objects, and another that evolved for representing the movements of agents with minds. However, an adaptationist account of folk dualism has also been offered by psychologist Jesse Bering (2006), making this the one trait for which both a by-product account and an adaptationist account have been offered. As a result, these two accounts come closest to manifesting a genuine empirical debate, and thus a counterexample to my claim that the theories explain different target phenomena.
Considering this case in light of the GEM, however, reveals that they do not, in fact, explain the same target facts. For while they do explain facts about the same trait, they do not explain the same facts about that trait. More specifically, what they explain are different stages in the evolution of the same trait—facts that could be separated by tens or even hundreds of millions of years. The GEM thus shows how these two theories make distinct contributions to the same causal narrative.

On Bering’s account, the adaptive value of folk dualism is due to the role it plays in helping people to manage their reputations in the community. If the bodiless minds being represented happen to belong to supernatural agents who punish norm violations and reward concern for others, then, Bering suggests, dualistic cognition may cause people to behave in ways that will contribute positively to their reputations. This then benefits the genetic fitness of the individual believer, because humans’ social relationships are very important for their genetic fitness, and their reputations are very important for maintaining good social relationships. On this view, what makes dualism an adaptation for religion, more specifically, is that the adaptive value of dualistic beliefs comes only in virtue of their outputs to social motivation. Representations of disembodied minds only contribute positively to genetic fitness if they succeed in motivating people to obey norms and inhibit selfish impulses.

According to Bloom, there are good adaptive reasons for representing agents differently from other physical objects, and this explains why we have two distinct cognitive systems capable of producing incommensurable outputs. I take this to be a plausible empirical claim about the ecological conditions of our ancestors, but it is important to note that these selection pressures would already have been in place a very long time ago—even by evolutionary standards. The selection pressures that explain why we represent agents as being different from
other objects in the environment would not be recent additions to the ecology of our lineage. Long before our ancestors resembled anything like a human being, and, in all probability, long before they resembled primates or even mammals, individuals who represented the movements of predators as being importantly different from the movements of other objects in their environment would certainly have had a fitness advantage over individuals who drew no such distinction. And our distant reptile ancestors would have had every bit as much reason as our recent primate ancestors did to make attributions of certain cognitive states to potential predators, because in either case a prey animal’s escape strategy often depends on whether it believes the predator has seen it. To be sure, a trait that first evolved back when members our lineage were fish or reptiles would not be a trait that evolved for religion, but there is room in evolutionary history for a lot to happen after this trait first evolved.

What Bloom’s account does not explain, as Sosis points out, is why a capacity for representing bodiless agents in general should be employed so frequently, across such a wide range of distinct religious systems, for representing bodiless agents of a very specific sort: agents who care about human social life, and who possess supernatural powers enabling them to direct natural events in ways that will punish and reward humans. Back when members of our lineage were still reptiles or four-legged mammals, there would have been nothing adaptive about beliefs involving bodiless agents with these particular properties. But by the time our ancestors began to resemble modern humans, ecological pressures would have changed in relevant ways. For example, as language capacities evolved, individuals would have acquired the ability to share information with each other about others’ actions. This, as Bering notes, would create a selection pressure for the effective management of one’s reputation. So these selection pressures would
make it adaptive to have the very specific type of dualistic beliefs that we actually observe across a wide range of distinct religious systems.

Thus, while the selection pressures to which Bloom appeals would most likely have been in place millions of years before the evolution of language, those to which Bering appeals could only have appeared after the evolution of language. Vast stretches of time could separate the facts explained by these two theories, yet both theories explain something important about the evolution of the same trait. The critical difference between the two accounts lies in the fact that while Bloom explains the original evolution of the trait in question, Bering explains instead how this trait subsequently acquired a new functional role, in relation to the traits of social motivation.

As a constituent trait of the religious phenotype, dualistic cognition is a functional part of the Goldberg device. What Bloom offers is an account of the original function of this part, the independent function it possessed before it was incorporated into the larger device. By contrast, what Bering offers is an account of how this trait came to be incorporated into the device. In other words, Bloom explains how a capacity for dualistic cognition came into existence in the first place, while Bering explains how this trait was later exapted to perform a specifically religious function.

3.8 Conclusion

I hope to have shown two things in the preceding discussion. The first is that a careful examination of the theories in question reveals them to be mutually consistent. These theories do not disagree about whether religion is an adaptation, because they do not take the terms ‘religion’ and ‘adaptation’ to refer to the same sets of empirical facts. As a result, the fact that they provide
opposing answers to the question “Is religion an adaptation?” does not show that they provide conflicting, competing explanations for the same natural phenomena.

The second thing I hope to have shown is that there already exists in the literature a mechanism for unifying these two theories. The Goldberg Exaptation Model, which has already been proposed independently by both an adaptationist and a by-product theorist, illustrates the critical difference between the task of explaining the function of the whole complex of traits that distinguishes religious from non-religious individuals and the task of explaining the particular functions of particular constituent traits within this complex. Even if no particular trait possesses a specifically religious function, it may yet be the case that the complex itself possesses a specifically religious function. Against this background, it is significant that adaptationist theories all agree about what that function is: to motivate cooperation even when cooperating is costly.

It must be pointed out that a few adaptationists who emphasize the role of cultural evolution have already recognized the potential for unification. While they have made no effort to diagnose the source of all the debate and disagreement, and while they do not employ the terms of the GEM, the positive account they provide is one in which cultural evolution has cobbled together a Goldberg device using the raw materials of genetic evolution. Anthropologists Scott Atran and Joseph Henrich say of their own theory: “This synthesis integrates insights from studies of the cognitive foundations of religion with evolutionary approaches to human cooperation to derive a deeper understanding of the origin and development of prosocial religions” (2010, p. 19). Likewise, psychologist Ara Norenzayan takes integration to be a major aim of his account: “The argument in this book is an attempt at integrating these two perspectives—the social and the cognitive—that are currently seen as
competing accounts” (2013, p. 11). He ultimately concludes that “successful religious groups—
the cultural ancestors of most human beings alive today—pieced together, step by step, a whole
cluster of psychological mechanisms that, building on supernatural monitoring and credible
displays of sincere faith, fostered and cemented social solidarity” (2013, p. 162).

Narrow adaptationists, who reject appeals to cultural evolution altogether, will not buy this. And perhaps they will reject McCauley’s metaphor on grounds that, while the religious phenotype is indeed a complex, it is not a device, because it is a complex that possesses no evolutionary function. But if narrow adaptationists are right about this, then something will still have to explain the formation of the complex. And if broad adaptationist theories do not succeed in doing this, it will not be because narrow adaptationist theories offer a better explanation. As Sosis rightly notes, narrow adaptationists have not even tried to explain why the various by-products they identify have come together to form the specific complex of traits distinguishes religious from non-religious individuals. On the other hand, even if this complex does turn out to be a Goldberg device with the characteristic function of motivating cooperation, this also would not show that that there is any flaw in the by-product theory, since the aim of by-product accounts is only to explain various constituent parts of the device. The GEM thus offers a way of understanding how existing theories from both sides of the purported debate actually make distinct and complimentary contributions to shared theoretical enterprise, forming distinct and complimentary parts of the same unified theory of religion evolution.
4.1 Introduction

Scientists studying the evolution of morality, religion and human cooperation in general are divided over an important methodological issue: the scope of adaptationist theory, or of theories based on selection. Traditionally, adaptationist theories of human behavior and cognition have adhered closely to the biological orthodoxy of inclusive fitness theory, appealing only to genetic selection, and only to selection occurring at the individual level (E. O. Wilson 1978; Dawkins 1976; Barkow, Cosmides and Tooby 1992; Pinker 2002).8 Recent years, however, have witnessed a growing number of psychologists, anthropologists, biologists, economists and philosophers appealing to cultural selection in addition to genetic selection, and to group selection in addition to selection at the individual level (D. S. Wilson 2003; Richerson and Boyd 2005; Atran and Henrich 2010; Haidt 2012; Norenzayan 2013; Bowles and Gintis 2011; Sterelny 2006; Mesoudi, Whiten, and Laland 2006). It has thus become necessary to distinguish between narrow and broad forms of adaptationist framework in the study of human evolution. The narrow framework confines its appeals to selection solely to one type of selection process: genetic selection at the individual level. By contrast, because genetic selection and cultural selection can both occur at either the individual or the group level, the broad framework recognizes not one but four distinct forms of selection process, each of which may, in principle, explain the evolution of a given trait. (See Figure 4.1.)

---

8 All parties involved agree that selection sometimes also occurs below the individual level, at the level of particular (sets of) genes (cf., Dawkins 1976).
I argue here against the narrow framework, and in defense of the broad framework, on grounds of scope. Appealing to additional forms of selection processes enables us to explain facts that fall outside the scope of the narrow methodology, leaving no motivation for imposing the constraints on selectionist theorizing that the narrow framework imposes. Several prominent adherents to the narrow framework appear to disagree, because they continue to present their own theories as competing explanations for the same phenomena explained by broad theories (Baumard, Andre, and Sperber 2013; Baumard and Boyer 2013; Krasnow et al. 2012; Pinker 2010). This assumption, however, is based on a mistaken understanding of the relationship obtaining between broad and narrow theories. Because the broad framework includes or encompasses the commitments of the narrow framework, there is no reason why specific empirical theories derived from these two frameworks should conflict, and, in fact, existing theories do not.
My argument proceeds by identifying facts of moral and religious psychology that need to be explained somehow, and that only the broad framework is capable of explaining. These facts concern the evolution of the social environment—the specific beliefs, norms, skills and practices that pervade particular social groups, and that play an essential role in structuring the psychology of individual group members. Cultural selection explains the prevalence and stability of certain patterns of cognition and behavior that persist within groups across many human lifetimes, and that figure essentially in explanations of individuals’ moral and religious psychology. By explaining the evolution of a group’s beliefs, norms and institutions, cultural selection explains features of the social environment that are necessary for explaining both what an individual learns from others and how others’ behavior affects her genetic fitness.

Among the traits of moral and religious psychology influenced by these features of the social environment are the specific contents of religious beliefs and moral norms, which differ from one group to the next, and which are necessary for distinguishing members of one group from those of another. The narrow framework may offer a partial explanation for the fact that the Abrahamic God is believed to take an interest in the moral behavior of human beings, but it cannot help us explain the most crucial psychological differences between Christians, Jews and Muslims. Inclusive fitness theory by itself cannot explain, for example, why Muslims believe that Jesus was a prophet, while Christians believe that he was the Son of God. And, for the same reasons, the narrow framework is also insufficient for explaining any individual’s attitudes toward abortion, same-sex marriage, or use of the death penalty. The phenomena of morality and religion that fall outside the scope of the narrow framework do not constitute a rare and obscure set of minor curiosities. Especially in the domains of morality and religion, they are among the most important phenomena to be explained. As a result, these domains are particularly well
suited to the task of illustrating the limitations of the narrow framework, and the benefits of expanding adaptationist methodology to include cultural selection and group selection in addition to genetic selection.

4.2 The Crux of the Issue: Genes vs. Culture

Narrow adaptationists take their accounts of morality and religion to be in competition with theories derived from the broad framework, and they present their own theories as the better of the two competing accounts (Baumard, Andre, and Sperber 2013; Baumard and Boyer 2013; Krasnow et al. 2012; Pinker 2010). But if this were the case, it would entail that broad and narrow theories provide different explanations for the same empirical phenomena, the same target facts. They do not.

The target phenomena of these two frameworks do overlap, because both purport to explain genetically inherited traits selected for at the individual level. In this domain, however, the theories they provide do not differ, and thus do not compete. All parties agree that the evolutionary mechanisms posited by inclusive fitness theory, such as kin altruism and reciprocal altruism, are fundamentally important for explaining human cooperation. But where the explanations provided do diverge, so too do the phenomena explained. When broad adaptationists appeal to cultural evolution and group selection, they do so in order to explain traits that individuals acquire through social learning. These facts fall outside the scope of the narrow framework, since it is only capable of addressing traits that individuals inherit

---

9 Wilson (2003) and Haidt (2012) are exceptions, since they appeal to genetic group selection in addition to cultural group selection. But as I argue in Chapter 1, they are not justified in doing so. They do not respond to the explicit arguments against such appeals, discussed in Section 2.1, that other broad adaptationists have provided (Richerson and Boyd 2005, p. 203).
genetically. As I explain further below, broad adaptationists assume that capacities for social learning are themselves genetically inherited traits. But to explain the function of a capacity for learning in general is not to explain the specific content of what is actually learned.

Indeed, the role of social learning lies at the very heart of this issue. Narrow adaptationists don’t deny that social learning occurs, or that it explains how individuals come to possess various beliefs, attitudes, and other traits of moral and religious psychology. Yet they reject the central methodological tenet of the theory of cultural selection, which is that social learning may be treated as cultural inheritance; an alternative, non-genetic mechanism by which psychological traits spread from one individual to another. This tenet makes it possible to explain cultural change by using the same models of selection that population geneticists use to track the selective retention of heritable traits. No matter how traits are inherited, selection models describe how they spread, why they are sustained at particular frequencies across generations, and how they gradually accumulate functional complexity through descent with modification. That said, it is important not to overstate the similarities between genetic evolution and cultural evolution, and even more important to recognize that the explanatory power of cultural selection models is not dependent upon an analogy with genetic selection. Changing the mechanism of inheritance from genes to learning radically alters the evolutionary dynamics, changing many things about how selection proceeds, and about the role it plays in explaining psychology. In the rest of this section, I focus on some important differences between them, in order to state more precisely what the differences between broad and narrow methodology are.
4.2.1 Setting Group Selection Aside

One crucial difference between cultural selection and genetic selection concerns the role of selection occurring at the level of groups, rather than individuals. Whether the traits in question are inherited genetically or culturally, group-level selection requires group-level variation. In standard models of group selection, this means that individuals from the same group share the same variants as each other, while individuals in different groups possess different variants from each other. If the variants that Mother Nature must choose between are located together within the same group, then when the whole group gets selected, all of these variants get selected together, in a package deal. That would not be selection between those individual-level variants. Further, for a process of group selection to count as genetic group selection, the relevant group-level variation in traits must be caused specifically by an underlying pattern of group-level variation in genes. Moreover, this genetic variation must be maintained continuously for the duration of the selection process. These conditions are difficult to satisfy in human populations (Bell, Richerson, and McElreath 2009; Langergraber et al. 2011).

