

UNIVERSITY OF CALGARY

Darwinizing Religion: William Irons' Evolutionary Theory of Religion

by

Edyta J. Czarnecki

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES  
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE  
DEGREE OF MASTER OF ARTS

DEPARTMENT OF RELIGIOUS STUDIES

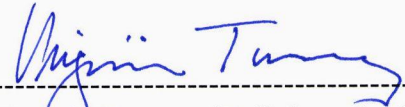
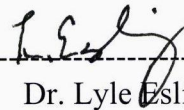
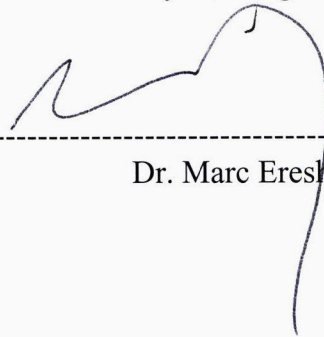
CALGARY, ALBERTA

SEPTEMBER, 2006

© Edyta J. Czarnecki 2006

THE UNIVERSITY OF CALGARY  
FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "Darwinizing Religion: William Irons' Evolutionary Theory of Religion" submitted by Edyta J. Czarnecki in partial fulfillment of the requirements for the degree of Master of Arts.

  
-----  
Supervisor, Dr. Virginia Tumas (Religious Studies)  
-----  
Dr. Lyle Eslinger (Religious Studies)  
-----  
Dr. Marc Ereshefsky (Philosophy)

Sept. 14/06

  
-----

Date

## Abstract

This thesis examines William Irons' costly signaling theory of religion as an adaptation for intragroup cooperation. I provide a thorough exposition of Irons' theories of morality and religion and argue that Irons' description of religion as an extension of morality is more aptly the superimposition of religion onto morality, leading Irons to make theoretical arguments that do not justify his adaptive explanation. Establishing that religion is built on pre-existent non-religion-specific adaptations and contending that it is the most powerful hard-to-fake signal of commitment are not consistent with the claim that religion is an adaptation. In comparison to the adaptive approach, alternative evolutionary perspectives allow for a more sophisticated conceptualization and empirical testing of religion, and a survey of the coherency of Irons' arguments within those competing evolutionary frameworks reveals that they are consistent with a bio-cultural approach.

## Acknowledgements

This thesis would not have been written without the support of my personal cheerleading squad. I owe each member a heartfelt thank you.

Virginia Tumas, for pushing without pushing and aptly dealing with my obsessive personality.

Eliza and Janusz Grabowski, for helping me swallow a frog I preferred to look at.

Kasia Hinks, Jackie Chambers, and Andrew Fernandes for listening to desperate rants.

My family and friends, for not seeing me for extended periods of time but still remembering my phone number.

## Table of Contents

Approval Page .....	ii
Abstract .....	iii
Acknowledgments .....	iv
Table of Contents.....	v
List of Tables .....	vi
Introduction .....	1
Chapter One: Identifying Adaptations .....	11
Evolutionary Science .....	11
Evolutionary Theories of Religion .....	26
Chapter Two: Irons' Adaptive Theory of Religion .....	35
The Evolution of Morality .....	36
The Evolution of Religion .....	42
Belief .....	43
Ritual .....	49
Chapter Three: Irons' Theoretical Evidence for Religion as the Target of Selection .....	62
Two Problems .....	64
Alternative Evolutionary Approaches .....	75
Do Irons' Arguments Support a Byproduct Explanation? .....	75
Do Irons' Arguments Support a Coopted Adaptive Explanation?.....	84
Do Irons' Arguments Support a Bio-Cultural Explanation?.....	91
Conclusion .....	106
Bibliography .....	114

## List of Tables

### Table 1

Conceptual and Evidentiary Criteria for Evaluating Evolutionary Categories ....20

## Introduction

Contemporary scientific interest in explaining the pervasiveness of religion by means of evolutionary theory dates to the mid-1970's. Sociobiologist E. O. Wilson is one of the earliest scientists to have applied population biology, comparative zoology, evolutionary theory, and genetics to the general area of human social behavior and more specifically to religion. Today numerous scientists focus their interest on applying evolutionary theory to explain religion. Some seek to expose the fitness-enhancing function for which religion evolved whereas others concentrate on depicting the evolutionary landscape of the mind that generates and maintains religious beliefs and behaviors. Some argue that religion is an adaptation, others postulate that it is an incidental consequence of non-religious adaptations, and others contend that it originated in a new use of either a pre-existing adaptation or a side-effect of an adaptation.

In this thesis I examine behavioral ecologist William Irons' evolutionary explanation of religion as an adaptation for enhancing within-group cooperation.<sup>1</sup> I evaluate his theory and method for consistency with an adaptive approach<sup>2</sup> to religion and conclude that his theoretical evidence, two arguments in particular, do not justify his theory that religion was the target of selection. His claims that religion is built on pre-existing social strategies and that it is the most powerful hard-to-fake signal of commitment are problematic within an adaptive approach to religion. I argue that these problems stem from his grounding of his theory of religion on top of his theory of morality as well as his commendable but hasty integration of the perspectives of behavioral ecology and evolutionary psychology. I analyze whether his arguments are more coherent in light of different, competing evolutionary approaches and explore how the adoption of a different approach would affect the conceptualization of religion and its empirical examination.

Although many evolutionary scientists are in the process of thoroughly investigating religion, this area of research seems to have slipped under the radar of religious studies scholarship. A brief survey of recent textbooks used in religious studies in the area of science and religion confirms that religious studies scholars do not often

engage contemporary scientific scholarship concerning the evolutionary nature of religion. For example, in their anthology Religion and Science: History, Method, Dialogue (1996) editors Mark Richardson and Wesley Wildman include William Irons' article "Morality, Religion, and Human Evolution" and theologian Philip Hefner's response to Irons entitled "Theological Perspectives on Morality and Human Evolution." However, Irons' article is one of his first ventures into extending his evolutionary theory of morality to religion; despite a length of twenty-four pages only three of these pages, including a brief mention in the conclusion, refer to his hypothesis concerning religion. Granted, this anthology is ten years old and the most significant amount of evolutionary scholarship on religion, including Irons' work, has occurred since then. However, if the editors' intention was to address the "substantive interactions" between science and religion in terms of biological and cultural evolution, they could have included more developed arguments from Stewart Guthrie or Pascal Boyer who at the time of the anthology had already published Faces in the Clouds: A New Theory of Religion (1993) and The Naturalness of Religious Ideas: A Cognitive Theory of Religion (1994), respectively.

The only references to any scientific evolutionary theories of religion in Ted Peters' anthology Science and Theology: The New Consonance (1998) occur in Francisco Ayala's essay, "So Human an Animal: Evolution and Ethics," and theologian Arthur Peacocke's paper, "A Map of Scientific Knowledge: Genetics, Evolution, and Theology." Both mention the work of E. O. Wilson but focus on his insistence on biological gain due to epigenetic moral codes and his overarching goal of rendering sociobiology *the* conceptual framework for attaining knowledge. Neither Ayala nor Peacocke provide any commentary on the general area of scientific theories of the evolution of religion.

Alister E. McGrath makes it clear that his book Science and Religion: An Introduction (1999) is "only an introduction" upon which "severe limits of space" are imposed (xii). His book understands the relationship between religion and science in terms of whether religion facilitates or hinders scientific development and does not consider scientific development on the subject of religion. In his chapter, "Issues in

Science and Religion,” with respect to the interaction of biology and religion he reviews “the impact of various forms of Darwinianism on religious thought” and when discussing the relationship between psychology and religion he reflects on “various approaches to understanding the origins and significance of religion” (178). McGrath considers Richard Dawkin’s arguments in the latter section, but it should be noted that Dawkin’s evolutionary theory of religion concentrates on showing how the notion of the process of natural selection “eliminates religious belief in God” rather than presenting a developed explanation of the evolutionary origin of religion. McGrath’s reference to the origin of religion in the context of psychology suggests that he may consider the current work of the cognitive science of religion but instead he recounts traditional 19<sup>th</sup> and 20<sup>th</sup> century essentialist conceptualizations of religion. He fails to mention current scholarship from biology or psychology regarding the evolutionary nature of religion in either the body of this chapter or in the list of recommended further readings.

Ian Barbour’s Religion and Science: Historical and Contemporary Issues (1997) is a standard textbook for science and religion in religious studies. In the context of “Religion and Human Nature,” Barbour devotes four pages to “The Evolution of Religion.” Here he summarizes Ralph Burhoe’s argument from a 1975 article: “religion has been the major force in fostering altruism and social cooperation extending beyond genetic kin [...] Religion has been selected because it contributed to the survival of the biocultural group.” This bears significant similarity to David Sloan Wilson’s current biocultural, group-selection theory of religion. However, Barbour’s discussion concludes that Burhoe developed this argument into the claim that, as religious belief is an expression of the current culture’s understanding of the world, to remain credible today, it must be reformulated “along strictly scientific lines.” In other words, Burhoe’s comments are not akin to the evolutionary theories that I consider in this thesis because they argue for “*evolutionary naturalism* as the religious philosophy best suited to a scientific culture” (263). Barbour rightly points out that as appeals to a metaphysical system, Burhoe’s claims are “not part of science itself” (264). Although Barbour mentions that, “anthropologists and sociologists have portrayed the functional role of religion in binding individuals in social groups and in *preserving the social order*,” he

does not call upon William Irons' theory or any other more current evolutionary theory of religion. Likewise, despite briefly considering the evolution of "the three basic features of religion," namely ritual, story, and religious experience, he does not mention any of the recent work that I analyze below (264-65). In the chapter "Evolution and Continuing Creation," Barbour provides a detailed exposition of the background and current debates in evolutionary theory, which is helpful to understand modern evolutionary explanations of religion, but instead of discussing such explanations, he looks at the theological implications of evolution.