Over the course of a single human lifetime, genetic mixing between neighboring groups could be minimal. But genetic selection takes many, many human lifetimes, and this leaves plenty of time for even low rates of migration to add up to large amounts of genetic mixing between sub-populations. This erodes the group-level genetic variation on which genetic group selection depends. In theory, genetic group selection can still occur, as long as selection pressures are strong enough, or migration rates are low enough. In fact, however, theorists of cultural selection Richerson and Boyd cite modeling work showing that “Even very small amounts of migration are sufficient to reduce genetic variation between groups to such a low level that group selection is not important” (2005, p. 203).

10 For a detailed analysis and review of group selection theory in general, see Okasha 2006).
Cultural group selection, however, does not face this obstacle (Henrich 2004; Baldini et al. 2013). Cultural migration also occurs, of course, since individuals do sometimes adopt beliefs and practices from foreign groups. But most people learn far more from in-group members than from foreigners, and as a result cultural evolution tends to generate many more differences than similarities at the group level. Indeed, it has been shown empirically that culturally inherited traits exhibit much more group-level variation than genes do (Bell, Richerson, and McElreath 2009).

This difference between genetic variation and cultural variation is crucial for understanding broad adaptationist theories. For while the broad framework recognizes four types of selection in principle, broad theories only appeal to three of these four in practice\textsuperscript{11}: genetic selection occurring at the individual level, and cultural selection occurring at both the individual and the group level. Genetic group selection is ruled out on empirical grounds, because of a lack of group-level genetic variation. Thus, in practice, all appeals to group selection in the broad framework are appeals to cultural group selection, not genetic group selection (see Figure 4.2).

\textsuperscript{11} Again, Haidt (2012) and Wilson (2003) are exceptions, but their appeals to genetic group selection are unwarranted (see footnote 2).
The difference between genetic variation and cultural variation thus makes the role of group selection in human psychology very different from the (highly controversial) role some biologists have attributed to it in other species (e.g., Nowak, Tarnita, and Wilson 2010). Indeed, from the broad adaptationist perspective, the role group selection plays in human evolution may be unique in the animal kingdom. But more importantly, this difference shows that the disagreement over group selection between broad and narrow adaptationists is entirely derived from a prior, more fundamental disagreement over appeals to cultural selection. Were it not for the fact that cultural inheritance provides far more group-level variation for selection to work with than genetic inheritance provides, broad adaptationists would agree that group selection is not important in human evolution. Accordingly, in what follows I will assume that the broad/narrow debate ultimately depends upon the more fundamental debate about the status of cultural selection. If appeals to cultural selection cannot pay for themselves independently, then there are indeed good reasons to reject appeals to group selection as well. But if selection among traits inherited by social learning does turn out to play a significant explanatory role, as I will...
argue below, then there is no reason not to appeal to group selection—as long as this is not genetic group selection.

4.2.2 Evolving Traits and Evolving Environments

Another crucial difference between cultural selection and genetic selection concerns their roles in explaining the evolution of human environments, which subsequently impact the psychology of individuals. Following Laland, Odling-Smee and their collaborators (2013; Laland et al. 2011), Sterelny (2003; 2006; 2012) has articulated in a number of ways the importance of niche construction in human evolution, and the importance of cultural selection in the construction of the human niche. Like beavers, ants and many other species, humans do not simply inhabit their environments—they also construct them, modifying their physical features in ways that are adaptive, and that influence the subsequent course of genetic selection. But while beavers construct ponds, and ants construct anthills (and even farms! (Suen et al. 2011)), humans construct cities that are nothing like their surroundings, and they cover millions of square miles with their crops, pastures and roads. Much to the chagrin of conservationists, we alter our physical environments to an extent that is orders of magnitude beyond that of any other species.

One important factor explaining the extreme degree to which we construct our environments is the accumulation of scientific and technological knowledge. This is information transmitted from one individual to another by social learning, so most of the beliefs we use today to build skyscrapers and interstates may be viewed as products of cultural selection. Over millennia, useful beliefs and technologies get retained, spread, and improved upon, increasing in sophistication and functional complexity. Less useful beliefs and technologies are rejected or forgotten. As a result, our physical environments have largely evolved under the guidance of
culturally selected, culturally inherited techniques and beliefs, including scientific beliefs. In turn, however, these physical modifications have also affected the subsequent evolution of scientific and technological knowledge. It’s much easier to conduct many kinds of research in buildings with electricity and running water than in other, less developed environments, and that’s not to mention large hadron colliders.

Importantly, however, features of the physical environment are not the only ones that are constructed largely under the influence of cultural selection. Humans also construct their social environments in ways that influence the subsequent course of evolution, and cultural selection is even more important for explaining some of these environmental features. In a process that Richerson and Boyd have dubbed gene-culture coevolution, selection on culturally inherited traits creates and maintains stable patterns of behavior within populations that subsequently influence the course of genetic selection in those populations (Laland, Odling-Smee, and Myles 2010; Richerson and Boyd 2005). The most well-documented example involves the evolution of genes for lactose tolerance in adults (Laland, Odling-Smee, and Myles 2010; Holden and Mace 2009). In certain groups, culturally inherited practices for keeping livestock created an environment in which a new source of nutrition was readily available: milk. When these practices first appeared, however, humans in general lacked the ability to digest milk, so there was no way to take advantage of this available source of nutrition. But chains of cultural transmission sustained and improved these practices of keeping livestock, maintaining the availability of milk in these environments over very long periods of time. Thus, culturally selected practices of pastoralism established strong and long-lasting selection pressures for the capacity to digest milk, and eventually the genes responded. Holden and Mace (2009) estimate
that over a period of approximately 8,000-9,000 years, individuals from certain pastoralist cultures evolved capacities for digesting milk.

Note that the selection pressures described in this example are due to the social environment, not the physical environment. This is not to say that the animals inhabiting a given environment don’t count as features of the physical environment, any more than it is to say that the other people within an environment don’t count as features of the physical environment. Rather, it is to say that domesticated livestock are not like the other animals inhabiting a given environment, because their presence depends upon human behavior. Herds of livestock are not independent features of the physical environment, but are among the modifications that people make to the physical environment. The herds that provided this new source of nutrition depended for their existence upon the pastoralist practices of the people who kept them, and these practices evolved by cultural selection, in the same way that scientific and technological knowledge evolves: individuals acquire information from their cultural ancestors and pass it on to their cultural descendants. Thus, it is not just human genes, such as those involved in the digestion of lactose, that evolve in response to the culturally selected social environment. The social environment itself evolves through cultural selection as well.

Moreover, people do not simply inherit information from their cultural ancestors, and pass it along to their cultural descendants. They also tinker with this information in the process, making improvements to the beliefs and practices they inherit before passing them on to others. Over time, these improvements accumulate, so the information transmitted gradually increases in functional complexity and sophistication, much as such complexity accumulates among genetically inherited traits. Boyd and Richerson (2005) refer to this as cumulative cultural evolution, and Tomasello (1999) has dubbed it “the ratchet effect,” noting that when social
learning takes the form of imitation, and when imitation occurs with high enough fidelity, 
selected information can be “held in place” long enough for additional innovations and 
improvements to be made. Thus, innovations represent new turns of the ratchet, while high-
fidelity social learning prevents the ratchet from slipping between cultural generations.

Not all social learning in humans occurs with high fidelity, of course; slip happens. But it is 
not unusual for traditional beliefs and practices to be retained within populations for hundreds or 
thousands of years, and this provides plenty of time for many small improvements to accumulate. 
Newton’s famous claim to have been standing on the shoulders of giants nicely illustrates the 
cumulative nature of scientific theorizing, but since most of this accumulation happens in small 
increments, most scientists are not giants in the relevant sense. (Newton himself, of course, was a 
giant.) Other examples abound. The first roads were not paved, and the first paved roads were 
not paved with asphalt. The same basic principles of democracy have been retained across many 
centuries, and have spread across many different social groups, but they have also been improved 
upon and developed in different ways, resulting in many different democratic systems. And the 
formal establishment of the Papal Inquisition in 1231 represents just one more turn of a ratchet 
that had already been turning for centuries, as the institution of the Catholic Church gradually 
accumulated a remarkable degree of functional complexity.

This last example illustrates the special significance of norms as features of the social 
environment that are constructed by cultural selection, and that subsequently alter the course of 
selection—both in cultural and in genetic evolution. By punishing and rewarding the behavior of 
others, we change those behaviors’ adaptive payoffs, and the Inquisition provides a salient 
example because the punishments involved were infamously fitness-reducing. While the 
adaptive consequences of most norms are less dramatic, the collective consequences of norms on
fitness are substantial. Violating a particular norm here or there may have little impact on one’s genetic fitness, and violating even a large number of minor norms probably won’t cause one to be burned at the stake—though it might make it hard to hold down a good job or attract a good mate. Nevertheless, anyone who lacked capacities for identifying and complying with norms in general would possess extremely low genetic fitness in almost any human social environment.

As features of the social environment, then, norms themselves highlight the adaptive significance of traits of norm psychology: the capacities involved in successfully negotiating an environment filled with other people who will reward and punish all kinds of behavior. Many of these capacities constitute central features of moral and religious psychology, and while some are inherited genetically (e.g., motivational capacities for guilt, sympathy or anger), others are inherited culturally (e.g., the belief that heretics should be killed in a way that makes an example out of them). Indeed, Sterelny (2003, p. 166), Sripada (2007) and Richerson and Boyd (1995) have all argued that our capacities for high-fidelity imitation are best thought of as genetic adaptations, because they provide individuals with access to the accumulated knowledge of whole cultural traditions. An Inuit individual needn’t start at the beginning when trying learn how to build a kayak using driftwood and seal skins. He can behave far more adaptively by adopting the skills and beliefs accumulated by his cultural ancestors. If so, then capacities for social learning themselves would be genetically inherited traits. Nevertheless, the particular content of what is learned—the specific beliefs, motivations, skills or practices one adopts from others—would be culturally inherited traits.

Thus, the social environment of the Inquisition would not only have affected individuals’ genetic fitness. It would also have determined, to some degree, what those individuals were capable of learning from cultural models. People inhabiting the environment of the Inquisition
would not only have learned to avoid saying anything heretical, but they would also have learned, most likely, that Jesus was the son of God. In many other social environments this would have been impossible, because there was no one else available from whom this belief could be inherited. When a child today learns from her parents or priests that Jesus was the Son of God, she is culturally inheriting a belief the content of which evolved centuries before she was born. The child would not be able to possess this belief—a belief with this specific content—were it not for the behavior of the adults who constitute her social environment. But the adults’ behavior, in turn, is explained by the fact that they, too, inherited the same belief in the same way, as did their own parents before them. Two thousand years ago there must have been innovators of this idea, individuals who formed this same belief in some other way, without exercising capacities for social learning. But once this belief appeared, it could be imitated by people who otherwise had no way of accessing the relevant facts and events—including people who lived thousands of years later. Even during Jesus’s own lifetime, however, when it was possible in principle to hear his words and observe his deeds, most people who formed supernatural beliefs about Jesus would have acquired most of those beliefs from others. Through the mechanism of cultural inheritance, then, this particular belief has not only remained strikingly common for a very long time, but it has also spread all over the globe, becoming prevalent in many populations that are otherwise very different from one another. So religious beliefs, like norms, are stable, prevalent features of the social environment, and their prevalence is maintained by cultural selection. Whatever effects these beliefs may have on the genetic fitness of individuals, they often have powerful effects upon traits of moral and religious psychology that individuals inherit by learning.
In particular, culturally selected features of the social environment are essential for explaining why individuals are members of the particular religions to which they adhere. The belief that Jesus was the Son of God is, for many, the essential, defining trait of the Christian system. If so, then it is only by explaining why individuals possess this belief, in particular, that we can explain why any given individual is a Christian, rather than a Jew, a Hindu or a Taoist. In the same way, the role of cultural selection in constructing the social environment is also essential for explaining why individuals adopt and enforce the particular norms they in fact adopt and enforce, including both moral and religious norms.

4.3 Broad Adaptationist Theories of Morality and Religion

Theories of moral and religious evolution—narrow and broad alike—proceed by identifying cognitive, motivational and behavioral traits that are simply assumed, on intuitive grounds, to count as moral or religious traits. The primary target for most narrow theories of religion, for example, has been belief in supernatural agents. But it is simply assumed that beliefs about supernatural agents count as religious traits, or as constituent features of a specifically religious phenotype. They fully recognize that a comprehensive account of religious psychology will require additional accounts of other constituent features as well, so the assumption is simply that belief in supernatural agents is one important trait among many. But there is no generally accepted analysis of the concept religion to which scientists can appeal in order to determine exactly which traits count as the religious ones. So there is plenty of room in the conceptual space of common-sense intuition for researchers to study a wide range of different traits and facts underneath this heading. The same is true for the traits and facts of the moral domain.
Similarly, narrow and broad theorists alike assume that, whatever else they may be, religion and morality are fundamentally social phenomena, involving cooperative interactions between individuals. Indeed, both frameworks explain the evolution of morality and religion by explaining the evolution of certain forms of cooperation, which are intuitively considered to be moral and/or religious. Moreover, both have attempted to explain the connection between morality and religion by explaining how moral norms and intuitions have influenced the cultural evolutionary history of religion. In his recent book *Big Gods* (2013), the broad adaptationist Ara Norenzayan focuses specifically on the evolution of religious systems containing gods who care about humans’ moral behavior, like the “big” God of Abraham. While a majority of the world’s population now believes in such big gods, the connection between religion and morality, Norenzayan argues, is a relatively recent occurrence, even by the standards of cultural evolution, rather than genetic evolution. He cites ethnographic studies showing that the religious systems of small social groups—like those inhabited by almost all human beings throughout most of human history—all posit gods, spirits and other supernatural agents. But very few of these agents know or care about the normative behavior of humans. Norenzayan’s book provides what is currently the most fully developed and carefully articulated version of the broad adaptationist account of religion, and it is literally *all about* the cultural evolutionary basis of the connection between religion and morality. Accordingly, it may be a bit surprising to learn that I will not be focusing on this state-of-the-art account, because it is, in a sense, too fancy for my purposes. Because my own aims are methodological, rather than empirical, my concern lies with the fundamental building blocks of the broad adaptationist approach to cooperation, which were first laid out in Boyd and Richerson’s *tribal instincts hypothesis*. Norenzayan adds a wealth of empirical details to these building blocks, but he does not change them.
Where broad and narrow theories really diverge is with regard to the particular forms of cooperation on which they focus. As I explain in the next section, narrow theories of morality and religion focus primarily on mutually beneficial forms of cooperation, since the benefits of mutualism explain what the relevant traits were selected for. Broad theories, by contrast, tend to focus on altruistic cooperation, because doing so helps to illustrate the independent contribution of cultural selection theory, over and above the explanations provided by genetic selection. This is not to suggest that the theoretical labor is simply divided, such that genetic selection explains mutualistic cooperation, while cultural selection explains altruistic cooperation. Genetic selection can explain—or explain away—at least some forms of altruism, and cultural selection can explain at least certain aspects of mutualism. But, as I explain in detail below, broad adaptationists have motivated their own theories of cooperation largely by identifying forms of cooperative behavior that narrow accounts have trouble explaining, and these have tended to involve cooperation that is costly, rather than beneficial.

Here I describe the tribal instincts hypothesis in general, as the foundation of the broad adaptationist approach to cooperation in humans. In doing so, however, I focus specifically on what that hypothesis explains about religion, morality and the evolution of the social environment. After addressing a crucial terminological matter, I first explain how cultural selection affects the relevant features of the social environment. I then turn to consider how the culturally selected environment subsequently impacts the moral and religious psychology of individuals.
4.3.1 ‘Altruism’ and ‘Prosociality’

One of the central concepts in the literature on the evolution of cooperation is that of altruism, which refers to behavior that is adaptively beneficial to others, but costly to the agent. Explaining how cooperation evolved requires explaining how it evolved in spite of its costs, and the costs of cooperation can sometimes be substantial, especially in humans. Unfortunately, “altruism” is not used with complete consistency in this literature, and, making matters worse, it has come to compete with a distinct but closely related term: “prosociality.” Here I explain how I will define these terms, and how they are related to one another.

First, like mutualism, altruism in this context is behavior that is actually costly, as measured in the currency of genetic fitness. It doesn’t matter how agents perceive their own actions, or whether the contents of their mental states can be said to count as genuinely other-regarding. Whatever it may mean for mental states to be altruistic, altruistic behavior is behavior that reduces the genetic fitness of the agent, while increasing the fitness of other individuals. There is no confusion about this much.

Beyond this, however, the boundaries of the concept of altruism become fuzzy, because of contingent historical facts about the literature on the evolution of cooperation. A great deal of evolutionary theorizing has been devoted to altruism because of the prima facie challenge it poses to inclusive fitness theory. If costs and benefits are measured in the currency of genetic fitness, and if altruism is costly by definition, then not only can altruistic behavior not be selected for, but it should have been selected against. Since altruistic behavior clearly occurs, it is a puzzle why the relevant genes have not been selected out of the gene pool. The primary response to this puzzle has been to show that behaviors that appeared altruistic at first turn out not to be so on closer examination. In other words, the primary response to the puzzle of altruism has been
not so much to explain it as to explain it away, and this approach has been highly successful in the biological world in general.

In what is known as *kin altruism*, or kin selection, the costs paid by the altruist benefit her genetic relatives. So while the altruist herself makes a sacrifice, *qua organism*, her genes do not make the sacrifice with her, and her inclusive fitness is increased, rather than decreased (Hamilton 1964). And in what Trivers (Trivers 1971) refers to as *reciprocal altruism*, behaviors that are costly in the short term remain beneficial in the long term, because the costs paid by the altruist in one interaction are repaid later by the beneficiary, in subsequent interactions. Neither of these accounts actually explains why a costly form of behavior occurs, because both accounts proceed by showing that certain forms of behavior are *not* costly (despite appearances). As Trivers puts it (1971, p. 35), “Models that attempt to explain altruistic behavior in terms of natural selection are models designed to take the altruism out of altruism.” His own models then demonstrate that reciprocal “altruism” is actually a form of mutualism, or cooperation for mutual benefit.

In the wake of these successes of inclusive fitness theory, altruism came to be defined largely (but not consistently) as behavior that is costly *overall* in the currency of genetic fitness. The theories of kin altruism and reciprocal altruism illustrated the need to distinguish between costs and *net* costs, and when certain forms of behavior were shown to be costly only in a limited or short-term sense, this was taken to mean that they didn’t actually count as altruism after all—the altruism had been taken out of them, in Trivers’ terms. Notice that there was no need to define altruism in this way, in terms of overall costs rather than limited or short-term costs. There is a perfectly respectable sense of ‘cost’ in evolutionary biology that does not imply net costs. For example, foraging is costly. In addition to metabolic costs for locomotion and cognition (if
food is hard to find), foraging exposes many animals to increased risk of predation. These are all costs, but they do not imply that foraging reduces fitness overall. And there is no reason that altruism could not be defined similarly, as behavior that is costly in the same short-term sense in which foraging is costly. Nevertheless, I will follow here what I take to be the dominant convention, and reserve ‘altruism’ specifically for behavior that exacts net costs, or that results in an overall reduction of an agent’s inclusive fitness.

From the broad adaptationist perspective, however, this convention is inconvenient. This is because broad theories need to be able to refer to costly forms of cooperation without presupposing anything about overall fitness value. Questions about the coevolutionary interactions between genetic selection and cultural selection are empirical questions, so for any given cultural trait it must remain an open empirical question what that trait’s impact on genetic selection might be. But if altruism already entails, by definition, that a trait reduces fitness overall, then to call a trait altruistic is already to assume that this empirical question has been closed—without conducting an empirical investigation. Since the standard definition of altruism does precisely this, broad theorists have tended to employ a different term instead: ‘prosocial.’

Like altruistic behavior, prosocial behavior is costly by definition. But unlike altruistic behavior (as it has come to be defined), prosocial behavior does not imply that these costs are net costs. To say that a behavior is prosocial is only to say that there is some sense in which it is costly. This leaves open the possibility that the same behavior is also, in some other sense, beneficial. Accordingly, it also leaves open the possibility that these benefits are greater than the relevant costs, resulting in a net fitness value that is positive, not negative. For example, if a mother is injured while fighting to defend her young, it is already clear that she has behaved prosocially. But it is not yet clear whether this prosocial behavior is also altruistic. If the injury
proves fatal, and the young cannot yet care for themselves, then the fighting behavior will be both prosocial and altruistic. But if the injury is minor, and everything is back to normal in a day or two, then the fighting behavior will be prosocial, but not altruistic.

In light of these considerations, much of what broad theories have actually focused on explaining is prosociality, rather than altruism. However, it is important to see that the broad framework can also explain full-fledged altruism, and not just by explaining it away. Indeed, it can explain altruism in a way that is entirely continuous with the account of (mere) prosociality it provides. Boyd and Richerson, who have developed many of the basic models and concepts of cultural selection theory, argue in detail that a wide range of culturally inherited traits really are fitness-reducing overall (2005, Ch. 5). For example, Pentecostal snake handlers learn from their cultural ancestors to place their fate in God’s hands, and some then die from snakebites. On the other hand, they also argue that culturally inherited beliefs and behaviors enhance individual fitness more often than they reduce it, and this is why capacities for high-fidelity imitation were selected for in the first place. By inheriting the accumulated knowledge of cultural ancestors, they argue, people are able to behave adaptively in a wide range of environments, including harsh environments that are very different from those in which their genes evolved. Thus, Richerson and Boyd describe the relationship between cultural selection and genetic fitness as follows: “Acquiring information from others allows people to adapt to a wide range of environments, but it also opens a portal into people’s brains through which maladaptive ideas can enter—ideas whose content makes them more likely to spread, but do not increase the genetic fitness of their bearers” (2005, p. 150). Or, in different terms, “A certain frequency of maladaptation inevitably results from design tradeoffs confronting an advanced cultural creature” (2005, p. 244).
In the broad framework, then, any given culturally inherited trait may be costly overall (altruism), but it also may not be (prosociality). Costly cooperation is sometimes prosocial, sometimes genuinely altruistic, and for any particular form of cooperative behavior it remains an open empirical question which label applies.

4.3.2 The Tribal Instincts Hypothesis

At the center of broad theories of cooperation is what Boyd and Richerson refer to as the *tribal instincts hypothesis* (Richerson and Boyd 1998, 2005; Boyd and Richerson 2006). This theory holds that humans possess genetic adaptations, or “instincts,” for cooperating with other members of an in-group, or “tribe,” which is just a social group that is significantly larger than an extended family—large enough to warrant being defined in abstract, symbolic terms, rather than by keeping track of relatives. They motivate this theory by pointing to what they call “the puzzle of human cooperation,” where the puzzle is to explain why cooperation in humans is so different from cooperation in other species. More specifically, the puzzle is that of explaining the scale of human cooperation. How did tribes grow to the incredible sizes of many modern nations? And what set of instincts can explain this, given that they haven’t evolved in any other species?

The concepts of kin altruism and reciprocal altruism are capable of explaining cooperative behavior that is impressively complex, such as the social hunting of lions, or the fungus farming of leaf-cutter ants. But family groups are small among primates, so kin selection does not explain the large-scale cooperation of humans. And Boyd and Richerson argue that reciprocity requires advanced and expensive cognitive capacities for tracking distinct histories of exchange with distinct partners. So while reciprocity-based mutualism does occur in other species, it cannot
explain cooperation among large numbers of interactants (Boyd and Richerson 1988). Yet humans regularly cooperate with perfect strangers whom they will never see again, and they do so by the millions. Some modern armies contain millions of soldiers, and these soldiers are paid and supplied by tax dollars collected from many more millions of civilians. All of these individuals work together to provide the collective good of national defense, but none of them will ever know the names or faces of the vast majority of their cooperative partners. This we do not find in other species. Ants and bees do cooperate in large networks, but the interactants are all members of the same immediate family. So how did humans evolve the psychological capacities that make it possible to form enormous groups bound together by something other than kinship and reciprocity?

The solution to this puzzle, Richerson and Boyd claim, is a coevolutionary interaction between cultural selection—at the group level—and genetic selection—at the individual level. On this view, cultural group selection explains the scale of human tribes because of the role culturally inherited beliefs, norms and institutions play in promoting prosocial cooperation. In this sense, it explains the composition and evolution of the social environment. Genetic selection then explains, subsequently, why individuals inhabiting these environments have evolved certain social “instincts,” or psychological capacities.

4.3.2.1 Cultural Group Selection and the Evolution of the Social Environment

Boyd and Richerson argue that humans are an exceptionally cooperative species because they are an exceptionally cultural species—where culture is defined strictly in terms of social learning. Like kin altruism and mutualism, social learning occurs in many species, from guppies (Brown and Laland 2003) to chimpanzees (Boesch and Tomasello 1998), so the claim is not that humans
have culture while other species do not. The difference, rather, is one of degree; no other species engages in social learning to anything like the extent that humans do. Boesch and Tomasello (1998, p. 593) point out that chimpanzees exhibit group-level variation in many forms of socially learned behavior, such as using sticks to extract marrow from the bones of prey. But they also point out that even in chimpanzees it remains unclear whether cultural information accumulates (p. 602). Social learning in chimpanzees rarely, if ever, takes the form of high-fidelity imitation, so if the ratchet effect occurs at all, then it does so only to a minimal degree. Humans, by contrast, are awash in culturally accumulated information. And as distinct norms, traditions and pools of information accumulate in distinct cultural groups, group-level variation increases, leaving plenty of scope for selection to occur among traits that (1) vary at the group level, and (2) are inherited culturally. This is cultural group selection.

Consider a simple model. Suppose that as a result of a long process of cumulative cultural evolution in an isolated group, everyone now adheres to a religious system specifying that no human being should ever be harmed. And suppose that in a different group, a different evolutionary history has caused its members to be committed to a religious system specifying that to kill an enemy in battle is the highest form of honor. Then suppose the warlike group loses a battle, is displaced from their land, and comes across the pacifist group—with its land and resources. If these groups are otherwise similar to ordinary groups in human history, the outcome of the encounter would be that the frequency of pacifist beliefs and values had dropped substantially in the overall population, while the frequency of warlike beliefs and values had increased. The primary cause of this change in frequencies would be that most of the pacifist individuals had simply been removed from the population. But even if the women and children are spared, these children will be much less likely to acquire their group’s pacifist values. For
some, this may be because the victimization they’ve suffered has caused them to explicitly reject these beliefs. For others, the problem may be that they no longer have the right kinds of cultural models to learn from. Their social environment is very different from the one in which their parents acquired pacifist commitments; now the only men around are warriors, not pacifists. Any of these mechanisms for changing the frequency of pacifist beliefs and values in the meta-population is an illustration of cultural group selection.

This example may seem a funny way of illustrating how morality and religion evolved, since warrior values don’t seem especially moral by our standards. But the example is constructed to illustrate selection for prosocial and altruistic norms, rather than norms that conform to modern, Western intuitions about morality. Warrior values are actually quite prosocial in the relevant, technical sense. Warriors are not people who make war on their own group members. Indeed, war requires not just a great deal of coordination and communication within the group, but also a great deal of sacrifice on the part of individual group members. As I’ve described warrior values here, achieving “the highest form of honor” requires one to go into battle, and that is a very costly, risky thing to do. It is thus important to see that, while warlike norms may motivate individuals to do horrible things to out-group members, they also motivate individuals to make great sacrifices for the sake of their in-group members.

These motivations are largely instilled in individuals by the normative behavior of the other people in their environment. Within a warlike group, individuals will be pressured by their families, friends, priests, teachers and others to adopt warrior beliefs and values. At the population level, this translates into selection pressure favoring the spread and stability of warrior traits. Likewise, in a social environment containing pacifist norms, individuals will be pressured to adopt pacifist traits, and there will be a selection pressure favoring the spread and
stability of pacifist behavior. Once certain norms have already become prevalent within a given
group, what Richerson and Boyd refer to as “moralistic punishment,” or the enforcement of
norms, will generate selection pressures to sustain their frequency in the population. However,
people can reward or punish just about anything, so the enforcement of norms can select for just
about any kind of behavior whatsoever. As Richerson and Boyd put it (2005, p. 201),

Moralistic punishment can stabilize *any* arbitrary behavior—wearing a tie, being
kind to animals, or eating the brains of dead relatives. Whether the behavior
produces group benefits is of no significance [for individual-level selection]. All
that matters is that when moralistic punishers are common, being punished is
more costly than performing the correct behavior, whatever it might be. When any
behavior can persist at a stable equilibrium, then the fact that cooperation is a
stable equilibrium does not tell us whether it is a likely outcome.