Barbour's When Science Meets Religion: Enemies, Strangers, or Partners? (2000) appears to be a condensed version of his above-mentioned work. As in his previous book, it contains a chapter entitled "Evolution and Continuing Creation." It includes a discussion of paleontologist Stephen Jay Gould's, biologist Richard Dawkins', and philosopher Daniel Dennett's presentations of current evolutionary theory as incompatible with, or in some instances independent of, theism. Although the contribution of these scholars to the current understanding of evolution is invaluable and I refer to each of them in this thesis, with the exception of Gould's undeveloped example of religion as an exaptation (see Gould, "Exaptation"), they do not put forward any full-blown theories of religion. However, Barbour's intention is not to look at the contemporary research in the evolutionary science of religion. His book's general purpose is to classify the relationship between science and religion as one of four categories: conflict, independence, dialogue, or integration, and his discussion of evolution attempts to do likewise for evolution and theists. Barbour's subsequent chapter, "Genetics, Neuroscience, and Human Nature," addresses "new scientific disciplines that present strong challenges to traditional religious beliefs concerning human nature" (121). Among these disciplines he includes "sociobiology," which he describes as "the study of the evolutionary origins of social behavior" (121). However, sociobiologists are not the only scientists to study the evolutionary origins of social behavior; numerous evolutionary scientists in addition to those working from a specifically sociobiological perspective agree that religion is a social behavior and study its evolutionary origins. Rather than including a discussion of any developed scientific arguments regarding the

evolutionary origins of the social behavior of religion, Barbour focuses on the extreme reductionism of E. O. Wilson's sociobiology and labels it as in conflict with religion. Barbour highlights both Wilson's conviction that his own evolutionary account of religious beliefs will render modern religion obsolete (13) and his proposal that religion's functions ought to be replaced by an evolutionary narrative (156). Barbour does not question the methodology that led Wilson to these conclusions. Although Barbour is aware that Wilson's sociobiology is sweeping and almost dictatorial (13, 155-156), he refers to Wilson's evolutionary explanation of religion several times while paying no attention to other contemporary evolutionary scientists offering comparatively more developed evolutionary theories of religion. By looking at any of these other explanations Barbour may have reached the conclusion that evolutionary explanations of religion are not necessarily in a relationship of conflict with religion.

One goal of this thesis is to create a platform for dialogue between religious studies scholars and evolutionary scientists. I intend to relate the scientific study of the evolutionary nature of religion to religious studies scholars as well as demonstrate to evolutionary scientists that religious studies scholars can detect problematic areas of their evolutionary appreciation of religion. Because I consider whether Irons' arguments are compatible with different, competing evolutionary approaches, this thesis also offers a single comprehensive review of key current evolutionary theories of religion and several of the most pertinent scientific debates.

My methodology consists of a critical conceptual analysis based on an extensive literature review of current evolutionary theories of religion. Although I concentrate on William Irons' particular adaptive explanation of religion, my critical analysis of his theory involves a comparative examination of the basic principles of various evolutionary approaches, including how they relate to the conceptualization and empirical testing of religion. To identify whether Irons' arguments are consistent with the adaptive approach to religion, I analyze the key evolutionary concepts of 'adaptation' and 'function' and appraise Irons' arguments in light of current understandings of those key concepts. To assess whether Irons' arguments might support a different evolutionary approach, I examine the conceptualization of traits, including their origin and relation to fitness, in

the different evolutionary categories of byproduct, coopted adaptation, and coopted byproduct. In identifying inconsistencies in the use of evolutionary terminology, this thesis touches upon issues that are pertinent to the philosophy of science.

This thesis is timely because the scientific project of establishing the evolutionary nature of religion is in its infancy and as of yet there is no consensus in the sciences regarding the evolutionary nature of religion. Between the tentative nature of science and the early stage of this scientific endeavor, these evolutionary theories of religion should be open to clarification and modification given new information and commentary. There is no consensus within evolutionary science regarding the subject of religion primarily because the existing research is conducted within the scientific fields of behavioral ecology and evolutionary psychology, which for the most part have proposed independent and opposing theories of the evolutionary landscape of religion. In general, behavioral ecologists focus on fitness consequences and propose that religion is an adaptation while evolutionary psychologists concentrate on the role of evolved mental modules in religious belief and practice and contend that religion is not an adaptation. Although these explanations are systematically connected because they employ evolution as their metatheoretical framework, they typically use the different evolutionary categories of adaptation and byproduct to explain religion. What makes dialogue between all of these scientists even more difficult is that they are not consistent in their use and hence with their definitions of key evolutionary terms such as adaptation and function. Not only is the current evolutionary study of religion not interdisciplinary in terms of a collaborative effort between evolutionary science and religious studies, but also it does not seem to be interdisciplinary even among the different disciplines that comprise evolutionary science. In short, the behavioral ecologists do not seem to be talking to the evolutionary psychologists. Dialogue between the two perspectives as well as a consistent terminology would only benefit the evolutionary study of traits, including religion.

Some of the scientists proposing evolutionary theories of religion recognize that they could supplement their existing knowledge of religion with insights from current scholarship on religion. For example, David Sloan Wilson extends an open invitation to non-science scholars to participate in the discussion: “Another reason to discuss the

survey at an early stage is to invite readers with appropriate expertise to join the game. Scholarship is, or should be, a communal effort, and there are many scholars of religion more highly qualified than I to evaluate these particular religious systems” (Cathedral 158). This thesis is a response to this call for engagement by religious studies scholars to increase the breadth and accuracy not only of D. S. Wilson’s specific theory, but rather the whole of the investigation into religion by evolutionary science. I aspire to make these theories known to non-science scholars of religion and to contribute to making the project of these scientists a truly interdisciplinary effort, not only among the hard sciences and social sciences, but also inclusive of the humanities.

As I mentioned above, several theologians have engaged in discussion with the sociobiologists who initiated the evolutionary study of religion. However, unlike theologians, religious studies scholars hold that the academic study of religion requires positing religion as a human phenomenon that can be studied without reference to the objective truth of revelation or transcendent reality. This important characteristic of religious studies allows the discipline to relate to this newly emerging scientific interest in religion. The collaboration of the evolutionary sciences with religious studies can only benefit the scientific enterprise. Unlike most scientists, religious studies scholars are well versed in the nature, theory, and multiple expressions of religion; they can contribute both their understanding of religion as well as their sensitivity to methodological implications to the project.

This thesis is divided into three chapters and a conclusion. Chapter One, “Identifying Adaptations,” is a comprehensive guide to evolutionary theory, concentrating on the concept of adaptation as well as the standard method of identifying adaptations. To prepare to assess the consistency of Irons’ arguments with his adaptive explanation of religion, I establish what it means to claim that religion is an adaptation as well as the nature of the theoretical evidence that is required to support such a claim. To provide the necessary background for an evaluation of Irons’ arguments within different evolutionary approaches to religion, a task I undertake in the third chapter, I also introduce the byproduct approach to religion and the coopted adaptive approach to religion.

In Chapter Two, “Irons’ Adaptive Theory of Religion,” I discuss Irons’ explanation of religion as an adaptation. Irons states that his theory of religion is an extension of his theory of morality; hence I begin with a detailed exposition of his evolutionary explanation of morality. In describing Irons’ evolutionary explanation of religion, I point out its weak claims and address those gaps with Joseph Bulbulia’s and Richard Sosis’ costly signaling models of religion, which they both claim are based on Irons’ explanation of religion as a hard-to-fake signal of commitment.

In Chapter Three, “Irons’ Theoretical Evidence for Religion as the Target of Selection,” I analyze whether Irons’ theoretical evidence supports his claim that religion is an adaptation. I explore his understanding of the relationship between his theory of morality and his theory of religion and identify two arguments that I consider to be problematic given his adaptive approach to religion. They stem from Irons’ failure to differentiate his theory of religion from his theory of morality as well as his attempt to integrate how religion is generated in terms of evolved psychology into his behavioral ecology explanation of why religion exists. He identifies the psychological mechanisms grounding religiosity as those evolved in the context of morality apparently without considering the criteria for adaptation within evolutionary psychology. Evolutionary psychology holds that grounding religion in ordinary mental modules evolved independently of religion prevents religion itself from being an adaptation. Stipulating that religion is the most effective expression of a non-religion-specific mechanism is not sufficient to demonstrate theoretically that religion is an adaptation. Since Irons’ arguments do not justify his theoretical claim that religion is an adaptation, I analyze them in light of the principles of competing evolutionary theories. I find that Irons’ basing of religion on pre-existent adaptations is explicable from the byproduct approach and the two exaptive approaches, but his contention that religion is the most powerful of one of these ordinary adaptations is tenable in none of those frameworks. I conclude, however, that both are consistent with a bio-cultural explanation of religion.

I intend to engage William Irons’ evolutionary theory of religion to show that his theoretical evidence does not justify his claim that religion is an adaptation. Irons’ understanding of the relationship between the evolution of morality and religion as well

as his seeming lack of expertise in the criteria for adaptation in terms of the evolution of psychological mechanisms render his explanation of religion inconsistent within an adaptive framework. If religion exists because of ordinary mental modules evolved in the context of morality and it has their same function of enhancing within-group cooperation, then religion itself is not an adaptation. Through considering Irons' theory within the larger picture of evolutionary science, this thesis has the lofty overarching objective of demonstrating that dialogue among different evolutionary approaches within evolutionary science as well as among evolutionary science and religious studies can lead to new insights in the academic study of religion that would not be attainable if these different areas of scholarship remained independent.