In other words, there is nothing especially cooperative or prosocial about normative behavior in
general—in theory, the norms of a group could function to *inhibit* cooperation among group
members, and to *prevent* tribes from growing very large. In fact, however, norms that increase
cooperation and prosocial motivation have become common universally across many groups.
The point Richerson and Boyd make here is that the connection between normative behavior and
cognition and prosocial motivation is explained by cultural selection among groups, rather than
among individuals within the same group. Selection between groups favors norms that
effectively motivate individuals to pay costs for the sake of cooperation. When individuals are
motivated to invest more heavily in cooperation, their groups reap greater benefits from
cooperation, and these benefits allow them to outcompete other groups. Thus, group-level selection not only favors norms for cooperation, but, more specifically, norms for costly cooperation—prosociality and altruism. One good way to increase prosocial motivation among individuals is by adopting norms that reward people for prosocial actions and punish them for selfish ones.

According to Boyd and Richerson, then, cultural group selection explains the scale of modern tribes by explaining how the individuals that make them up became so willing to invest in cooperation and collective action. Prosocial norms motivate prosocial behavior, and prosocial behavior spreads by group selection. But all of the traits to which they appeal in this group-level selection process are inherited culturally, by social learning, rather than genetically. So cultural group selection, rather than genetic group selection, explains why prosocial norms have become universal features of human social environments.

In a sense, these appeals to group selection are based on the same fundamental principle that has for decades motivated group-selectionist theories of altruism in biology. Biologists E. O. Wilson and D. S. Wilson (Wilson and Wilson 2007) summarize this principle thus: “Selfishness beats altruism within groups. Altruistic groups beat selfish groups. Everything else is commentary” (p. 345). However, the tribal instincts theory also differs in two crucial ways from traditional group-selectionist accounts of altruism, and both of these ways render the tribal instincts theory more empirically plausible. First, it is an account of prosociality, not altruism, and this relaxes the requirement that the traits in question reduce fitness overall. Prosocial traits can be selected at the group level whether the costs involved are net costs or just ordinary costs. Second, the traits in question are acquired by learning, rather than genetic endowment. This relaxes the requirement, discussed in Section 4.2.1, that group-level variation in phenotypes be
explained by an underlying pattern of group-level variation in genotypes. Instead, the burden of explaining this variation is discharged by appeal to a suite of psychological capacities, which are the tribal instincts to which Boyd and Richerson refer. And it is much more likely that our cultural ancestors possessed these capacities than it is that our genetic ancestors exhibited group-level variation in genes for altruism, and did so for long enough periods of time to affect the slow process of genetic group selection.

4.3.2.2 The Culturally Evolved Environment and Moral and Religious Psychology

According to the tribal instincts theory, cultural group selection explains the composition of tribes, or social environments, while tribal instincts are explained by the genetic selection occurring among the individuals within a given tribe. For the normative behavior of the other people in an individual’s environment not only impacts the beliefs and motivations that individual acquires during the process of acculturation. It also impacts her genetic fitness. As a result, individual-level genetic selection explains why we possess certain psychological capacities for negotiating culturally evolved social environments. As Richerson and Boyd put it, “Cultural evolution created cooperative, symbolically marked groups. Such environments favored the evolution of a suite of new social instincts suited to life in such groups, including a psychology which “expects” life to be structured by moral norms and is designed to learn and internalize such norms” (2005, p. 214). Tribal instincts, then, are genetically inherited traits produced by gene-culture coevolution, the same evolutionary process that gave us the ability to digest dairy.

As we’ve seen, however, once certain norms have already evolved within a group, they also influence the subsequent course of cultural selection, or selection among culturally inherited
traits. So there are two important ways in which the norms in a particular environment impact the psychology of individuals within it. Across many lifetimes, they influence certain genetically inherited traits. But during each individual’s lifetime, they also influence various culturally inherited traits. Both types of impact have many important consequences for social psychology, but I focus here on two traits that are particularly important for the study of morality, religion, and the relationship between them: (1) *intrinsic prosocial motivation*, and (2) the role of religious cognition in activating prosocial motivations.

Foremost among the tribal instincts are capacities for acquiring the norms of one’s own group, so as to avoid the costs of punishment and maximize the benefits of normative reward. This requires cognitive capacities for identifying what the norms are in a local environment, but it also involves motivational capacities for complying with those norms, and, in addition, for punishing and rewarding others. And these motivations must be substantial, because they must often succeed in overcoming other, countervailing motivations. By definition, prosocial norms require sacrifices, and doling out punishments and rewards brings more costs still. But like all animals, we usually possess motivations to *avoid* making sacrifices whenever we can help it. So normative motivation must be strong enough to inhibit and overcome these competing desires.

Sripada and Stich (2006), who provide a detailed model of norm psychology, note that we are sometimes motivated to follow norms for practical, instrumental reasons. For example, we might obey a norm simply to avoid upsetting grandma. But this is merely to show sensitivity to *someone else’s* norms, which one has identified, but not actually internalized or adopted. Genuine normative commitment is characteristically *not* based on such instrumental, pragmatic considerations, as Kant’s distinction between hypothetical and categorical imperatives suggests. For this reason, Sripada and Stich argue that norms characteristically provide *intrinsic*
motivation: “Our claim is that people are disposed to comply with norms even when there is little prospect for instrumental gain, future reciprocation or enhanced reputation, and when the chance of being detected for failing to comply with the norm is very small” (p. 285). To genuinely adopt a norm is to acquire a motivation to do the right thing for no reason other than that it is the right thing.

And motivations of this kind can, indeed, be incredibly strong. Psychologists Atran and Ginges have extensively studied the motivations behind violent religious conflict, and they observe that “humans will kill and die not only to protect their own lives or defend kin and kith, but for an idea—the moral conception they form of themselves, of ‘who we are’” (Atran and Ginges 2012, p. 855). To explain these extreme forms of intrinsic motivation, they follow Tetlock et al. (2000) in appealing to sacred values, or preferences that cannot be traded off against one’s practical interests: “Once a value becomes sacred, distinct decision-making rules are thought to apply. For example, it becomes forbidden to think of sacred values as just another fungible value with a given utility that can be measured along a common scale. One should not, for instance, barter one’s child for money” (Sheikh et al. 2012, p. 111). Refusal to barter one’s own child for money is perhaps not the best example, because that could be explained by intrinsic motivation to protect one’s kin, rather than intrinsic motivation to follow a norm. But the example can easily be modified, since one really shouldn’t barter anyone else’s child for money, either. In this sense, to agree that there is no normatively appropriate answer to the question, “How much does this kid cost?” is to agree that some values are sacred. Sadly, of course, not everyone does agree about this, and in many historical contexts there actually were normatively appropriate answers to this question. Sacred values clearly differ from culture to culture.
There is nothing inherently religious about sacred values, but religious cognition is one common trigger of these motivational responses. Sheikh et al. (2012) found that among Americans and Palestinians alike, people who participated in religious rituals more often were also more likely to treat a wide range of preferences as sacred values. And Tetlock et al. (2000) demonstrated the connection between religious cognition and sacred values in a very interesting way, by having fundamentalist Christians respond to a set of “heretical counterfactuals,” or descriptions of ways in which sacred events might have taken a different course. After responding to a scale measuring fundamentalism, Christians were asked to respond to statements that “apply causal schemata that are routine in everyday life but profoundly controversial when extended to the sacred founders of religious or political movements” (p. 864). One item suggested that “If Jesus had not chosen Judas as one of his 12 disciples, Jesus would not have been betrayed or crucified,” while another stated that “If Mary had given birth to more children after Jesus, she could not be portrayed as the Holy Virgin central to Christian beliefs” (p. 864). Participants were asked both to judge the plausibility of such claims and to evaluate a fictional author who had made the claims in a recent book. They were also asked to make the same judgments about “secular counterfactuals,” which described the same counterfactual events occurring in the lives of ordinary good people. The researchers found that fundamentalists were more likely to reject or deny heretical counterfactuals than secular ones. But more importantly, they also found that fundamentalists’ exhibited a coherent suite of motivational responses that had independently been shown to be characteristic of sacred values: “Fundamentalists were most disgusted by heretical counterfactuals, most prone to penalize those who endorse such propositions, and most pained and strained by such propositions. There was no relationship, however, between fundamentalism and reactions to secular counterfactuals” (p. 865).
Intriguingly, though he makes no mention of any of these scientific studies, Leiter (2013) offers an analysis of religion in a very different context that dovetails nicely with them. In the course of an argument against legal exemptions made on religious grounds, he claims that an essential feature of religious belief is the “categoricity” of its commands:

The categoricity of religious commands accounts for both one of the most admirable and one of the most frightening aspects of religious commitment—namely, the willingness of religiously motivated believers to act in accordance with religious precepts, notwithstanding the costs… Thus we find the devoutly religious among those who were at the forefront of domestic resistance to Nazi oppression in the 1930s, and the injustice of apartheid in South Africa from the 1960s onward and in American in the 1950s and ‘60s. We also, of course, find the devoutly religious among those who bomb abortion clinics and fly airplanes into buildings. These religiously inspired individuals risk (and often suffer) death, injury, and prison in order to comply with their religious conscience. (p. 36)

This concept of categoricity thus captures well the relationship between religious cognition and intrinsic, prosocial motivation. Indeed, while most of the behaviors motivated by religious cognition are merely prosocial, some are downright altruistic.

According to the tribal instincts theory, social environments that evolved by cultural group selection may affect genetically inherited psychological traits of many kinds. That is, the hypothesis is that cultural group selection leads to gene-culture coevolution, in a way that impacts many features of social psychology. In order to identify facts about morality and religion
that are not explained by the narrow framework, however, I have focused specifically on traits of normative motivation, and their relation to religious cognition. While the particular norms inherited by an individual are culturally inherited traits, Richerson and Boyd claim that capacities for norm acquisition in general are inherited genetically, much as capacities for language acquisition are. It is thus important to distinguish capacities for learning from the capacities that are learned. Children are born with capacities for learning some language or other, but capacities for speaking English, in particular, are the result of social learning. In the same way, the tribal instincts theory claims, children are born with capacities for learning some set of norms or other. But the capacity to see specifically that selling children for money is wrong, and in a deep sense that may be characterized as a sacred value, is a culturally inherited trait.

4.4 Narrow Adaptationist Theories of Morality and Religion

While they do explain certain forms of prosocial behavior, narrow adaptationist theories of morality and religion do not focus primarily on costly forms of cooperation. Instead, they focus on mutualism, or cooperation that is mutually beneficial for the interactants involved. They do this because they assume that a certain form of mutualistic cooperation also counts as a specifically moral form of cooperation. More specifically, Baumard, Andre and Sperber (2013) present an account of the psychology of fairness, which they take to be based on intuitions and judgments about proportionality between the costs paid and benefits received from cooperative interactions. Baumard and Boyer (2013) then apply this account of the evolution of moral psychology to a complimentary account of the cultural evolution of religion, observing that norms of fairness are important features of many of the most successful religious systems. They
take as their primary explanandum the spread of “moralizing religions,” and they explain the prevalence of moralizing religions by appealing to universal, genetically inherited intuitions about proportionality, making no mention of cultural selection, or of traits inherited by learning. Accordingly, I refer to these as “proportionality theories” of morality and religion. They represent the most recent, sophisticated, and fully developed accounts of moral and religious psychology provided by narrow adaptationists.

Proportionality theories show that genetic selection can explain more about moral and religious psychology than one might initially suppose. That is, they dispel two assumptions about the scope of the narrow framework that may be naive, but that are nevertheless common. The first we have already encountered: the assumption that genetic selection only explains self-interested behavior. The normative commitments of morality and religion typically require the regulation and suppression of self interest, but appeals to genetic selection typically focus on the promotion of self interest. So it may seem that the narrow framework has little to say about the sacrifices that morality and religion characteristically require. Proportionality theories, however, are based on the same assumptions that gave rise to the original puzzle of altruism, and the theories in question here are based on the same solution to that puzzle that Trivers proposed four decades ago. The adaptive value of fairness intuitions, according to proportionality theories, is due to the role such intuitions play in extracting profits from mutually beneficial relationships, and this frequently requires subjugating short-term interests for the sake of long-term gains. Thus, proportionality theories are well equipped to take the altruism out of an important class of (apparently) altruistic behaviors—those in which behaving fairly is costly in the short term, but beneficial in the long term. In other words, they explain why fair ways of behaving are often prosocial, but not actually altruistic.
The second common assumption dispelled by these theories is that genetic selection only explains universal forms of behavior and cognition. Narrow adaptationists have argued for many years that patterns of cross-cultural variation can easily be explained by appeal to genetic adaptations, without appeal to traits inherited in a specifically cultural way. But since the notion of cultural inheritance is necessary for explaining cultural evolution in terms of Darwinian selection, this means that it is possible to explain cultural evolution—changes in the relative frequencies of cultural traits—without appealing to cultural selection. And it is possible to do so not only for universal forms of behavior and cognition, but also for forms of behavior and cognition that vary from culture to culture.

In the rest of this section, I first describe the general strategy narrow adaptationists have employed for explaining how universal, genetically inherited traits give rise to phenotypic differences across cultures. I then explain how proportionality theories employ this strategy in the domains of moral and religious psychology. This enables them to explain how genetically inherited traits produce prosocial cooperation without appealing to group selection, while also enabling them to explain cultural evolutionary outcomes without appealing to cultural selection. My aim here is to show how much proportionality theories do explain, before going on to argue that, nevertheless, this is not enough.

4.4.1 Evoked Culture: How Genetic Similarities Produce Cultural Differences

Consider two different ways in which a belief may spread through a population, or group: through convergent individual learning, or through social learning. We do not learn from others that the sun rises every morning, even if we do learn from them to use words like “sun” to describe what we see every morning. Rather, individuals just see this for themselves, and they all
form the same belief because they are observing the same sun using the same visual systems. These similarities in the functioning of individuals’ visual systems are the results of shared genes shaped by a shared history of genetic selection. In this mundane way, genetic selection constrains the content of individual learning in ways that cause individuals’ beliefs to converge. So the belief that the sun rises every morning occurs with very high frequency in human populations, but social learning has nothing to do with this.

Other shared beliefs, by contrast, are the results of social learning, not convergent individual learning. Virtually all educated people today believe that the earth orbits the sun, but this is not because each of them has independently made the relevant observations and inferences. Rather, educated people today share this belief precisely because they have been educated—they have learned from other people that the earth orbits the sun. That is social learning, or cultural inheritance.