---

Introduction Notes

<sup>1</sup> I focus on Irons' theory as he has elaborated his costly signaling theory of religion in several articles, and well-known fellow costly-signaling theorist, Joseph Bulbulia, credits Irons as the first theorist to describe religious behavior as a commitment device ("Review" 669). Moreover, anthropologists Richard Sosis and Eric Bressler claim that, "Irons's work moved beyond earlier research by focusing on the selective pressures that may have shaped human systems of belief and favored religion as a universal strategy to promote cooperation" (212). Sosis and Bressler do not refer to anthropologist Lee Cronk's and economic sociologist Laurence Iannaccone's theories that pre-date Irons' first article in which he theorizes about the evolution of religion, namely "Morality, Religion, and Human Evolution" (1996), (see Cronk's "Evolutionary Theories of Morality and the Manipulative Use of Signals" (1994) and Laurence Iannaccone's "Sacrifice and Stigma: Reducing Free-Riding in Cults, Communes, and other Collectives" (1992)). However, Irons cites Cronk and Iannaccone as, like himself, holding that, "Practicing a religion often entails considerable cost, and thus religion can serve as a hard-to-fake signal of commitment to a particular social group" ("Co-Creator" 781). Despite this similarity, their arguments are sufficiently different from Irons' such that Irons' can be considered to comprise the first evolutionary commitment theory of religion. Cronk admits that in "Evolutionary Theories of Morality and the Manipulative Use of Signals" (1994) his concern "was neither with the origin of religion nor with how religions are used to enhance the solidarity of relatively egalitarian communities [...]" ("Moralistic" 351). Similarly, although Iannaccone's research addresses commitment in religion, his focus is on economic rationality and not commitment itself.

<sup>2</sup> By adaptive approach, I refer to the evolutionary framework that posits religion to be an adaptation. The adaptive approach is not any evolutionary approach that explains any fitness-enhancing trait. Instead, it is the approach that explains a particular trait that originally evolved because of its fitness-enhancing function. In other words, although the adaptive approach as well as the exaptive approaches of coopted adaptation and coopted byproduct explain traits with adaptive value, only the adaptive approach explains adaptations. I reiterate this distinction in the following chapter.

## Chapter One: Identifying Adaptations

As in this thesis I intend to analyze whether behavioral ecologist William Irons' adaptive theory of religion is justified by his own theoretical evidence, I first explain what it means to claim that religion is an adaptation. This chapter sets the stage for the following chapters with a discussion of evolutionary theory and its application to the study of religion. In the first section, "Evolutionary Science," I outline the fundamental principles of evolution and the mechanism of natural selection according to contemporary scientific scholarship and discuss at length the concept of adaptation. I introduce 'adaptationism' as a methodology and address criticisms of its employment as a valid scientific means to detect adaptations. This discussion provides the background for the second section, "Evolutionary Theories of Religion," in which I introduce how evolutionary theory is employed in the scientific study of religion.

### Evolutionary Science

Evolutionary science seeks to explain the natural world through the theory of evolution. The University of California Museum of Paleontology's Understanding Evolution website claims that, "The 'Theory of Evolution' is an evidence-based, internally consistent, well-tested explanation of how the history of life proceeded on Earth." It provides scientists with a framework within which to propose hypotheses and to test specific falsifiable predictions.

Following Charles Darwin, contemporary evolutionary scientists understand evolution as descent with modification. Biological evolution occurs whenever there is a change in gene frequency in the gene pool of a population over time. Darwin's most important contribution to evolutionary theory was identifying the mechanism of evolution as "natural selection." Changes in gene frequency can occur via natural selection or 'genetic drift.' Unlike genetic drift, which is the cumulative effect of random sampling in the production of offspring, natural selection "is overwhelmingly the most important force in evolution and the only one that assembles and holds together particular ensembles of genes over long periods of time" (E. O. Wilson, Sociobiology 33).

According to Darwin, natural selection is the inevitable outcome of ‘phenotypic variation,’ ‘inheritance,’ and ‘fitness.’ Organisms<sup>1</sup> in a local population in a certain environment vary in the particulars of their physical, behavioral, or psychological traits or tendencies; in other words, they exhibit phenotypic variation. A simple example of a variable physical trait is the human femur: all humans have femur bones that facilitate bipedal movement, but some are longer and thicker than others. The processes of ‘mutation’ (new genetic variants), ‘genetic recombination’ (shuffling of pre-existing genetic variants), and ‘gene flow’ (migration) supply natural selection with genetic variation from which to select among within a population (E. O. Wilson, Sociobiology 32).

‘Fitness’ refers to the survival and reproductive success of an organism in comparison to competing organisms. It is the result of the interplay between a particular trait of the organism and the existing selection pressures in a more-or-less bounded environment. Since natural selection depends on the differential survival and reproduction of competing organisms, fitness is a relative concept (D. S. Wilson, Cathedral 38). This means that it is not important how well an organism survives and reproduces in absolute terms. Instead, it is important how well an organism survives and reproduces relative to competing organisms in the same environment.

The ‘environment’ is an all-encompassing concept meant to include any local conditions that affect the survival and reproduction of its inhabitants. These conditions can be external to the population, such as geographic location or ecological niche, or internal to the population, such as social or political structure. One of the most obvious constraints or selection pressures an environment exercises on its population is the limited availability of resources. As organisms in populations tend to produce more offspring than their environment can sustain, “there must in every case be a struggle for existence” (Darwin, Origin 117). Not all fertile organisms will have a maximal number of offspring and be able to support them until they in turn have their own offspring.<sup>2</sup> More importantly, not all fertile organisms with a reproductively significant<sup>3</sup> trait possess equal chances of having reproductively successful offspring in a particular environment. Differential reproduction arises because some organisms in a population possess

reproductively significant variants of a trait that fit better with the demands of their local environment relative to the other existing variants. As such, these phenotypic variants, upon interaction with persistent features of the environment, offer their carriers an advantageous edge in the competition for survival and reproductive success. Natural selection is the preservation of favorable variations, however slight, and the rejection of injurious variations (Darwin, Origin 131).

In addition to phenotypic variation and differential fitness, Darwin's depiction of natural selection includes the principle of inheritance, i. e., that the phenotypes of offspring tend to resemble those of their parents more than those of the parents' contemporaries. For a trait to have a positive fitness bias and increase in frequency in successive generations of a population, it must be heritable. Phenotypic traits and tendencies are coded in genes that can be passed on from parent to successive generations. Over time, those organisms that have more offspring relative to others because of their advantageous trait variant will amplify the presence of that trait in the next generation.

To put the three conditions considered above together in a concise statement of the process of natural selection, natural selection stipulates that the phenotypic variant of a trait that is most advantageous (or least detrimental) to the survival and reproductive success (fitness) of an organism in a particular environment, will become more common relative to alternative variants in the subsequent generations of that population in that environment.

The biological goal of all organisms is to pass on more of their reproductively significant traits to subsequent generations than fellow organisms sharing the local environment. Evolutionary psychologists John Tooby and Leda Cosmides state that, "the causal process of natural selection builds organic machines that are 'designed' to serve only one very specialized end: the propagation into subsequent generations of the inherited design features that comprise the organic machine itself" (53). Although passing on one's heritable trait variant is accomplished typically by having offspring, an organism can increase its fitness without having more offspring than others or having any at all. An individual can successfully propagate its own reproductively significant trait

variants by helping to increase the fitness of close kin (Hamilton). According to Tooby and Cosmides, “design features that promote both direct reproduction and kin reproduction, and that make efficient trade-offs between the two, will replace those that don’t” (53).<sup>4</sup>

Natural selection cannot make phenotypes evolve out of nothing (Gould, Structure 141) nor can it fix anything (Dennett 214). Evolutionary biologist George C. Williams cautions that, “Selection has nothing to do with what is necessary or unnecessary, or what is adequate or inadequate, for continued survival. It deals only with an immediate better-vs.-worse within a system of alternative, and therefore competing, entities” (31). Selecting and discarding cannot produce the absolute most perfect of all conceivable solutions to a specific problem posed by any aspect of the inclusive environment. The fact that evolution tends to be opportunistic in its tinkering with existing alternatives should not be misinterpreted as evidence of preordained progress toward a grand design (see Williams 34-35).

Despite the fact that natural selection lacks foresight (Dennett 215) and can select only variations of a trait that already exist in a given population, its process is positive. This means that natural selection’s successive modifications of pre-existing structures result in increasingly sophisticated solutions to persistent fitness problems (Tooby and Cosmides 60-61). Over evolutionary time, natural selection directs change by retaining the relatively advantageous variants in the given environment (Tooby and Cosmides 57). This results in the increased prevalence of phenotypic traits that are optimal because they are the best of the *available* trait variants for that environment (Sober, “Six” 73).

The trait that ends up being selected for is an ‘adaptation.’ Paleontologist Stephen Jay Gould remarks that adaptation is the process of transforming environmental data into internal changes of form, physiology, and behavior (Structure 157). The product of the process of adaptation too is called an adaptation.<sup>5</sup> Tooby and Cosmides provide the following comprehensive definition of an adaptation: it is

- (1) a system of inherited and reliably developing properties that recurs among members of a species that (2) became incorporated into the species’ standard

design because during the period of their incorporation, (3) they were coordinated with a set of statistically recurrent structural properties outside the adaptation (either in the environment or in the other parts of the organism), (4) in such a way that the causal interaction of the two (in the context of the rest of the properties of the organism) produced functional outcomes that were ultimately tributary to propagation with sufficient frequency (i. e., it solved an adaptive problem for the organism). (61-62)

An adaptation is a genetically transmitted, predominantly species-typical trait that was directly selected for because of its fitness-enhancing function in the given environment. It is not a correlated response to selection for another trait nor is its function fortuitous (Williams 9).<sup>6</sup> An adaptation evolved solely because, relative to the alternative traits, it solved a persistent specific problem encountered by ancestral populations (Tooby and Cosmides 62) such as finding mates, protecting offspring, fleeing predators, or pursuing prey (Atran, *Gods* vii).<sup>7</sup> Various scholars argue that the essential features of a biological adaptation are fitness-enhancement and “special design” (Williams; Schmitt and Pilcher 643-44; Buss et al. 536; Cosmides and Tooby 165): the adaptation increases the survival and reproductive success of its bearer relative to other alternative attributes and it is complex, economic, efficient, reliable, precise, and functional.<sup>8</sup>

‘Adaptationism’ is the process of identifying adaptations through ‘reverse-engineering.’ It is a “tendency in evolutionary biology to reconstruct or predict evolutionary events by *assuming* that all characters are established in evolution by direct natural selection of the most adapted state, that is, the state that is an optimum ‘solution’ to a ‘problem’ posed by the environment” (Lewontin qtd. in Dennett 238). This method is called ‘reverse-engineering’ because the evolutionary scientists who employ it start from the assumption that traits are adaptations and work backwards to figure out what fitness problems they solved. The subject of this thesis, anthropologist William Irons, employs this adaptationist framework in his study of religion.

I call, as best as I can, on the theory of biological evolution to seek the most basic explanations of a phenomenon such as religion. I begin to seek theoretical

explanation in terms of adaptation, which means I ask how a particular trait could have benefited human ancestors in past environments. (“Inquiry” 358)

Irons reverse-engineers the evolutionary origin of religion by considering what fitness-advantage it may have provided human ancestors. By examining the nature of religion and the nature of the reproductive challenges plaguing the evolution of *Homo sapiens*, Irons develops the theoretical argument that religion was selected for because it functioned as a reliable communication of commitment to one’s group that facilitated the formation and maintenance of human sociality.

Evolutionary scientists, including William Irons, identify traits as adaptations by keeping in mind both Williams’ special design features and several core principles of evolutionary biology (see Schmitt and Pilcher 644). The special design features suggest how adaptations are phenotypically manifested: more often than not they will be expressed in behaviors that are universal, functional for specific tasks, complex, energetically conservative, require interaction with the environment for activation, and increase the organism’s survival and success in reproduction. Like design specificity, central theories within evolutionary theory direct research of adaptations down certain avenues. For example, inclusive-fitness theory (Hamilton) suggests adaptations for familial aid, and reciprocal altruism theory (Trivers) points to adaptations for coalition formation (Schmitt and Pilcher 644). The heuristic value of the principle of reciprocal altruism led Irons to contend that there must be an adaptation for limiting free-riding within groups. As he argues that costly signals of costly commitment deter free-riding within groups and that religion is a powerful hard-to-fake signal of commitment, he claims that religion was selected for because it functioned to facilitate reciprocity and thereby promote intragroup cooperation.

A common disagreement among evolutionary theorists concerns the ubiquity of natural selection and the prevalence of adaptations (see Sober and D. S. Wilson, Unto 101). Is natural selection as pervasive as adaptationists assume? Do adaptive explanations hold an unwarranted monopoly over non-selective explanations of evolutionary change? Does adaptationism sometimes or often lead to the mistaken identification of non-adaptive traits as adaptations?

Gould and geneticist Richard C. Lewontin criticize adaptationist reasoning in “The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme” (1979). In this “oft-cited, oft-reprinted, but massively misunderstood classic” (Dennett 238), Gould and Lewontin argue that the adaptationist framework is problematic because adaptationists, like Voltaire’s Dr. Pangloss, assume that every trait is an optimal adaptation designed by natural selection for a specific purpose and, like Rudyard Kipling’s “just-so stories,” adaptationism leads to fantastic, unfalsifiable explanations of traits as adaptations.<sup>9</sup>

Gould and Lewontin criticize the adaptationist program because in a non-Darwinian manner it holds natural selection to be the omnipotent power and unrestrained ultimate cause of nearly everything: “This program regards natural selection as so powerful and the constraints upon it so few that direct production of adaptation through its operation becomes the primary cause of nearly all organic form, function, and behavior” (76). Without a doubt many traits were selected for because of their fitness-enhancing function; certainly natural selection mediates change and is an important cause of trait evolution. However, phyletic, developmental, and architectural constraints of the body plans of organisms, as well as environmental constraints, significantly restrict the possible changes to certain paths (Gould and Lewontin 85). Not all traits are functional or the products of natural selection. Due to a common misinterpretation of his theory, Darwin himself had to set the record straight that natural selection was not the only mechanism of descent with modification (Gould and Lewontin 81). In the final sentence of his introductory chapter in *Origin*, he states that, “I am convinced that Natural Selection has been the main *but not exclusive* means of modification” (Darwin *Origin* 69; emphasis added).<sup>10</sup> Gould and Lewontin claim that adaptationists do not consider alternatives to natural selection in the construction of traits, not because they think that non-selective evolutionary forces and constraints on natural selection do not exist or they are not aware of them, but instead because they circumscribe the domains of action of those other factors so narrowly that they do not have any actual force to affect evolutionary change (77). For Gould and Lewontin, architectural constraints are primary in explaining structures while adaptations are secondary (76).

Gould and Lewontin contend that adaptationists are non-Darwinian in their methodological practice of failing to give up the belief that all traits are adaptations. The Panglossian assumption that all present traits are adaptations directs adaptationists not only to propose evolutionary explanations of traits in terms of their selected function but also to continue to explain a trait in terms of adaptation even when the explanation lacks empirical support. Within the adaptationist program, evidence is not difficult to come by, as often the sole criterion for an adaptationist explanation is theoretical consistency with natural selection: an adaptive story for a trait is considered justified if it presents a coherent combination of phenotypic variation, heritability, and fitness such that the trait positively affects the fitness of its carrier. Anti-adaptationists claim that adaptationists do not consider non-adaptive possibilities and, on the rare occasions that they submit their explanations to the scrutiny of empirical tests, often these tests are not falsifiable. In the even rarer case that the results falsify the adaptationist explanation, rather than inspiring suspicion that a non-adaptationist explanation might be required, they replace one story with another adaptive story. If another story is not immediately available, then the adaptationist assumes that it will be forthcoming. Gould and Lewontin claim that adaptationists have been known to attribute this temporary ignorance to an imperfect understanding of where an organism resides and what it does rather than to consider alternative non-selective factors (78-79).

Panglossianism and just-so storytelling are not the only problems with adaptationism. Gould and Lewontin argue that adaptationists cannot assume that the detection of a common biological structure or process producing a certain fitness benefit is sufficient reason to infer that this benefit is the naturally selected function of this structure or process (84). Although fitness-enhancement and design specificity are characteristics of adaptations, not all traits that are universal and currently fitness-enhancing are adaptations. Gould and Lewontin contend that the universality of a trait does not necessitate that the trait is an adaptation. For instance, even if religion is “one of the universals of social behavior, taking recognizable form in every society from hunter-gatherer bands to socialist republics” (E. O. Wilson, Human 169), this is not enough to demonstrate that religion is an adaptation. Although evolutionary scientists studying

religion agree that religion is a universal phenotypic trait, they disagree as to whether or not it is an adaptation.

Like universality, Gould and Lewontin contend that the adaptationist tendency to focus on the current utility of an organic structure or behavioral strategy conflates the structure's current utility with its historical genesis (79, 85): "One must not confuse the fact that a structure is used in some way [...] with the primary evolutionary reason for its existence and conformation" (79). The key issue is "whether differential reproductive success historically influenced the design of a given trait, not whether the trait currently influences differential reproductive success" (Symons 150).<sup>11</sup> To prevent the conflation of current utility and historical origin, Gould and Vrba coined the term "exaptation" to designate a trait that is currently fitness-enhancing but was not historically constructed via direct selection for its present utility. They agreed with Williams that, as the fortuitous utility of exaptations is not the same as the historical role for which the trait evolved, the utility is not the function of the trait but rather its "effect" (Williams, 9; Gould, Structure 1231). Unlike the process of adaptation whereby a trait is "crafted" for its function, Gould calls the process through which a pre-existing trait is appropriated for a novel effect "cooption."<sup>12</sup> The trait being coopted was shaped previously by either the direct or indirect action of natural selection. In other words, the coopted structure with a current positive fitness effect (an exaptation) may have been originally an adaptation selected for a function unrelated to the one it presently serves (in this case the coopted structure is a "coopted adaptation") or it may have been a non-adaptation, with no function at all (in this case the coopted structure is called a "coopted byproduct") (Gould, Structure 1232-33; Buss et al. 539; see Table 1 below, p. 20).<sup>13</sup>

Table 1

Conceptual and Evidentiary Criteria for Evaluating Evolutionary Categories<sup>a</sup>

Differentiation Criteria	Categories			
	Adaptation	Exaptation		Byproduct or Spandrel
		Coopted Adaptation	Coopted Byproduct	
<b>Origin and Maintenance</b>	new trait directly selected for use	pre-existing adaptation selected for new use	pre-existing byproduct selected for use	incidental consequence of selection for adaptation
<b>Process</b>	adaptation	cooption	cooption	spin-off of adaptation
<b>Fitness</b>	correlated with fitness: solved adaptive problem at time of its evolutionary origin	currently correlated with fitness: solves current problem unrelated to original adaptation's evolution	currently correlated with fitness: solves current problem unrelated to original byproduct's evolution	not correlated with fitness
<b>Usage</b>	function	effect	effect	none
<b>Critical Features</b>	adaptation with function	pre-existing adaptation coopted for new effect	pre-existing byproduct coopted for effect	byproduct with no function and no effect
<b>Examples of Evolutionary Theories of Religion</b>	William Irons; David Sloan Wilson	Candace Alcorta and Richard Sosis	Scott Atran and Ara Norenzayan	Stewart Guthrie; Pascal Boyer

<sup>a</sup> This table is an extensively modified version of Buss et al.'s Table 2 in "Adaptations, Exaptations, and Spandrels" (545) that incorporates aspects of Gould's Table 11-1 in *The Structure of Evolutionary Theory* (1233) as well as my own additions. The main amendment other than my addition of examples is that, like Gould, who heeded Williams' recommendation, I contend that exaptations do not have 'functions' but rather have 'effects.' This table provides a concise summary of the evolutionary terms I employ throughout this thesis.