The distinction between these two ways of explaining shared belief applies not only to beliefs, but also to norms, skills, practices, motivations and many other forms of psychological activity. And, of course, it illustrates the difference between similarities among individuals that are due to cultural inheritance, on one hand, and similarities due instead to genetic inheritance, on the other. While broad and narrow theorists agree that genetic inheritance is important, because it enables selection-based explanations of population-level evolution, narrow adaptationists deny that cultural inheritance and cultural selection are also important. More than twenty years ago, Tooby and Cosmides began an article entitled *Evolutionary Psychology and the Generation of Culture* by claiming, “The most productive application of evolutionary biology to the question of culture will not ultimately lie in framing cultural change as an evolutionary process parallel to organic evolution” (Tooby and Cosmides 1989, p. 29). More recently, Pinker
(2010) has claimed that, “Given the undeniable practical advantages of reasoning, cooperation, and communication, it seems superfluous, when explaining the evolution of human mental mechanisms, to assign a primary role to macromutations, exaptation, runaway sexual selection, group selection, memetics, complexity theory, cultural evolution (other than what we call “history”), or gene–culture coevolution (other than the commonplace that the products of an organism’s behavior are part of its selective environment)” (p. 8996). Of the eight “radical revisions” to evolutionary theory that Pinker lists here, three are processes to which broad adaptationists appeal: group selection, cultural evolution (by which Pinker means cultural selection, specifically), and gene-culture coevolution. As we’ve seen, however, all three of these alternative processes ultimately rest on the same fundamental methodological assumption: that psychological similarities produced by social learning may usefully be treated as a mechanism of inheritance supporting a distinct, non-genetic form of selection process. Like Pinker, Tooby and Cosmides, the proponents of proportionality theories of morality and religion reject this methodology. They present their theories as accounts of morality and religion that compete with broad adaptationist theories, and they make no mention of specifically cultural selection processes capable of interacting causally with specifically genetic selection processes.

In the article just mentioned, Tooby and Cosmides introduce the concept of evoked culture to explain how particular beliefs, skills, practices, etc. can spread through populations—and vary between populations—without appealing to cultural inheritance. This concept emphasizes the role of the environment in explaining patterns of similarity and difference among the psychological traits of individuals. On this view, cultural variation can be explained by the fact that different environments evoke different responses from individuals who share the same genetic adaptations. Just as similarity in beliefs about the sun result, in part, from the fact that
genetically similar individuals are faced with the *same* environmental stimulus, *differences* in belief likewise result from the fact that genetically similar individuals are faced with *different* environmental stimuli. For example, we do not need to appeal to social learning to explain why groups located in the tropics do not build igloos, while groups located in the arctic do not take siesta. People do not learn from other people that the weather is hottest at mid-day, or that the hottest part of the day in the arctic still isn’t hot enough to warrant staying inside. And people from the tropics would probably end up making igloos themselves if they found themselves in the arctic, where building materials other than ice are hard to find. On this view, patterns of group-level variation are simply *evoked* by the environment, because the same genetic adaptations manifest themselves in different specific forms of adaptive behavior.

The concept of evoked culture is central to the narrow framework, because cultural variation must be explained somehow, and without a specifically cultural form of inheritance there can be no specifically cultural form of selection. In theories based on evoked culture, the traits that spread through human populations do not spread by social learning, and thus they do not spread by inheritance. On this view, people do not *inherit* psychological traits *from* other people. Rather, people who are already psychologically similar, because of shared genes, happen to *re-create* the beliefs, skills and practices of their fellow group members. Two genetically similar individuals living on distant continents will form the same belief about the daily rising of the sun, but not because one of them has inherited that belief from the other one. These similarities are simply evoked by environmental similarities, so there is no appeal to cultural inheritance, and thus no appeal to cultural selection. And since different environments can evoke different responses in the same way that similar environments can evoke similar responses,
universal genetic adaptations can be used to explain patterns of variation across cultural groups in the same way that they explain patterns of similarity across groups.

4.4.2 Proportionality Theories of Religion and Morality

In his book *Religion Explained* (2001), Boyer develops a detailed and extensive account of religious evolution based on evoked culture. He argues that the same universal adaptations are responsible both for forming ordinary concepts about natural agents and for forming religious concepts about supernatural agents. If so, this account can not only explain why supernatural agent concepts appear almost universally among religious systems, but it can also account for the different supernatural agent concepts that appear in different religious systems. On this view, different specific environmental stimuli evoke different specific beliefs about different supernatural agents. But they do so because of functional similarities that are explained by genetic selection.

Boyer describes a specific mechanism explaining how religious beliefs are evoked. Innate and universal patterns of memory and inference, he argues, favor the spread of *minimally counterintuitive* (MCI) beliefs. Religious representations of all kinds violate our predictive intuitions, he suggests, because these intuitions are formed from a set of innate and universal templates—abstract and fundamental conceptual categories that structure individual learning, shaping the acquisition of more specific and concrete concepts. For example, the template ANIMAL structures the acquisition of a large number of more specific concepts, such as *dog*, *cow* and *iguana*. And it is because of general features of the ANIMAL template, which are shared among all humans, that all humans who learn about iguanas share the intuition that iguanas don’t talk. If this intuition is then violated by some representation of a supernatural
iguana that does talk, then that representation will strike people as particularly interesting and memorable, leading them to think about the talking iguana concept more frequently than other representations about iguanas, and causing them to remember this idea better. As a result, Boyer claims, talking-iguana concepts are likely to spread through populations more easily than other iguana concepts, which are less interesting. Of course, if counterintuitive concepts are too counterintuitive, then they will simply be incoherent, difficult to process, and difficult to remember. This hinders their ability to spread through populations, rather than facilitating it. Hence, Boyer concludes, religious ideas spread when they are counterintuitive, but only minimally so.

On this account, what determines which concepts are minimally counterintuitive are the features of our conceptual templates, which are genetically inherited traits. Talking-iguana representations evoke the same responses across individuals because genetic selection has endowed each of those individuals with the same cognitive adaptations. As Boyer puts it, “Templates are one of the devices that allow minds to reach similar representations without having a perfect channel to ‘download’ information from one mind to another” (2001, p. 45). Here he is adopting Sperber’s model of culture as an “epidemiology of representations.” This model uses the evoked culture approach to avoid treating psychological similarities between individuals as heritable traits. Instead, Sperber argues, representations spread through cultures by evoking the same “symptoms” across individuals, just as a virus evokes the same symptoms of response across individuals. As he puts it, “Your understanding of what I am saying is not a reproduction in your mind of my thoughts, but the construction of thoughts of your own which are more or less closely related to mine” (1996, p. 58).12

---

12 By distancing themselves from the “reproduction” or “downloading” of thoughts across individuals, Sperber and Boyer mean to distance themselves from meme-based approaches to cultural evolution,
Interestingly, Boyer’s early work on religious cognition had little to say about the social aspects of religion, or about the relationship between religion and morality. To address this gap, he has recently teamed up with Baumard to develop an account of moralizing religions, or religious systems that emerged only recently, and only in certain large-scale societies, and that “connected beliefs about imagined superhuman agents with highly specific moral prescriptions” (Baumard and Boyer 2013, p. 272). They note that “[religious] traditions from very different places describe morality in highly similar terms. In particular, they all include prescriptions based on the same principle of proportionality” (p. 272). To illustrate, they point to the explicit formulations of the Golden Rule that appear in Buddhism, Confucianism, Hinduism, Taoism, Judaism, Christianity and Islam. They also note the role of proportionality intuitions in fitting punishments to transgressions, whether those punishments occur in this life or in the afterlife.

They do not mention templates in this account, but intuitions about proportionality play the same role in explaining the spread of moralizing religions that templates play in explaining the spread of minimally counterintuitive concepts. As they put it, “evolved moral intuitions provide a background against which certain types of information (in this case from moral and religious doctrines) are more easily communicated than others. Because moral intuitions often presuppose proportionality, explicit statements advocating proportionality seem more compelling and natural than possible alternatives, leading to the cultural spread of such doctrines” (p. 275). In other words, moralizing religions all evoke the same intuitive responses, based on the same genetic adaptations, causing these religious systems to spread.
This proportionality account of religion is founded upon a prior proportionality account of morality, which Baumard has developed along with Sperber and Andre. This account uses models of *partner choice* to explain the evolution of intuitions about proportionality and fairness, and partner choice models are based on the same adaptive logic as Trivers’ theory of reciprocal altruism: mutually beneficial interaction. Perhaps the most obvious examples of mutualism in humans involve ordinary economic exchanges, which are governed by norms according to which the value of goods or services received should be proportional to the value of the costs paid. An individual who cheats another in a particular interaction may gain unfair benefits in the short-term, but in doing so she also risks alienating the partner, who has been forced to pay the costs of cooperation without receiving the benefits. Thus, the short-term costs of fair dealing in a particular interaction are balanced out by the long-term benefits of the reciprocal relationship in general.

Mutualism, however, need not take the specific form of reciprocity, because many forms of mutualism do not depend on multiple interactions that are distributed across time. An ecologically relevant example of mutualism would be the hunting of big game. This example also illustrates the difference between proportionality and equality in the distribution of benefits. A fair distribution of benefits does not necessarily imply an equal distribution. Rather, it implies that benefits be proportional to costs, so that, in the end, each individual’s benefits are proportional to her costs. Selection for mutually beneficial interactions will, these researchers suggest, produce cognitive adaptations for identifying and comparing the costs and benefits accruing to the various participants, generating intuitions about proportionality that are manifested in norms of fairness.
Anyone who ends up with a disproportionately small share of the benefits in a given interaction is likely to take measures to ensure that future interactions go better for her. One way of doing this is by breaking with her partners, and choosing others instead. Partner choice models explain mutualistic cooperation by assuming that some partners are better than others, and that individuals have some say about who their cooperative partners will be. When the cooperative task requires some costs—as it almost always does—opportunities for cheating arise, since individuals may find ways to free ride on the behavior of others, skimping on the costs of collective action while still receiving the benefits. Such situations are called collective action problems, and they are typically modeled in economic game theory by the payoff structure of the Prisoner’s Dilemma.\(^\text{13}\) In the Prisoner's Dilemma, each participant faces a decision between cooperating and defecting, and the costs of either choice depend on what the other participants choose. When everyone cooperates, everyone enjoys substantial benefits, because the costs of cooperation are appropriately distributed across interactants. But when one person’s decision to cooperate is \textit{not} matched by her partners, the costs are not distributed appropriately. The cooperators pay too heavily for the benefits that are generated, suffering net losses. On the other hand, if everyone defects, then no cooperation occurs at all. No one pays the costs, but no benefits are received either. So while mutual defectors do not do as well as mutual cooperators, they at least come out ahead of cooperators who are the victims of others’ defection, since they at least do not suffer net losses. And when one person’s decision to defect is \textit{not} matched by her partners, she obtains the highest payoffs of all—even higher than cooperators who band together. She enjoys a free ride, obtaining the benefits of everyone else’s cooperative activity without

\(^{13}\)Skyrms (2004) and Pacheco et al. (2009) point out that many interesting forms of collective action are better modeled by the payoff structure of the Stag Hunt, rather than the Prisoner’s Dilemma. The Stag Hunt, however, describes situations in which cooperating is not costly, relative to the alternative option of defecting. Since everyone studying morality and religion wants to explain prosociality, or costly forms of cooperation, the Prisoner’s Dilemma is the appropriate model.
having to pay the costs herself. Since this payoff structure renders participants vulnerable to exploitation by free riders, it is important for individuals to choose reliable cooperators as partners, rather than defectors.

The need to identify good partners, then, creates what is called a *biological market*, in which individuals compete with one another to be chosen by others as cooperative partners. When people are free to choose their partners, victims of defection can “punish” those who cheat them by simply taking their business elsewhere, or by refusing to choose the defector again as a partner in the future. Thus, just as Coca Cola and Pepsi compete with one another to be chosen by customers as partners in mutually beneficial exchanges, “cleaner” fish, who eat the parasites off of other fish, compete with one another to be chosen by “client” fish. Cleaners can cheat in this otherwise mutualistic arrangement, by eating their clients' useful mucus, rather than their parasites. But this causes clients to avoid defecting, mucus-eating cleaners, and to preferentially approach cleaners who are more reliable cooperators. Indeed, Bshary and Grutter (2006) show experimentally that client fish engage in “image scoring,” observing the interactions that cleaner fish have with other clients, and using this information to choose their partners wisely. By analogy with an economic market, this is the equivalent of a customer’s observing the interactions of the Pepsi Company with its customers before deciding whether to buy Pepsi.

Of course, there are plenty of differences between the biological markets of humans and those of other species, but one difference is of special importance for present purposes: humans possess *third-party* reputations. Unlike fish, our linguistic abilities enable us to broadcast the image-scoring information we acquire, sharing with others the information we've gathered ourselves about individuals' reputations for cooperating reliably. This helps cooperators to find each other within a sea of defectors, since individuals may avoid defectors even when they’ve
never interacted with them before, and even when they’ve never observed their interactions with others.

The consequences of this for defectors may be dramatic. As we’ve seen, the emergence of large groups is a recent occurrence in evolutionary history, and gossip can travel fast in small groups. As Baumard et al. note (p. 66), “A mutualistic social life is dense with information relevant to everyone’s reputation.” As a result, in a worst-case scenario an egregious act of cheating could become juicy gossip, and its perpetrator could gain a reputation as a defector among all of his potential cooperative partners within a matter of days, or even hours. While human cheaters are often given second chances, this worst-case scenario shows that reputations make partner choice an especially powerful mechanism for stabilizing mutualism in humans. The costs of gaining a bad reputation are simply too severe to make cheating profitable.

As a result, the adaptive value of the target traits explained in proportionality theories is based on the role proportionality intuitions play in managing an individual’s reputation. These traits are assumed to be features of moral psychology, specifically, because they are responsible for motivations to treat others in a fair, impartial manner, even when doing so is costly. And they are assumed to be features of religious psychology, specifically, because of the role they play in explaining the spread of some of the world’s most prevalent religious systems. On this view, moralizing religions evoke universal, genetically inherited intuitions of proportionality.