In the same vein as Gould and Lewontin, Williams warns that the concept of adaptation is “special and onerous” and should only be invoked when necessary and as a last resort (4, 5, 11). Despite being an adaptationist himself, Williams states that to limit the misuse of adaptation, adaptationists should not appeal to natural selection and adaptation in the following two cases.<sup>14</sup> First, do not invoke adaptation when non-selective lower-level explanations are available. All causes, besides natural selection, of currently useful features must be exhausted before a feature can be hypothesized to be an adaptation. The less arduous principles of physics, chemistry, cause and effect, and chance must be shown as inadequate to provide a complete explanation of the trait before one can hypothesize that it is an adaptation (Williams 11). Evolutionary psychologist Lee Kirkpatrick states that the burden of proof should be on those who make adaptationist claims (926): before positing that religion is an adaptation, the adaptationist should first rule out all other well-known ordinary adaptations that could account for religion.

Second, the adaptationist should not invoke adaptation when a feature is a “byproduct” of another adaptation (Williams 9).<sup>15</sup> A trait is a byproduct if it is not itself selected for but rather is a side-effect of another trait that was selected for (see Table 1 above, p. 20). Gould and Lewontin borrow the architectural term “spandrel” to denote a byproduct. A spandrel is a tapering triangular space formed as an unintended consequence, or byproduct, of a right angle intersection of a dome on top of two rounded arches. Although a spandrel may be used as a canvas for a creative ornamental effect in a cathedral, claiming that a spandrel is an adaptation that functions “to house the evangelists” ignores the fact that its existence is attributable only to the non-adaptive incidental architectural consequence of the primary decision to build a dome atop rounded arches in a cathedral. In the framework of evolution, a spandrel is a metaphor for a characteristic that is an inevitable offshoot of the construction of an adaptation rather than an adaptation itself. An adaptation has a selected for function, but if a byproduct has some beneficial usage (in this case, a coopted byproduct), it is not a ‘function,’ but rather a ‘fortuitous effect’ (see Williams 9).

Today’s adaptationists contest Gould and Lewontin’s and Williams’ sweeping criticisms of the adaptationist program. They claim that Gould and Lewontin and

Williams make it sound as though all adaptationists make all of the ‘errors’ mentioned above all of the time, but this is an exaggeration. Although these theoretical attacks on Darwinism have had an immense impact on the outside world, Dennett argues that in truth they are only minimal corrections to Darwinian orthodoxy presented in an over-dramatized manner with a distorting effect (263-64). Certainly some adaptationists are guilty of the charges laid by these anti-adaptationists. However, not only adaptationists comprise the guilty and the adaptationists that are guilty are a minority.

Adaptationists rebut Gould and Lewontin’s criticisms of adaptationism with a variety of arguments. One counter-argument contends that Gould and Lewontin are capitalizing on difficulties that exist in any approach to the study of humans and make it appear as though only evolutionary biologists make unwarranted leaps from hypothesis to conclusion, propose just-so stories, and fail to test hypotheses (Alexander 18). Any scientific field is potentially open to questionable methodologies employed by a minority of its scientists. The question is whether these bad methodologies are intrinsic to the science or are mistakes on the part of a few of the scientists who work within the science (Sober, “Six” 76). It is useful to separate the proposition of adaptationism from the adaptationists who employ it. For example, the mistake of some adaptationists in assuming that a trait’s current utility is the reason that it initially evolved is not inherent to adaptationist reasoning but is rather the misunderstanding of few adaptationists who use adaptationism.

Another criticism is that Gould and Lewontin focus on cases that are dismissed within evolutionary biology and ignore those that are exemplary of the best work in the area, i. e., those that cannot be faulted easily (Alexander 17).

[T]hey give the impression that their mission is to locate and emphasize the weakest parts of the arguments of individuals they are criticizing, then use those weakest parts to declare that the entire enterprise with which those parts are connected must be discarded – i. e., that the weakest components in an intellectual edifice can be used to prove that there is no core of accuracy and correctness, no possibility of a cumulative growth of knowledge (Alexander 18).

From the beginning adaptationism was awarded a bad reputation on the basis of a few cases that the majority of adaptationists themselves considered to be instances of bad adaptationist explanation. Moreover, those few bad cases that warrant Gould and Lewontin's attack are ones that were made long ago and not by modern evolutionists (Mayr, "Adaptationist" 327). In the field of current evolutionary inquiry, the majority of the scientific research enabled by the adaptationist program has yielded fruitful results. For this reason, it should continue to be employed.

Adaptationists question Gould and Lewontin's claim that it is wrong to assume that natural selection is the primary force of evolutionary change. Since the time of Darwin, Darwinians have routinely shunned the assumption that all traits are adaptations and that all coopted traits were once adaptations (Dennett 281): "The thesis that every property of every feature of everything in the living world is an adaptation is not a thesis anybody has ever taken seriously, or implied by what anybody has taken seriously" (Dennett 276). However, in a sense, all traits exist because of the action of natural selection, either direct or indirect. Gould's concepts of byproducts and exaptations (either coopted adaptations or coopted byproducts), even though introduced as alternatives to adaptations, still rely on natural selection (Mayr, "Adaptationist" 330): they all "invoke selection at some point in the causal sequence" (Buss et al. 546). In adaptations, selection is responsible for producing the adaptation; in coopted adaptations, selection explains the original adaptations as well as the cooption which modifies them to serve a new fitness-enhancing role; in byproducts, selection is required to produce the adaptations of which byproducts are incidental side-effects; finally in coopted byproducts, selection operates to construct the adaptations of which byproducts are incidental side-effects as well as to coopt the original byproducts to fulfill a fitness-enhancing task (see Table 1 above, p. 20). Because of the direct or indirect role of natural selection in the construction of a trait belonging to any evolutionary category, Gould and Lewontin's critique of adaptationists for assuming the operation of natural selection only holds if this assumption takes the specific form of the claim that whenever natural selection operates it leads to adaptations (Buss et al. 546). However, no legitimate forms of the adaptationist program make this claim.

Contrary to Gould and Lewontin's arguments, adaptationists argue that they do not employ unorthodox Darwinian principles. They acknowledge the existence and role of constraints and do not argue that traits are absolutely optimal. Good adaptationist reasoning does not neglect either the random fixation of genes or developmental constraints and considers both of them integral to any explanation (Dennett 270, 278). Despite holding that, "Natural selection has been the only important cause of most of the phenotypic traits found in most species," adaptationists recognize that "other forces" prevent selection from "creating" optimally designed adaptations (Sober, "Six" 72).<sup>16</sup> Although Gould and Lewontin criticize adaptationism for being a Panglossian paradigm, adaptationists do not contend that natural selection, like supernatural design in natural theology, must produce perfectly designed adaptations (see Mayr, "Adaptationist" 327). Numerous developmental, mechanical, and phylogenetic constraints limit optimal design, such as lag in time between a new adaptive problem and the evolution of the trait that solves it, the step-by-step nature of selection that prevents selection from going through "deep fitness valleys" to arrive at a better design on top of a "neighboring mountain," the lack of available genetic variation for selection to act upon, the costs involved in the construction of adaptations, and the necessity of coordination with other traits (Buss et al. 538). All of these constraints as well as various chance events prevent adaptations from being optimal designs in general, but they do not prevent them from being optimal designs relative to all of the available competing alternatives in that environment at that point in time.

Adaptationists object to Gould and Lewontin's accusation that adaptationism automatically leads the researcher to conclude that a trait is an adaptation. Adaptationism as a method of doing biology is separate from adaptationism as a claim about nature (Sober, "Six" 84). Asking "What would organisms be like if they were well adapted?" is not necessarily a commitment to the position that the trait under investigation is actually well adapted (Sober and D. S. Wilson, *Unto* 11; D. S. Wilson and Sober, "Reintroducing" 588). Questions such as "What is the function of a trait?" or "Why does this trait exist?" have great heuristic value for designing studies to examine the evolutionary foundation of the trait. Adaptationists urge evolutionists not to prejudge the correctness of adaptationist

explanations and to realize the importance and value of asking adaptationist questions like the ones mentioned above (D. S. Wilson and Sober, “Reintroducing” 607). “The point is not that reverse engineering is always *sufficient* to deliver the right solution but that it is always *necessary*” (Boyer, Explained 26). Only empirical tests determine whether an adaptationist prediction is a valid explanation.

Finally, Gould and Lewontin’s criticism that adaptationist explanations are not subject to empirical and falsifiable tests is mistaken. Adaptationist explanations are not accepted solely on the basis of theoretical consistency with natural selection. It is “excessive to regard functionalism in general as a giant compendium of just-so stories that are somehow immune to scientific inquiry” (D. S. Wilson, Cathedral 71; see also Mayr, “Adaptationist” 328). As all scientific hypotheses can be interpreted as stories at some level, the key issue in evolutionary explanations is whether adaptive hypotheses are formulated in a precise manner consistent with the principles of evolutionary biology, generate specific empirical predictions, and parsimoniously account for the data in a more compelling fashion than competing hypotheses (Buss et al. 543). For instance, for the most part, evolutionary psychologists employ the adaptationist framework but rigorous experimentation often leads them to posit byproduct explanations rather than adaptive explanations.