Where environments are similar, these traits result in similarities across cultural groups. But differences across groups can also be accommodated, because environmental differences are predicted to evoke different intuitions about what counts as a proportional, or fair, interaction. Thus, this is a selectionist account of genetic evolution, since it explains the functions of genetically inherited psychological traits by appeal to a history of genetic selection. But it is a
non-selectionist account of cultural evolution, because it explains the spread of religious ideas by appealing to evoked culture, rather than by appeal to selection among traits inherited in a specifically cultural manner. Because they are based on genetic selection, proportionality theories do not explain the costliness of behaving fairly as a form of genuine altruism. But they do explain why fairness often requires costly, prosocial behavior. Partner choice within biological markets shows that individuals must pay the short-term costs involved in behaving fairly, because the alternative is to face the even higher long-term costs of having no one to cooperate with.

4.5 How Are Broad and Narrow Theories Related?

I argued above that because the broad framework simply adds appeals to cultural selection onto a theoretical framework already grounded in genetic selection, the broad framework can include or incorporate theories of morality and religion derived from the narrow framework. As I have defined them, the broad and narrow frameworks conflict, as methodologies. But this does not mean that the particular theories derived from these frameworks conflict. Broad adaptationists have no reason to deny, for example, that kin altruism explains the enormous sacrifices that human parents make for their children, or that much of human cooperation is based on family ties. Likewise, broad theorists assume that mutualistic cooperation is important in humans for the same reason that it is important in many other species. Indeed, they have no reason to deny even the claim that mutualistic cooperation is more important in humans than in any other species.

Speaking for a team of broad adaptationists (including Boyd and Richerson) who have developed an updated form of the tribal instincts theory, Gintis (2013) states explicitly that this
theory and the proportionality theory of morality are “in broad agreement.” This broad theory is called the Beliefs, Preferences and Constraints (BPC) model, and Gintis specifies that, “all of the human behaviors affirmed by [Baumard et al.] fit nicely into the BPC model, and are in no way in conflict with our stress on altruistic cooperation and punishment” (p. 91).

Baumard et al. disagree, however. They respond directly to Gintis’s claim, arguing that “the stress on altruistic cooperation and punishment that Gintis mentions implies…giving the main role in the evolution of cooperation and morality to group-level selection, and, on this, we beg to disagree… In our framework, group selection is not necessary to explain the existence of human morality” (p. 105). This objection reveals some significant assumptions these theorists are making about the scope of their theory. (Also, note that what both theorists mean by “altruistic” cooperation” here is what I have defined as prosocial cooperation.)

Baumard et al. believe their theory explains “the existence of human morality,” and that it does so by giving the “main role” to individual-level selection, rather than to group selection. And this, they believe, entails that their narrow account of morality is in conflict with the BPC model. But for any two empirical theories to conflict, they must provide different explanations for the same phenomena, the same empirical facts. Baumard et al. evidently assume that the two theories in question here do conflict in this way, in virtue of the fact that both purport to explain “the existence of human morality.” While their theory does this by giving the main role to individual-level selection, they suppose, the broad theory does it by giving the main role to group selection instead. This assumption would only be safe, however, if the domain of human morality constituted a particular set of facts. It doesn’t.

Baumard et al. are correct to point out that in their framework, group selection is not necessary to explain morality. But this is only because in their framework, there is nothing more
to morality than fairness. The empirical evidence they rely on explains only those features of moral psychology that are connected to intuitions about proportionality, which are then explained by appeal to the adaptive value of mutualistic interaction. I see nothing wrong with the assumption that proportionality intuitions map onto what we ordinarily and intuitively think of as fairness, and I also see nothing wrong with the assumption that concerns about fairness are moral concerns. Thus, I agree that the target facts their theory explains fall within the moral domain, and in this sense they are justified in claiming that that their account explains the existence of human morality. But that is just to say that they explain one important set of facts about morality. This provides no reason to think that broad theories do not explain other important facts about morality.

Like religion, morality is an exceedingly complex, multifaceted natural phenomenon, which can only be studied by analyzing it into manageable pieces. Scientific theories of moral psychology proceed by analyzing the moral domain into more specific sets of facts or traits, which they then take as their target explananda. Baumard et al. have chosen judgments and decisions about fairness as their explananda, carving out a manageable but important sub-domain to explain in detail. But broad adaptationists like Gintis, Boyd and Richerson have done the same thing, explaining the prevalence of prosocial norms and prosocial motivations in general—and not only in the sub-domain of fairness.

In order to provide a charitable interpretation of the claim that the proportionality theory conflicts with the BPC model, I will grant for the sake of argument that fairness is the most important sub-domain within the larger moral domain. Baumard et al. claim that individual-level selection, rather than group selection, played “the main role in the evolution of cooperation and morality,” and I take it that only one theory can play the main role here. Thus, they appear to be
assuming that fairness constitutes the main set of explananda—the core of the moral domain, so to speak. Since there is no generally accepted definition of the concept of morality to which one can appeal in policing such claims, I don’t know how such a claim could be defended beyond mere appeal to intuition, and I don’t find it intuitive myself. But the alternative interpretation of this claim seems to me less charitable. This would be that the sub-domain of fairness actually exhausts the moral domain—that, in fact, there is nothing more to morality than fairness. This strikes me as an extreme and counterintuitive view of the moral domain, and I don’t think Baumard et al. must be committed to it, so I will attribute to them instead the less extreme, more intuitive assumption that fairness comprises the core of the moral domain.

Note, however, that it would not harm my argument even to grant the more extreme view. For if it turned out that phenomena falling outside the domain of fairness did not count as facts about morality, this would not make those facts disappear. It would not show, for example, that prosocial norms outside the domain of fairness are not explained by cultural group selection, or that they can be explained away, or that they do not stand in need of explanation at all. These facts need to be explained whether they can be called “moral” or not. Parallel remarks apply to the explananda of the religious domain.

Moreover, proportionality theories do not even explain everything that needs to be explained within the sub-domain of fairness. This is because, as I argue in the next section, evoked culture theories in general do not actually explain cross-cultural variation. They merely accommodate it, providing explanations for cultural similarities that do not conflict with facts about cultural differences. The next section, then, is devoted to identifying two important types of moral and religious explananda that are not explained by proportionality theories: (1) those that reside outside the boundaries of the sub-domain of fairness, and (2) those that reside inside
this sub-domain, but that differ across groups because of the role social learning plays in structuring individuals’ intuitions about fairness.

Since the broad framework shares with the narrow framework the assumption that facts about individual-level genetic selection are fundamentally important, it has no difficulty accepting that genetic selection explains facts that lie at the core of the moral domain. Even if there were a significant sense in which individual-level genetic selection explains the most important facts about morality and religion, it would still be necessary to explain a host of other important facts that genetic selection alone does not explain. It is in this sense that the broad framework includes narrow theories. Because the broad framework explains the evolution of traits inherited both genetically and culturally, while the narrow framework explains only traits inherited genetically, the broad framework explains everything the narrow framework explains and much more.

4.6 What Needs to be Explained about Morality and Religion?

In the last section I identified two different ways in which the proportionality theory of morality, and the accompanying theory of religion, are limited in scope. First, they only explain the psychology of fairness, and there is much more to moral and religious psychology than fairness. Second, however, proportionality theories are limited even within the domain of fairness, because they only explain traits that are inherited genetically.

The first type of limitation is specific to proportionality theories. Because they only explain psychological facts related to intuitions about proportionality, these theories only explain one sub-domain of the larger domain of morality in general. Consequently, they also explain only a
limited set of facts about the cultural evolution of moralizing religions. I take these limitations to provide grounds for objecting, not to these theories themselves, but to the theoretical framework from which they are derived. I see it as no objection to a theory that it has limits. All theories have limits, and no selectionist theory of any kind can explain why lightning causes thunder, or why weather is warmer at lower latitudes. So it is important to keep in mind that I am objecting to a methodology—the narrow framework—and not to particular theories derived from this methodology. My argument is based on scope, not empirical adequacy, so I am assuming for the sake of argument that proportionality theories are straightforwardly true.

I have chosen to focus on proportionality theories because they represent the most fully developed attempts by narrow adaptationists to explain morality and religion, and the best that framework has to offer with regard to explanations of prosocial norms and normative behavior. It is thus worth recalling here, briefly, how much they do explain. For while they can only explain genuine altruism by identifying systematic forms of malfunction, or systematic mismatches between current environments and the environments of the evolutionary past, proportionality theories can say much more about prosocial forms of cooperation that are not actually altruistic. Collective action problems, as defined by the Prisoner’s Dilemma, are a fundamentally important feature of human cooperation and human social organization. While proportionality theories do not explain everything about collective action, it is important that punishment by partner choice, and competition in biological markets, can motivate individuals to pay out substantial costs for the sake of collective action. And like Richerson and Boyd, proportionality theorists claim that the normative, prosocial motivations they identify are intrinsic, rather than instrumental. Within the domain of fairness, then, proportionality accounts really do stand to make an important contribution.
The second type of limitation, however, does not apply specifically to proportionality theories. Theories based on evoked culture in general explain only the genetically inherited traits of individuals, and this is a limitation that applies not just within the sub-domain of fairness, but to all sub-domains within the moral domain, as well as to the domain of religion. Once we have identified the role of social learning and cultural selection in explaining facts about fairness, it becomes apparent that even if narrow adaptationists go on to devise additional theories outside the domain of fairness, then these theories, too, will only be able to provide partial accounts of the facts in their target domains.

4.6.1 Mapping the Moral Domain

Many aspects of moral psychology have nothing to do with fairness, and nothing to do with mutually beneficial interactions. Most strikingly, this includes beliefs, norms and motivations regarding causing harm to others. If a mugger holds someone up in a dark alley, he is not treating his victim unfairly, and he is not cheating on a mutually beneficial exchange. For his victim clearly would not benefit no matter how the interaction turned out. So while it may be true that virtually all violations of fairness also cause some kind of harm, norms against harming others do not, in general, have anything to do with fairness, proportionality, or mutualism.

It might seem that this claim can be resisted by appealing to the Golden Rule. Perhaps what makes it wrong for a mugger to harm others is that it would be wrong for him to be mugged himself. But while this would explain how intuitions about proportionality support normative judgments against harming others, it would do so in a way that separates the psychological role attributed to them from the evolutionary account on which proportionality theories are based. Baumard et al. are able to explain the evolution of proportionality intuitions only in contexts of
mutually beneficial cooperation, because it is only in those contexts that long-term benefits are said to outweigh the short-term costs of fair, prosocial behavior. Mugging is not a form of mutually beneficial interaction, and nor are a wide variety of interactions in which one individual might have reason to harm some other individual. Mugging is predation, not cooperation, and the benefits to the mugger do not depend on some long-term reciprocal relationship with the victim. From an adaptationist point of view, muggers have no reason to do unto others as they would have others do unto them, for the same reason that predators in general have no reason to do unto their prey as they would have their prey do unto them. If intuitions about proportionality are used to explain harm outside of the context of mutualist cooperation, then we no longer have an adaptationist explanation for those intuitions themselves.

Moreover, since Baumard and Boyer (2013) cite the appearance of the Golden Rule in the Bible as evidence that Christianity is a moralizing religion, it is important to point out that, in the domain of harm, the Bible also advises followers not to treat others in a fair, proportional manner. To “turn the other cheek” and to “give someone the shirt off your back” are both common sayings taken from a passage (Luke 6:29) in which Jesus advises followers not to respond to others in a manner that is proportional to the manner in which they have been treated. In less biblical terms, the same point is captured by the saying, “Two wrongs don’t make a right.” The second wrong might be fair, but it still isn’t right. These normative views demonstrate that even if norms of fairness produce certain forms of prosocial behavior, other prosocial norms have nothing to do with fairness. And, given the particular evolutionary explanation offered by proportionality theories, this is obviously because many forms of interaction have nothing to do with mutually beneficial cooperation.
Further, proscriptions of harm are not the only sorts of moral and religious activities that proportionality theories cannot explain. Note that \textit{none} of the Ten Commandments addresses fairness (and nor do any of the Five Pillars of Islam). The first few commandments are more religious than moral, specifying what one’s attitude toward God himself ought to be. Most of the others then focus on harm, including murder, bearing false witness against others, and stealing and coveting others’ property (including wives). Then, there is one specifying that children should honor their parents, and another against adultery. This last, of course, represents just the tip of an iceberg of norms governing sexual behavior. Prostitution is a form of mutually beneficial exchange, but questions of fairness do not explain moral attitudes and religious injunctions against it. Rather, prostitution has traditionally been considered wrong because it is a specific form of fornication. And general norms against fornication and adultery shade off into an ocean of more specific sexual norms specifying what people should wear, whom they should interact with, and what forms these interactions should take. Modern debates about homosexuality and same-sex marriage, for example, are neither about harm nor fairness.

I assume, on the basis of intuition, that most norms about harm, sex and fairness count as moral norms. But they are only among the most obvious major categories of moral cognition and behavior. Haidt (2012) maps the moral domain according what he calls “moral foundations,” which are basic matters of normative concern found universally across cultures, classes, religious groups and political orientations. On the basis of a wealth of empirical data, he identifies six of these foundations, each of which represents a distinct sub-domain within the moral domain. Only one of them, the Fairness/cheating foundation, concerns fairness. A second is the Harm/Care foundation. Attitudes toward sexual behavior fall primarily under the Sanctity/Degradation foundations, and the other three foundations are Liberty/Oppression, Loyalty/Betrayal and
Authority/Subversion. So, for example, the political debate over abortion in the United States can be explained largely by the fact that abortion activates different responses in different groups: while political conservatives tend to see the issue in terms of the sanctity of life for the unborn child, liberals tend to see the issue as a matter of liberty for the mother to make her own decisions about her body.

In addition to these moral norms, as we saw in the case of the first few of the Ten Commandments, there are also specifically religious norms that have nothing to do with fairness. In these cases, the intrinsic motivations and normative beliefs that individuals acquire depend specifically on religious cognition. When a Christian possesses the normative belief that it is wrong to work on the Sabbath, she does so on the basis of a prior set of descriptive beliefs, according to which this command was given directly to Moses by God. Likewise, when Ginges et al. (Ginges et al. 2007) find that Palestinian students treat sovereignty over East Jerusalem as a sacred value, which cannot be traded off for monetary compensation in peace accords, this depends upon descriptive beliefs they hold as well. For example, Muhammad is believed to have ascended to heaven to meet with Allah from a site located in East Jerusalem, now the site of one of Islam’s holiest mosques.