Despite Gould and Lewontin’s narrow appraisal of adaptationism, it has been and continues to be a productive research tool for investigating nature and predicting the properties of organisms (Sober and D. S. Wilson, Unto 11). As Dennett articulates so well in the following passage, adaptationism serves an invaluable function in all the life sciences:

Adaptationist reasoning is not optional; it is the heart and soul of evolutionary biology. Although it may be supplemented, and its flaws repaired, to think of *displacing* it from its central position in biology is to imagine not just the downfall of Darwinism but the collapse of modern biochemistry and all the life sciences in medicine. (Dennett 238)

### Evolutionary Theories of Religion

Several times in the above discussion of biological evolution and adaptation I have mentioned evolutionary theories of religion. An evolutionary explanation of religion is a claim about the evolutionary landscape of religion. It identifies how natural selection was involved in the origin of religion and whether religion has a function. David Sloan Wilson and Joseph Bulbulia have proposed independently that there are two dominant research strands in the evolutionary study of religion that are differentiated by whether or not the scientists contend that religion is an adaptation (D. S. Wilson, Cathedral 44; Bulbulia, “Review” 656). On the one hand, the evolutionary scientists who claim that religion is a biological adaptation generally work from within the scientific fields of behavioral ecology and evolutionary biology. The adaptive approach concentrates on answering the question, why religion exists, in terms of the function it serves, i. e., why it was favored by selection. Although religion may have numerous current fitness-enhancing effects, the function of religion is the one that provided religious organisms with a favorable status in the differential survival and reproduction of competing organisms in human evolutionary history.<sup>17</sup> Evolutionary scientists employing the adaptive approach to religion contend that religion was selected for in human evolutionary history because, relative to available alternatives, it best solved the recurrent fitness problem of within-group conflict. They claim that religion has the social function of facilitating social exchange between group members. This thesis examines behavioral ecologist William Irons’ adaptive story that religion was selected for because it solved the problem of free-riding that disrupted group cohesion at a time in our evolutionary history when group warfare was rampant and the largest, most united groups had the advantage in survival.

On the other hand, evolutionary scientists who adopt the byproduct approach and identify religion as an indirect product of natural selection predominantly work within the “cognitive science of religion” (CSR). This subfield of evolutionary psychology arose in the 1990’s to study empirically the cognitive resources required to acquire, generate, represent, and transmit religious concepts and actions.<sup>18</sup> Unlike adaptive evolutionary scientists who focus on fitness measures and function, byproduct theorists explain the

evolutionary landscape of religious belief and practice in terms of evolved psychological mechanisms. Evolutionary psychologist Lee Kirkpatrick explains that the focus is on psychological mechanisms rather than behavioral or psychological traits because they are the information-processing modules that are designed by natural selection to solve specific, recurrent, fitness-relevant problems by generating behavioral, cognitive, and emotional outputs: “It is these mechanisms, not their behavioral or psychological products, that are adaptations” (923). Byproduct theorists claim that there is no religion-specific mechanism. Instead, religion is the incidental consequence of the direct selection for various non-religion-specific mental modules. Because religion depends on pre-existent, mundane psychological mechanisms rather than a religion-specific mechanism, byproduct theorists claim that religion is not an adaptation. An examination of these ordinary psychological mechanisms explains the various recurrent features of religious belief and practice.

Religion’s engagement of only pre-existing ordinary mental modules prevents understanding religion as an adaptation but it does not necessitate that religion has no adaptive value. Byproduct theorists claim that as a side-effect of the selection for other traits, religion was not selected for any function and has no fitness-enhancing effect. However, a byproduct can become fitness-enhancing and therefore have adaptive value. The byproduct of religion can be coopted for fitness-enhancing effects such as alleviating debilitating anxiety or facilitating social cohesion; in either of those cases, religion would be a coopted byproduct. As a coopted byproduct, religion may be fitness-enhancing, but this role cannot be understood as responsible for the origin of the mechanisms that generate religious belief and practice (Atran and Norenzayan 718).

Coopted adaptation is another category of evolutionary traits that can explain how a trait can be grounded in ordinary mechanisms and yet have adaptive value. Here, it is not the byproduct of an adaptation that is coopted, but rather it is the adaptation itself that is coopted. Anthropologists Candace Alcorta and Richard Sosis claim that the determination of whether religion has adaptive value does not hinge on whether religion incorporates pre-existent mental modules, but rather it depends on whether there is evidence that these modules have adapted to solve new ecological problems (326-27).<sup>19</sup> If

so, then religion is a coopted adaptation: it is the product of the cooption of pre-existing ordinary mental systems rather than having arisen from a novel, functional mutation as claimed in the adaptive approach. This coopted adaptive approach to religion incorporates byproduct theory's emphasis on pre-existing ordinary mental modules and empirical evidence of these mundane adaptations but it maintains the argument of the adaptive approach and the coopted byproduct approach that religion is fitness-enhancing.

Evolutionary approaches are distinguishable not only on the basis of the proposed evolutionary origin of the mechanism generating the trait, but also in the manner in which they explain the trait. In other words, they differ in their consideration of why and how religion exists. Adaptive, coopted adaptive, and coopted byproduct approaches concentrate on identifying why religion exists in terms of what problem it solves while byproduct approaches focus on identifying how ordinary mechanisms in the evolved human mind create and maintain religious beliefs and practices. Nevertheless, adaptive explanations sometimes consider the 'how aspect' when they incorporate a treatment of the mechanisms that were selected for religion to carry out its selected function. Likewise, coopted adaptive and coopted byproduct theories must explain the non-religion-specific mechanisms that religion coopts, as well as the adaptive value of religion. Byproduct theories often address the 'why aspect' because part of the evolutionary explanation of how human cognitive architecture is active in religiosity demands an explanation of why the mundane adaptations of which religion is a byproduct were selected for in the first place. Evolutionary explanations that address both why a trait exists as well as how the evolved mind is active in generating the trait are commendable because the two approaches are complementary and both are required for a complete understanding of an adaptation (Mayr, "Cause" 1503; Irons, "Adaptively" 198; D. S. Wilson, Cathedral 170-177; D. S. Wilson, "Testing" 392).<sup>20</sup>

Because all evolutionary categories of traits, including religiosity, appeal to natural selection and adaptation at some point, they are subject to Gould's and Lewontin's as well as Williams' criticisms. However, because of the differences in their appeals, some are less open to those criticisms than others. In principle, byproduct theories do not claim that the trait in question is an adaptation with a selected-for function

and therefore are not as susceptible to such criticism as adaptive or even coopted claims. Evolutionary psychologist Lee Kirkpatrick claims that the byproduct approach is “the more theoretically conservative position” because, as Williams argues, the burden of proof is on those scientists who claim that religion has adaptive value (926). To identify a feature as an adaptation, a scientist must demonstrate that it cannot be accounted for by other well-known mechanisms and that it has design specificity. Unlike adaptive explanations, in relinquishing functionality, byproduct explanations are not Panglossian, cannot confuse current trait utility with the historical genesis because they do not posit the trait to have adaptive utility, and do not invoke adaptation because the lower-level explanation of byproduct suffices. As they do not argue for a religion-specific mechanism, they do not need to demonstrate fitness-enhancement or design specificity in religion. One of the most difficult characteristics of religion to explain as adaptively advantageous to the fitness of the believer is the high cost of religious belief and practice. How does bodily mutilation and sacrificing much-needed food sources to gods increase the survival and reproduction of the believer? This uneconomic side of religion seems to beg for an evolutionary explanation of religion that is non-adaptive and non-functional. However, as I discuss in the next chapter, Irons turns the situation upside down and argues that it is the costliness of religious ritual that was selected for because it makes signals of commitment to the community easy to monitor and more reliable and facilitates the formation and maintenance of large, well-united groups of nonrelatives. For Irons, religion is economic despite its costs because the benefits of sociality outweigh the costs of religious signals.

Bulbulia remarks that byproduct theories are desirable because appealing to several mundane cognitive mechanisms to explain religious thought is actually to propose a simpler cognitive system than to postulate a religion-specific psychological mechanism (“Religious” 4). It is easier to demonstrate empirical evidence of adaptation of a domain-specific ordinary psychological mechanism than of a multifaceted and complex cultural phenomenon like religion. A psychological mechanism for agency detection or intuitive ontology seems easier to define and empirically test than a category like religion, which adaptive theorists as well as most scholars of religion explicitly acknowledge is difficult

to pinpoint precisely (Irons, “Hard-to-Fake” 296; D. S. Wilson, Cathedral 220-22).<sup>21</sup> Boyer concurs that explaining religion as a byproduct rather than an adaptation is “more prudent and empirically justified” (“Natural” 31).

In the following chapter, I focus my analysis on behavioral ecologist William Irons’ adaptive explanation of religion. In the third chapter, I consider the byproduct approach and the coopted adaptive approach to explaining the evolutionary foundation of religion. My intention overall is to determine whether Irons’ argument that religion is an adaptation is justified by his own theoretical evidence. To carry this out, I examine his arguments for internal coherence and for consistency with the concept of biological adaptation, and finally I assess their fit within the byproduct and coopted adaptive frameworks of explanation.

---

Chapter One Notes

<sup>1</sup> I point out that there is no standard for the use of the terms ‘organism’ and ‘individual’ within evolutionary theory. I use the terms ‘individual’ and ‘organism’ in the same manner as evolutionary biologist David Sloan Wilson and philosopher Elliot Sober. In “Re-introducing Group Selection to the Human Behavioral Sciences” (1994), Wilson and Sober state that the word individual refers to a single flesh and blood creature whereas the word organism refers to any biological entity whose parts have evolved to function in a harmonious and coordinated fashion (see note 1 on page 606).

Stephen Jay Gould considers ‘individual’ to denote the generalized unit of selection at any hierarchical level while ‘organism’ refers to a single organic body (600-01); he is aware that Sober and Wilson designate the reverse meanings for these terms. William Irons and Richard Dawkins use the terms individual and organism to refer to a single creature.

<sup>2</sup> In Origin of Species, Darwin made it a point to specify that the survival of the individual carrier of a particular trait as well as its success in leaving progeny are important in understanding the struggle for existence and calculating fitness (116).

<sup>3</sup> Selection only acts on those phenotypic features that have a significant effect on the individual’s reproduction in a specified environment (Tooby and Cosmides 52). Darwin states that, “variations neither useful nor injurious would not be affected by natural selection” (Origin 131). Pressures in the individual’s environment determine which traits are reproductively significant; for example, as female peacocks select male mates with flamboyantly large and colorful tails, the male peacock tail is a reproductively significant trait.

<sup>4</sup> Another means through which an organism can increase its fitness without necessarily having more babies than others is make sure that the babies it does have are successful in reproduction. This is why in my discussion of the three tenets of evolution I state that relative fitness refers to having offspring that themselves have offspring. “Unbridled fecundity” is seldom the means to maximize personal reproductive success (Williams 161). Dawkins points out that bearing children may be half the fitness battle because for the offspring in turn to survive long enough to reproduce successfully, in many species, there must be at least some care invested in the offspring by the parent(s) (109). Various mixes of child-bearing and child-caring are evolutionarily stable. Williams states that, “selection will adjust the amount of immediate reproductive effort in such a way that the cost in physiological stress and personal hazard will be justified by the probability of success [in rearing an offspring to a reproductive age]” (172). See Irons, “Morality” 56-58 for a further discussion of the “absurdity” of the theory of maximal reproduction.

<sup>5</sup> Adaptation can be understood as both an evolutionary process and a product of natural selection. Here, it designates the evolutionary process in which natural selection modifies the traits in a population to better suite the demands of the local environment. From this point onward, I use the term ‘adaptation’ primarily to define the phenotypic trait natural selection is selecting for. In other words, adaptation refers to the functionally designed historical end-product of the process of evolution. See Schmitt and Pilcher 643 for a more thorough discussion of the term adaptation as a process and product.

<sup>6</sup> Like Williams and Gould, for the remainder of this thesis I reserve the term ‘function’ to designate the fitness-enhancing operation of a useful trait built by selection for its current use (see Gould, Structure 1230). It should be noted that Gould is inconsistent with his use of the terms ‘function’ and ‘effect.’ For example, in “Exaptation: A Crucial Tool for Evolutionary Psychology” (1991), he claims that exaptations, like adaptations, have functions. However, in his seminal work, The Structure of Evolutionary Theory (2002), he makes it clear that function is restricted to adaptation whereas effect is reserved for exaptations. See Buss et al. for a discussion of Gould’s inconsistency in his use of several evolutionary terms.

<sup>7</sup> The claim that adaptations are solutions to *past* environments is an issue of contention (see Lloyd and Feldman's "Evolutionary Psychology: A View From Evolutionary Biology" (2002) and Irons' "Adaptively Relevant Environments Versus the Environment of Evolutionary Adaptedness" (1998)). The exact relationship between adaptation and environments is not as clear as I present it. On the one hand, evolutionary psychologists such as evolutionary psychologist Donald Symons argue that the time period and environment (the environment of evolutionary adaptedness or EEA) in which humans biologically evolved was the Pleistocene (143). Symons claims that as the majority of human evolution occurred in the Pleistocene, most human adaptations are adaptations to the selection pressures that plagued the fitness of small groups of nomadic hunter-gatherers ancestors who foraged in the African savannah between 1.81 millions years ago and 11 550 years ago. As our environment is no longer the Pleistocene, most adaptations to this past environment are no longer reproductively advantageous in the novel environment we currently inhabit. On the other hand, others argue that to understand an adaptation, we need not necessarily focus on all of the conditions of the Pleistocene. Irons contends that what is important is the evolved adaptation's "adaptively relevant environment" (ARE), i. e., "those features of the environment that the mechanism must interact with in order to confer a reproductive advantage" (198). If those features of the environment have not changed since the Pleistocene, then the adaptation is still adaptive in the modern environment and measuring current fitness can shed light on the evolutionary existence of the adaptation.

<sup>8</sup> Williams set out these criteria in Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought (1966) and these other evolutionary scientists refer to him. For example, Cosmides and Tooby cite Williams as stating that, "Standards for recognizing special design include factors such as economy, efficiency, complexity, precision, specialization, and reliability, which – like a key fitting a lock – render the design too good a solution to an adaptive problem to have arisen by chance" (165).

<sup>9</sup> Evolutionary biologist Massimo Pigliucci and philosopher Jonathan Kaplan's Box. 1 provides a concise summary of the issues in the adaptationist program that Gould and Lewontin attack.

<sup>10</sup> See also the concluding chapter in a sixth edition of Origin. Here, Darwin points out that the statement was present in the first edition of Origin in an attempt to dispel the misrepresentation that he attributes the modification of species exclusively to natural selection.

<sup>11</sup> Focusing on current utility often leads to measurements of reproductive differentials to demonstrate adaptation. However, Symons states that correlations between individual variation in trait expression and current reproductive success do not illuminate adaptations effectively nor unambiguously for the following reasons: an adaptation designed to solve fitness threats may correlate negatively with fertility; a genetically fixed adaptation may vary with environmental variation such that its expression is unlikely to correlate with reproductive success; over evolutionary time, current seemingly insignificant or undetectable correlations may have great selective importance; trait covariance with reproductive success may be caused because both are correlated with a third variable; a trait may correlate positively with successful propagation in current conditions because its effect is different than its intended evolved function; and an adaptation is not necessarily presently adaptive (148-49). As Symons is one of the foremost advocates of understanding the environment of all human adaptations as the Pleistocene, it is not surprising that he holds that reproductive consequences measured in today's individuals are irrelevant to determining whether a trait is an adaptation and that the fitness of human behaviors must be measured in hunter-gatherer societies.

For an example of research in evolutionary theory that decouples current utility and historical genesis, see D. S. Wilson and Sober's discussion of Hutterites as a human example of extreme group-level functional organization in "Re-introducing Group Selection to the Human Behavioral Sciences" (1994). More specifically, they cite Symons in stating that they plan to "focus more on the design features and what they would have meant in ancestral environments than on the present day consequences of the design features" ("Re-introducing" 603).

Although I do not consider this category of evolutionary theories of religion in this thesis, Wilson points out that one theory explicitly stands by this distinction with its claim that religion was an adaptation to past environments but is maladaptive in modern environments (see Wilson, Cathedral 45).

<sup>12</sup> Gould actually calls the term “cooptation,” but in the literature it is typically referred to as “cooption.” For consistency I employ the term “cooption.”

Buss et al. state that the causal mechanism of adaptation and cooption is natural selection but add that an existing cognitive and motivational mechanism may be responsible for the coopting (542).

<sup>13</sup> Anthropologist Scott Atran differentiates adaptations, byproducts, and exaptations (more specifically, coopted byproducts) in the following manner: “Evolutionary adaptations are functional biological designs naturally selected to solve important and recurrent problems in ancestral environments, such as teeth for masticating food. Evolutionary by-products are necessary concomitants of adaptations that were not initially selected to have any direct utility, such as the whiteness of teeth (owing to much calcium and little bacteria). Nevertheless, by-products can acquire or co-opt functions [effects] for which they were not originally designed, such as white teeth as a sign of health for attracting mates” (Gods 265). In such a case, white teeth that increase fitness in mate selection become a coopted byproduct.

<sup>14</sup> Although Williams states that adaptation is an onerous concept, he was an advocate of “organic adaptations” to the extent that one purpose of his book Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought (1966) was to convince the reader that the study of the general nature of adaptation is important enough to warrant a special branch in biology (258). Teleonomy would try to answer the question, “What is the function of this biological phenomenon”? It appears that Williams fell into the very trap of the adaptationist program that Gould and Lewontin criticize (see Gould, Structure 550).

<sup>15</sup> Not everyone agrees with Williams, Kirkpatrick, and Gould and Lewontin that studies of the evolutionary foundation of traits should begin by assuming that the trait in question is explicable in terms of non-selective forces. Ernst Mayr argues that the evolutionist must first attempt to explain all biological phenomena as products of natural selection and only after such attempts fail resort to explaining them in terms of byproducts. In other words, Mayr suggests that adaptationist reasoning should assume that the trait is an adaptation until empirical evidence suggests otherwise. In fact, he states that even if one selectionist explanation of a feature is discredited, the evolutionist must test novel adaptationist solutions before resigning to explain the trait as a product of chance. Mayr justifies this reasoning with the argument that no one can “deduce the probability of causation by chance,” but one can deduce the probability of causation by selection (“Adaptationist” 326).

<sup>16</sup> To prevent a skewed exposition of Sober’s arguments I must point out that despite acknowledging the existence of non-selective forces, Sober asserts a monistic view of adaptationism: natural selection is the only important cause of trait evolution to the extent that non-selective processes may be “safely ignored” (“Six” 84). He illustrates his reasoning with an analogy to Newtonian mechanics (“Six” 73-74): the Earth’s gravitational force accelerates free objects towards the Earth’s surface at  $9.81\text{m/s}^2$ . As the Earth’s surface is not in a vacuum, the force of gravity on any object competes with air resistance. If the object is a bowling ball, the force of air resistance is negligible relative to the force of gravity. However, if the object is a feather, the force of air resistance is significant relative to the force of gravity. Depending on the object, either a pluralistic perspective or monistic perspective is required. In the case of the bowling ball, a monistic perspective is acceptable. Sober argues that organisms are like bowling balls in the sense that the force of non-selective forces in the evolution of an adaptive trait is negligible relative to the force of natural selection. He considers a monistic natural selection perspective acceptable when studying the evolution of traits.