Broad adaptationist theories of moral psychology have no difficulty accounting for any of these phenomena that fall outside the scope of fairness. This is because the broad framework provides a general theory of prosocial norms and religious beliefs, according to which any kind of group-beneficial behavior might be selected for. Norms that facilitate mutually beneficial interactions among group members may accomplish this, but so too may norms against adultery and murder, and beliefs about gods and prophets.
4.6.2 Cross-Cultural Variability: Explanation vs. Accommodation

The considerations just mentioned show that large and important sub-domains of the moral domain do not concern fairness, and are not explained by intuitions about proportionality. But this generates two important new issues. First, in arguing against the narrow framework, I have been focusing specifically on the scope of proportionality theories, since I take them to be the most fully developed attempts from narrow adaptationists to explain prosocial cooperation. But to show that *these* theories do not explain certain phenomena is not to show that the narrow framework in general cannot explain these phenomena. Clearly, individual-level genetic selection in general *does* explain other aspects of moral psychology. Second, I have claimed that while the frameworks themselves are in conflict, the particular theories derived from them are not, because these theories explain different target phenomena. Yet Henrich et al. (2010) provide a particular theory from the broad framework that explains norms of fairness, and this appears, on the surface, to provide a counterexample to my claim. I address both of these outstanding issues in this section, by comparing proportionality theories to an alternative account of cultural evolution based on inclusive fitness theory: Shaun Nichols’ account of the evolution of harm norms (2004, Ch. 7).

Nichols’ theory is not strictly based on the narrow framework, but, like proportionality theories, it is an account of cultural evolution based on evoked culture, rather than cultural inheritance. Consequently, it can be used to illustrate how facts outside the domain of fairness might be explained in narrow adaptationist terms, without appealing to cultural group selection or gene-culture coevolution. However, close examination of Nichols’ account reveals that even if the evoked culture approach can explain many important phenomena outside the domain of fairness, it will nevertheless leave other important phenomena unexplained. As a result, the
account can also be used to identify certain limitations of the narrow framework that also apply within the domain of fairness.

Nichols does not reject appeals to cultural selection, but he clearly offers his theory in the spirit of other narrow theories of cultural evolution. He describes it briefly as “an epidemiological account,” alluding to Sperber’s analogy for illustrating the notion of evoked culture, discussed in Section 4.4.2. Moreover, he emphasizes evidence showing that his target traits are inherited genetically (p.156). The traits in question are capacities for emotional sensitivity to the pain and suffering of others, and Nichols emphasizes the affective and motivational aspects of these sensitivities, in particular. His central claim is that “affect-backed” norms will, all else being equal, spread and persist better than other norms that are not “backed” by affect. Because humans are universally prone to experience negative affective responses to the pain and suffering of others, he suggests, norms regulating harmful action “resonate” with our innate intuitions, and this helps them to spread. The upshot is that the scope of harm norms has continually expanded throughout cultural evolutionary history, in a way that looks very much like what other philosophers have considered moral progress. This affective resonance model, then, provides an evoked culture account of the “ubiquity” of harm norms, which Nichols opposes to accounts of harm norms based on moral progress. Thus, affective responses play the same role in this theory that intuitions of proportionality play in proportionality theories, as genetically inherited traits that constrain individual learning and development in ways that cause individuals converge upon the same normative beliefs and motivations. No mention is made of inheritance through social learning, cultural group selection, or gene-culture coevolution.

Importantly, however, Nichols also says nothing to imply that his own theory competes with theories based on cultural selection. And he notes explicitly that his own appeals to evoked
culture provide only a partial explanation for the ubiquity of harm norms (p. 148). In fact, when focusing specifically on questions about how to explain cross-cultural variation in harm norms, he briefly mentions “cultural processes” and “cultural fitness” in a way that might be interpreted as a reference to cultural selection processes. Near the beginning of the essay, Nichols’ notes explicitly (p. 142-43) that any complete account of harm norms must “accommodate” facts about cross-cultural variation, in addition to facts about cultural universality. The following passage, near the end of the essay, is where he explains how his affective resonance model accomplishes this:

It is worth noting that the above explanation of the ubiquity of harm norms is fully consistent with rich diversity in harm norms. For the claim is simply that harm norms will have enhanced cultural fitness. This allows for considerable normative diversity, because it concedes that cultural processes play a vital role in the development of norms. Because cultural processes implicate a complex and variegated set of forces, it is hardly surprising, on this view, that there is so much diversity in the norms found in different cultures. (p. 156-57)

While he never mentions cultural inheritance, Nichols concedes here that “processes” other than the ones he identifies are also important for explaining cultural evolution. Crucially, then, to accommodate facts about cultural diversity, in the way he does here, is not to actually explain them. Nichols observes correctly that variation in harm norms is “fully consistent” with the manner in which his appeals to genetic selection explain the ubiquity of harm norms. But there is
much more to providing an explanation than mere consistency, and Nichols says nothing further about how to do this for cross-cultural diversity among harm norms.

This brings out an important point about evoked culture accounts in general. As Nichols’ own terminology suggests, it is necessary to distinguish between providing a positive explanation for a given phenomenon and providing an “explanation” that merely accommodates the phenomenon, or is consistent with it. Evoked culture accounts proceed by observing that individuals who are genetically similar may nevertheless behave differently in different environments, and then assuming that individuals from different groups are, in fact, confronted with different environmental stimuli. But this is not to explain the differences in behavior that are evoked, because it simply pushes the explanation back onto the relevant environmental stimuli. Nichols does not attempt to identify the relevant environmental differences, so he does not attempt to provide a positive explanation for cross-cultural differences among harm norms. Neither do Baumard and Boyer, in their account of moralizing religions. The closest they come to this is in their discussion of proportionality between the severity of violations and the severity of the punishments imposed: “Obviously, proportionality (greater punishment for greater violations) all depends on people’s estimates of the respective costs of violations and sanctions. So what is an appropriate sanction to one group may seem unduly severe or lenient to another” (2013, p. 273). Again, this is just to push the explanation back onto differences in the social environment for which no explanation is even attempted. Once we have distinguished between, on one hand, the genetic adaptations that cause different responses to be evoked, and, on the other hand, the environmental stimuli responsible for evoking them, it becomes apparent that theories of evoked culture do not actually explain cross-cultural diversity, even if they do accommodate it. It is to Nichols’ credit that he recognizes this.
Broad theories, of course, do offer a positive explanation for cultural variation among harm norms, and they do so by developing mathematical models designed to make sense of the additional “complex and variegated set of forces” to which Nichols refers. Indeed, because appeals to cultural group selection depend on group-level variations among culturally inherited traits, broad adaptationists have devoted considerable attention to this matter. As we saw in Section 4.2.1, while migration removes group-level variation in genes, rendering genetic group selection insignificant in humans, this problem does not arise for cultural group selection. Because individuals learn from in-group members far more than they learn from foreigners, and because rates of cultural evolution are much higher than rates of genetic evolution, the trajectories of cumulative cultural evolution unfolding in distinct social groups are largely decoupled from one another. It is only a matter of time, then, before group-level variation emerges. And once it does, a number of factors work together to maintain and amplify these differences. Richerson and Boyd mention two of these factors (2005, p. 203-6), but recently Baldini et al. (Baldini et al. 2013) have expanded this list to include four additional factors maintaining group-level variation. These factors jointly explain why empirical investigation reveals far more group-level variation among culturally inherited traits than among genetically inherited traits (Bell, Richerson, and McElreath 2009). All of this well-maintained group-level variation makes appeals to cultural group selection empirically plausible. By contrast, there are no similar mechanisms at work in genetic group selection, counteracting the effects of migration in removing this variation.

One of the factors responsible for maintaining group-level variation among cultural traits is particularly important for understanding the evolution of normative behavior: moralistic punishment, or norm enforcement. People punish and reward the behavior of their fellow group
members in ways that maintain conformity within groups and variation between them, and this applies to harm norms as well as to other kinds of norms. An interesting example comes from an unusual norm instated in the Mongol armies of Genghis Khan. By the Khan’s decree, all members of a squad of ten soldiers were forced, on pain of death, to either fight together or retreat together (Plano Carpini 1928). As a result, if an individual deserted, he would knowingly be condemning the rest of his squad to death at the hands of their own leaders. Desertion was thus prevented not in the usual way, by pitting fear of battle against fear of punishment, but rather by pitting fear of battle against motivations of guilt and friendship (in addition to punishment, presumably). Soldiers in most other groups have not fought in such tightly-knit units, but that is because they were not forced to do so by other members of their own group. The enforcement of this norm thus caused individuals in the Mongol group to behave similarly to one another, but differently from individuals in other groups. In other words, the punishment involved in enforcing this norm had the effect of maintaining group-level, cross-cultural variation.

Obviously, genetic differences between the Mongols and members of other groups are not responsible for this form of group-level variation. But it is equally obvious that these differences were not evoked by features of the physical environment. The physical environment might well explain other special features of the Mongol group, such as the fact that they fought on horseback, instead of on foot. The geography of the Asian steppe lends itself much more to nomadic pastoralism than to settled agriculture, and since horses facilitate the nomadic lifestyle, universal genetic adaptations gave the Mongols the same reasons that anyone else would have had to ride horses. But the special form of squad organization exhibited by the Mongol army is not plausibly viewed as a special behavior evoked by a special physical environment. If this
behavior is to be explained as a universal response to a special environment, then it is a special social environment, not a special physical environment. More specifically, the relevant environmental features are the norms laid down by Genghis Khan, and adopted by everyone else through social learning. Mongol leaders learned from Genghis Khan that they ought to enforce this special form of squad organization, and Mongol soldiers learned from their leaders what forms of behavior counted as complying with the norms.

Accordingly, broad adaptationists can actually explain, at least in part, the group-level variations in fighting behavior between the Mongols and other groups. This is because they can explain not just the relevant genetic adaptations, but also the relevant differences in the social environment. This is not to say that the broad framework explains everything about the relevant social environment. From the perspective of cultural selection theory, Genghis Khan’s invention of a new form of squad organization is a cultural mutation, and just as the theory of genetic selection does not explain where mutations come from in the first place, the theory of cultural selection also does not explain this. What it does explain is what happens after the mutation occurs: it spreads quickly throughout the entire Mongol army, where it is enforced, but it does not spread to other groups. As a result, Mongol soldiers behave differently on the battlefield than members of other groups do. They fight more cohesively. Narrow adaptationists cannot explain this difference between the fighting behavior of the Mongols and the fighting behavior of other groups, because they cannot explain the normative differences—the differences in the social environment—that give rise to these differences in fighting behavior.

What Nichols’ concession illustrates is that to accommodate cross-cultural variation is simply to assume the existence of some kind of environmental variability. Some of this environmental variability falls outside the scope of the broad framework as well as the narrow
But many other environmental differences do fall within the scope of the broad framework, and outside the scope of the narrow one. And these include important differences between the religious beliefs and moral norms that are characteristic of particular social groups, and that individuals inherit by social learning. Like other theories of evoked culture, proportionality theories proceed by identifying genetically inherited traits that are universal, and thus do not differ across individuals from different groups. This is clearly consistent with group-level variation in norms and judgments about fairness, but it does not explain such variation. Baumard, Andre and Sperber (2013) explicitly address some well-known studies conducted by broad adaptationists, which use game-theoretic experiments to measure cultural differences in fairness judgments. These studies demonstrate clear patterns of cross-cultural variation that can easily be explained by cultural selection (Henrich et al. 2006; Henrich et al. 2010). But when Baumard et al. discuss these patterns (p. 70-71), their discussion shows only that they can accommodate them. They provide no more explanation of cultural variability in the domain of fairness than Nichols’ account does in the domain of harm.

Baumard et al. correctly point out that individuals from different cultures are interpreting the experimental situation in different ways. That is, individuals import the normative assumptions of their distinct cultures into the experimental situation with them, causing the same experimental stimuli to evoke systematically different responses in individuals from different groups. Thus, the observed differences in fairness judgments are explained by appeal to prior, independent differences among the norms of their respective groups. But the proportionality account has nothing to say about why those groups have different norms. For example, the Lamalera of Indonesia make more fair offers in the Ultimatum Game than do the Tsimane of Bolivia, or the Machiguenga of Peru. In accommodating this fact, however, Baumard et al.
simply fall back upon the same explanation offered by the broad adaptationists who conducted
the study: “In accord with their cultural values and practices, Lamalera proposers in the
Ultimatum Game think of the money as owned in common with the recipient, whereas Tsimane
and Machiguenga proposers see the money as their own and feel entitled to keep it” (p. 71). The
cultural values and practices referred to here consist in normative beliefs and motivations that
study participants have acquired through social learning, in virtue of the fact that, in the social
environments in which they were acculturated, there were many cultural models who
antecedently possessed those same beliefs and motivations. And these cultural models most
likely acquired the same traits in the same way, as did their cultural ancestors before them.
Clearly, this does not explain the innovations, or cultural mutations, that contributed to this
evolutionary process, so even the broad adaptationist framework does not explain everything
about the evolution of different social environments. But selection among these cultural traits
offers a way of explaining certain facts about the evolution of the social environment that are
essential for explaining the religious and moral psychology of the people who inhabit those
environments.

4.7 Conclusion

For approximately the last 10,000 years, genetic selection in humans has been tracking a moving
target, as our physical environments have evolved much more rapidly than our genome. Most
human beings alive today inhabit physical environments that are nothing like the undeveloped
lands around them, while 15,000 years ago there was virtually no distinction to be made at all
between developed and undeveloped lands. There is probably little point in attempting to state in
general terms how great the impact of genetic selection may have been on our cognition and behavior during this time period. But the magnitude of the differences between modern environments and those of 15,000 years ago suggests that some of the new selection pressures generated by civilization may be quite strong, and recent work in genetics has suggested that natural selection can act much more quickly on human genes that we had previously supposed (Laland, Odling-Smee, and Myles 2010). So by the standards of evolutionary timeframes, it appears that modification of the physical environment may have become a major factor in the construction of the human niche, but only very recently. And as I noted in Section 4.2.2, the accumulation of culturally selected knowledge, skill and technology is a major factor explaining how our physical environments have changed.

Over the same period of time, however, our social environments have also changed dramatically, as groups have evolved from small-scale foraging societies to large-scale, industrialized nations. The mounting evidence of gene-culture coevolution suggests that the recent evolution of the social environment, too, may be a major factor in the construction of the human niche. And the accumulation of culturally inherited information has also been a major factor explaining how our social environments have evolved.

In particular, selection acting on culturally inherited religious beliefs and moral motivations has established stable patterns of behavior within particular social environments, and this has influenced, in turn, both the genetic fitness and the social learning histories of individuals inhabiting those environments. Unlike genetic selection, cultural selection includes mechanisms for maintaining variation at the group level, so these stable patterns of behavior vary in many ways from one social environment to the next. And since the groups of individuals who constitute these different environments often compete in various ways, some of this cultural
selection has taken place at the level of groups, rather than individuals. As a result, the overall meta-population is now filled with religious beliefs and moral norms that generate prosocial motivation, causing people to behave in ways that are costly for themselves, but beneficial for their groups.

Cultural group selection thus explains why people learn to invest in their social groups, or societies. Human beings often possess intrinsic motivations to pursue various “causes,” “higher purposes” and “divine natures,” which benefit “the greater good.” Sometimes the greater good consists in the welfare of other individuals, but sometimes it consists instead in the group’s beliefs, norms and institutions—its “way of life.” Often, prosocial behavior is not explicitly intended to have either of these types of benefit. Choosing not to steal may be explicitly intended to please God, avoid punishment, or just to do the right thing. But whatever the agent’s explicit intentions, by choosing not to steal she is paying an opportunity cost from which the would-be victim benefits. Prosocial motivations like these probably increase genetic fitness most of the time, in the long run. But nevertheless, they are also costly most of the time, and sometimes these costs are not recovered. An individual’s history of learning and development within a particular social environment can explain a lot about both the conditions under which she engages in costly behavior and the magnitude of the costs she is willing to pay in doing so.

Even if the capacities for acquiring prosocial goals are genetically inherited traits, genetic selection does not explain what, specifically, is acquired. One cannot become motivated to make sacrifices for the sake of what is right until one has learned which ways of acting are the right ones, and these details vary across social environments in ways that can’t be explained by genetic inheritance. Even if all humans are genetically disposed to behave prosocially for the sake of impartial, fair dealing, what counts as fair dealing depends in many ways on the
particular norms, beliefs and practices that have evolved in particular environments. Genetic selection does not explain these evolutionary details, even if it can accommodate them. Similar remarks apply for the other sub-domains of the moral domain.

Likewise, even if all humans are genetically predisposed to prefer religious systems espousing proportionality principles such as the Golden Rule, this cannot explain why some individuals’ prosocial motivations are activated by beliefs about Jesus, while others’ prosocial behavior depends on beliefs about Karma instead. It’s possible that genetically inherited traits explain why people are likely to be religious, likely to be moral, and likely to favor moralizing religions, just as they explain why people are likely to learn some language or other. But just as genetic selection does not explain which language an individual will acquire, it also does not explain why or how Christian individuals acquire the particular psychological traits that make them Christians, rather than Hindus or Jews. These facts about religious psychology depend upon chains of cultural transmission that stretch back for thousands of years, chains that have remained unbroken during that time period because of selection, but not genetic selection. Genetically inherited traits vastly underdetermine the content of moral and religious belief, and of normative motivation.

These are observations about the role social learning plays in determining which forms of prosocial behavior and cognition an individual will exhibit. But social learning may also play a role in determining how strong the relevant motivations should be. The costliness of prosociality means that prosocial motivations must often compete with other motivations not to be prosocial. It helps to think that God is on one’s side when going into battle, but that does not make the fear go away. In order to be properly acculturated to a particular social environment, the magnitudes of one’s normative motivations must correspond to the punishments and rewards at stake. To
have a basic functional understanding of the norms in a given environment requires knowing the
difference between minor gaffes and unspeakable crimes. We learn these things from parents,
other family members, peers, priests, teachers, movies, books, television and many other sources.
They are not evoked by features of the physical environment.

The Christian emphasis on temptation nicely illustrates the role of religious norms,
specifically, in influencing prosocial motivations. In his famous “soul-making” theodicy, Hick
(1978) eloquently describes how Abrahamic norms emphasize “moral effort,” regardless of
which specific forms of behavior one is supposed to put this moral effort into:

The value-judgment that is implicitly being invoked here is that one who has
attained to goodness by meeting and eventually mastering temptations, and thus
by rightly making responsible choices in concrete situations, is good in a richer
and more valuable sense than would be one created *ab initio* in a state of
innocence or of virtue. In the former sense, which is that of the actual moral
achievements of mankind, the individual’s goodness has within it the strength of
temptations overcome, a stability based upon an accumulation of right choice, and
a positive and responsible character that comes from the investment of costly
personal effort… human goodness slowly built up through personal histories of
moral effort has a value in the eyes of the Creator which justifies even the long
travail of the soul-making process. (1978, p. 292)

The culturally transmitted beliefs and norms of particular religious systems not only specify what
people should and should not do. They also specify how much effort they should put into doing
it. And the more individuals are willing invest in their group and its ideas, the more likely those ideas are to spread and persist, and the more prosocial the social environment becomes.

According to the tribal instincts hypothesis, capacities for identifying prosocial norms, and for developing intrinsic motivations that are sufficiently strong, are genetically inherited traits that have evolved by individual-level selection. Thus, when proportionality theories claim that intuitions about proportionality are an important source of prosocial motivation that evolved by individual-level genetic selection, broad adaptationists may heartily agree, as Gintis says. However, from the broad adaptationist perspective proportionality theories describe only one “half” of what is in fact a coevolutionary process, the other “half” of which is constituted by cultural group selection acting on particular norms and beliefs—even within the domain of fairness. To attempt to explain the genetic evolution of norms (in any domain) without appealing to facts about cultural selection is like attempting to explain the evolution of flowers without appealing to selection in bees. Coevolutionary relationships, where they exist, should be recognized as such. In the domains of morality and religion, the narrow adaptationist framework makes this impossible.

This does not show that proportionality theories are in any way false or inaccurate. These considerations provide no reason to reject these theories or any other accounts based on evoked culture. There is, however, good reason to reject the methodological framework from which such theories have been derived. Accordingly, I do not disagree with the explanations provided by proportionality theories, but I do disagree with Pinker’s methodological claim that the “radical revisions” to evolutionary theory proposed by broad adaptationists “seem superfluous.” We have proportionality theories of religion and morality to thank for exposing the central flaw in narrow
adaptationist reasoning, because they are responsible for illustrating just what the narrow framework is not capable of explaining.
In focusing on the intersection between two emerging scientific disciplines, my aim has been to identify ways in which each discipline benefits from the other. Because both fields are still quite young and heavily interdisciplinary, both are still grappling with fundamental conceptual issues. But the overlap between them provides opportunities in both domains for clarifying claims and commitments, identifying theoretical relationships, and responding to important challenges.

Chapters 2 and 3 focused primarily on defending claims about the evolution of religious psychology, and they did so by appealing to cultural evolutionary theory. Conversely, however, Chapter 4 focused primarily on defending claims about cultural evolutionary theory, and did so by appealing to claims about the evolution of religious psychology. I briefly review here the major contributions each discipline has made to the other.

Among the most important implications of cultural evolutionary theory for the study of religion are those concerning the role of group selection. Broad adaptationists universally agree that group selection is of pivotal importance in understanding the psychological functions of religious beliefs and the social functions of religious systems and institutions. Yet they do not all agree about whether group selection explains only the cultural traits of religion, or whether, because of gene-culture coevolution, genetic group selection may also occur. Chapter 2 addressed this debate, concluding that group selection explains only the culturally inherited, socially learned traits of religious altruism. Importantly, however, the reasons behind this conclusion generalize, showing not only that genetic group selection has not been significant in the evolution of religion, but that it also has not been significant in the evolution of other
psychological traits. Thus, while the primary aim of this chapter is to identify the evolutionary dynamics responsible for religious evolution, it also yields an important conclusion about the evolution of cultural phenomena in general.

This debate is located within the broad adaptationist camp, but Chapter 3 addresses a broader, more general debate. And though this debate is distinct from the debate between broad and narrow adaptationists, it mirrors that debate in an important way. All of the theories that present religious belief as an evolutionary by-product, or as a product of adaptations possessing non-religious functions, are based on the narrow adaptationist framework. And virtually all adaptation theories, which present religion as an adaptation in its own right, are based on the broad framework. It thus appears that theorists’ fundamental methodological and theoretical assumptions about evolution in general are influencing their empirical views about religion, specifically. By-product theorists deny that religion is an adaptation, but this is because they take a very narrow view of the process of adaptation itself. Adaptation theorists, by contrast, hold that religion is an adaptation, but this is because they have a broad and pluralistic view of adaptive processes, which involves more than just genetic selection. Thus, while the adaptation/by-product debate itself is an empirical debate about the evolution of the religious phenotype, it appears to be driven by prior commitments about the forms of selection relevant for explaining human behavior.

Once the empirical debate between adaptation theories and by-product theories is clearly distinguished from the methodological debate between broad and narrow adaptationists, two important things follow, which are the topics of Chapters 3 and 4, respectively. First, with regard to the adaptation/by-product debate, a natural resolution emerges, according to which there is no winner or loser. Instead, as the Goldberg Exaptation Model illustrates, these empirical theories
can be integrated. Second, with regard to the broad/narrow debate, a winner emerges: broad adaptationism. Broad adaptationism explains everything that narrow adaptationism explains, and much more besides, and there are no compelling reasons to constrain appeals to selection strictly to genetic selection. Adding cultural selection to genetic selection simply makes it possible to integrate and organize a much wider range of social and psychological facts from the perspective of a single, coherent evolutionary framework. I see this conclusion as a contribution to the science of cultural evolutionary theory in general, since the debate between broad and narrow adaptationists is a debate about the role of cultural selection and group selection in explaining human behavior in general. And I see the resolution of the adaptation/by-product debate as a contribution to the science of religious evolution, specifically.

In addition, however, the Goldberg Exaptation Model may be able to contribute in other ways, which suggest directions for future work. The GEM was identified in order to emphasize an important distinction between questions about the religious phenotype in general, as a whole system possessing its own functional role, and questions about the constituent traits or parts that comprise this system. The by-product theory usefully addresses “parts-level” questions, but it is the adaptation theory that addresses the “system-level” questions. Existing by-product theories explain the properties and activities of certain traits that form important parts of the system, but it is the adaptation theory that addresses etiological questions of selection history, which make it possible to identify the function of the system itself in a way that is based on empirical theory, rather than intuition. Very roughly, this function appears to be that of using descriptive beliefs about supernatural agents to activate normative motivations. But developing a less rough, more precise account of the function of the whole religious phenotype could prove a fruitful avenue to pursue in future research.
Moreover, with an answer to this question in hand, answers to other system-level questions may follow, and these, too, provide opportunities for future research. These questions concern the specifically religious roles played by the parts in the system, which depend upon what the overall function of the system itself is. They are interesting because they suggest that etiological functionalism might be used to develop a general theory of religion as a natural phenomenon—a principled way of identifying and individuating the traits, facts, behaviors, functional roles, etc. that count as religious phenomena, specifically. As by-product theorists emphasize, the genetic traits they identify possess non-religious functions—the ones that enabled them to evolve in the first place. But if there is some interesting sense in which these traits count as religious traits, then there must be some sense in which they function, psychologically, in a specifically religious way. At the moment, as I explain in Chapter 4, traits, functions and facts in general are assumed to count as religious simply on the basis of common-sense intuition. A great deal of work has been devoted to explaining cognition involving supernatural agents, and this is simply because beliefs involving supernatural agents seem, intuitively, to be religious beliefs. But clearly there are also non-religious beliefs about supernatural agents as well, and while we can make intuitive decisions about which ones count as the religious ones and which ones do not, we still lack any theory-based, scientific way of drawing such distinctions. There is no generally accepted definition of the concept religion available, and nor is there a general theory or framework capable of identifying which phenomena count as the religious ones.

But in a field in which everyone is already committed to the etiological method for identifying traits and functions in general, I see no reason why claims about which traits and facts count as the religious ones could not be based on something more substantial than folk intuition. An adaptationist account of the function of the religious phenotype in general could be
used to replace intuition-driven attributions of religious functioning with attributions based on empirical facts about histories of selection. From the system-level perspective afforded by the GEM, it is easy to see that many traits of the religious phenotype perform more than one function, and all of the relevant functions may be identified in etiological terms. Accordingly, in addition to determining which traits count as traits of the religious phenotype—the ones that function together to perform the overall function of the Goldberg device of religion—the GEM also provides a way of determining when any given trait is functioning in a specifically religious manner, and when it is performing other functional roles. On this view, a religious trait is functioning religiously only when it is interacting with the other parts of the system. If system-level questions in general can be answered in etiological terms, then attributions of function in the science of religion in general can be grounded in empirical facts about selection history, rather than in functionalist intuitions. Thus, when intuitions conflict regarding which functional roles are relevant in the study of religion, it would be possible to appeal to an independent authority. And more importantly, a general framework would elucidate the theoretical relationships obtaining among research programs devoted to studying different aspects of religious belief and behavior, enabling researchers to articulate the significance of their findings relative to the basic aims of the discipline.

In light of all this, I see the Goldberg Exaptation Model as a contribution to the study of religious evolution, but one that depends upon a distinction that is essential to the study of cultural evolution: that between genetic selection, in particular, and Darwinian processes of selection in general. And this distinction, in turn, is the central issue at stake in the debate between broad and narrow forms of adaptationist methodology. Accordingly, I see the argument against narrow adaptationism, and the identification of the broad/narrow distinction itself, as
contributions to the field of cultural evolution that happened to emerge naturally from an examination of theories of religious evolution. In Chapter 4 I argue explicitly that the broad framework is needed to explain the evolution of religion, but I also show that the same concerns apply as well to the study of morality. In part, this is motivated by the fact that morality is so intimately related, psychologically speaking, to religion. But it is also important to see that the broad framework is necessary not just for understanding morality in relation to religion, but also for understanding morality itself, as its own form of cultural phenomenon. And, though this is not a point I argue for in Chapter 4, it seems likely that if religion and morality are cultural phenomena that are best explained by an adaptationist methodology including cultural selection and group selection in addition to individual-level genetic selection, then so too are other cultural phenomena.

Separating the methodological and ontological commitments of the broad and narrow frameworks in general from the particular empirical theories derived from those frameworks makes it easier to see that, contrary to the claims of those who remain skeptical about cultural selection, the science of cultural evolution cannot live on evoked culture alone. This point is crucial for understanding cultural phenomena in general, and religious phenomena in particular, from a scientific perspective capable of integrating theories of psychology and social science with the rest of scientific knowledge.
Bibliography