<sup>17</sup> I use the more inclusive term ‘organism’ rather than ‘individual’ because the adaptive approach can be broken down into claims that religion is an adaptation at the individual level and claims that religion is an

adaptation at the group level. In this thesis I am primarily concerned with William Irons' individual selection theory of religion. However, David Sloan Wilson offers a group selection theory of religion which I discuss in Chapter Three not in order to present his group-level argument but rather to consider his suggestion that religion may be both a biological and cultural adaptation.

<sup>18</sup> Although the cognitive science of religion is only two decades old, byproduct theories of religion are not. Charles Darwin, the founder of natural selection, ruled against an adaptive approach to religion well over one hundred years ago. Like today's cognitive scientists of religion, Charles Darwin proposed that religion evolved as a byproduct of selection rather than a target of selection (Darwin, Descent 96-99; Bulbulia, "Review" 657-58). His evolutionary appraisal of religion in The Descent of Man is limited to three pages and he presents his argument in a less refined manner than those put forth by current cognitive scientists of religion. Nevertheless, Darwin makes it clear that religion did not evolve because it enhanced reproduction. He posits that there is no special or independent domain of religious thought: religious cognition is an amalgam of several ordinary elements including love, dependence, submission, fear, reverence, gratitude, and hope for the future. It manifests itself differently depending on cultural variables and may be entirely lacking. Although the complexity of religious experience requires moderately high intellectual and moral faculties, including well-developed imagination, curiosity, and reason, the tendency to animate the world with spiritual essences is shared by animals. Religion is the byproduct of ordinary animal cognition that serves no adaptive function: "These miserable and indirect consequences of our highest faculties may be compared with the incidental and occasional mistakes of the instincts of the lower animals" (Darwin, Descent 99).

<sup>19</sup> Alcora and Sosis imply that if a trait is built on pre-existing psychological mechanisms, then it is not an adaptation *per se*. However, as such a trait can be fitness-enhancing, it can have adaptive value. In other words, all fitness-enhancing traits (adaptations and exaptations) have adaptive value, but only those that originate because of their own function are adaptations (see above Table 1, p. 20; see note 1 to the Introduction).

<sup>20</sup> This discussion introduces the distinction between 'ultimate causation/explanation' and 'proximate causation/explanation' in evolutionary theory. Ultimate causation concerns why traits are favored by natural selection whereas proximate causation addresses how the trait works in the sense of the nature of the mechanisms responsible for its generation and utility (see Mayr, "Cause"). I chose not to use these terms because proximate explanation typically refers to a detailed treatment of the chemistry and physics involved in the mechanisms that generate behaviors while ignoring the evolutionary reason for the existence of the mechanisms in the first place. In light of this fact, I contend it would be too simplistic to label byproduct theories as proximate explanations and adaptive theories as ultimate explanations. Although byproduct theorists focus their research on identifying the psychological mechanisms active in religious thought and behavior and explaining how they are involved, they do consider the evolutionary origin of these psychological mechanisms and hence consider ultimate causation within their proximate explanations. Moreover, byproduct theorists do not typically delve into neuroscience and genetics in order to provide a thorough explanation of the proposed psychological mechanisms. Because neither the adaptive nor the byproduct evolutionary approach fit as a perfect example of proximate or ultimate explanations, I chose not to use these terms. However, adaptive explanations tend to answer primarily why religion exists and secondarily how it is generated while byproduct explanations tend to answer primarily how religion is generated in terms of cognitive mechanisms and secondarily why it exists in terms of natural selection.

<sup>21</sup> I elaborate on the difficulty associated with defining the 'fuzzy' category of religion in Chapter Three.

## Chapter Two: Irons' Adaptive Theory of Religion

Religions are, I suggest, built up over many generations and consist of numerous symbols that create, express, and maintain commitments that are essential to human social life.

Irons, "Inquiry" 367

In this chapter, I outline William Irons' individual-level evolutionary explanation of the origin and function of religion. Irons argues that religion offered its believers a selective advantage in human evolutionary history ("Inquiry" 358). I describe and analyze his adaptationist contention that religiosity reduced within-group conflict in an evolutionary environment of human warfare in which natural selection favored large well-united groups. Irons explains that religion is a hard-to-fake signal of commitment to one's community and its code of ethics that facilitates within-group cooperation. I refer to Joseph Bulbulia's and Richard Sosis' costly signaling models of religion in areas of Irons' theoretical explanation of religion that I identify as underdeveloped and incomplete.<sup>1</sup>

In order to analyze the quality of Irons' theoretical evidence for his adaptive claim of religion in the next chapter, in this chapter I trace Irons' development of his arguments. Irons holds that his evolutionary theory of religion is "a logical extension" of his evolutionary theory of morality ("Evolution" 392; "Inquiry" 358, 360). As such, in the first section of this chapter, "The Evolution of Morality," I provide an exposition of the social strategies upon which Irons grounds his evolutionary theory of morality.<sup>2</sup> I review Irons' argument that zoologist Richard D. Alexander's identifies morality, or "indirect reciprocity," as the solution to the fitness-relevant problem of conflict among nonrelatives in groups. Also, I look at Irons' assessment of "punishment" and "hard-to-fake signals of commitment" as the social strategies that facilitate indirect reciprocity in the establishment of large well-united cooperative groups.

In the second section, "The Evolution of Religion," I demonstrate how Irons' evolutionary theory of religion combines indirect reciprocity and the game theory

concept of commitment to explain religion's origin and function. The repetition in these sections mirrors the duplication in Irons' theories of morality and religion that I explore in the next chapter. My discussion and analysis of Irons' explanation of religion is divided into two subsections: "Belief" and "Ritual." In "Belief," I address Irons' claims regarding what religion is a commitment to and how belief in an unseen order motivates commitment. In "Ritual," I focus on Irons' arguments concerning religious signaling of commitment.

### The Evolution of Morality

Irons argues that social strategies other than kin selection and reciprocity had to evolve for humans to live in large groups. Irons concurs with Alexander that the primary reason for the evolution of human sociality was defense against fellow human predatory groups ("Shaped" 2, 6).<sup>3</sup> Alexander explains that in human evolutionary history, group-living offered individual group members adaptive advantages in terms of shared food, collective hunting, and defense against and war with enemies. He claims that because larger, better-united groups were more successful on average in warfare among groups, such groups were selected for.<sup>4</sup> However, larger groups were difficult to hold together because the force of selection among individuals created conflicts among group members (Irons, "Evolve" 67). As groups became larger and individuals had to interact socially with increasing numbers of nonrelatives, kin selection, i. e., the selection for altruism among genetic relatives (Hamilton), no longer sufficed to maintain cooperation. Reciprocal altruism, i. e., the selection for altruism toward those who reciprocate (Trivers), also could not maintain cooperation because despite all group members benefiting from mutual cooperation, individual group members benefited most by "free-riding," i. e., defecting when his social exchange partner cooperates (Sosis, "Badges" 5). Free-riding was always a problem in maintaining cooperation, but when the groups were small, group members could easily monitor one another's tendencies in social exchange. However, as groups increased in size and became less composed of relatives, free-riding became difficult to police efficiently. The risk of free-riders, the need to monitor free-riders, and the existence of outbreaks of internal conflict hampered direct exchanges of

aid (Irons, "Evolve" 67). Irons contends that the social mechanisms of "indirect reciprocity," "punishment," and "commitment" facilitated the formation and maintenance of large social groups of nonrelatives (Irons, "Co-Creator" 777).<sup>5</sup>

Alexander asserts that indirect reciprocity, or a moral system, is the adaptive solution to the problem of within-group conflict that is inherent to the formation of large, unified groups of nonrelatives (Irons, "Evolve" 67). Indirect reciprocity emerges as the result of direct reciprocity operating in a society where repeated interactions are common and at least some members keep tabs on who reciprocates and who cheats in social exchange (Alexander qtd. in Irons, "Shaped" 3-4). Irons agrees with Alexander that in human evolutionary history some individuals became "interested audiences" who monitored each others' tendencies in various repeated social exchanges in order to limit altruistic behavior toward those who played nice ("Evolution" 384). Through observation individuals could gather large amounts of social information about others without having to risk personally interacting with cheaters ("Hard-to-Fake" 296).

According to Alexander, the introduction of indirect observation and reputation into reciprocity facilitated the establishment of within-group cooperation in human evolutionary history (Irons, "Evolution" 384). In societies based on indirect reciprocity, the cooperative behavior of members was not as threatened by free-riders as it was in direct reciprocity because the cooperator's fitness advantage lay in the effect of her own good social reputation rather than in the direct reciprocation of aid from the benefactor (Irons, "Evolve" 67-68).<sup>6</sup> The good social reputation of reciprocation had the effect of drawing observers into reciprocal social exchanges. It also had the effect of drawing fellow members with good reputations into allied subgroups in which direct exchanges of favors could proceed unhindered over long periods of time (Irons, "Evolve" 67).<sup>7</sup> Alexander contends that very early in the evolution of the hominid line reputation became important as it made possible moral systems which limited the ways in which members of the same community who were not relatives nor direct reciprocators could damage other group members' reproductive interests (Irons, "Evolve" 68).

Alexander states that as a moral system, indirect reciprocity further limits intragroup conflicts by defining certain behaviors as right or wrong, where right

coincides with actions that do not do “too much damage to others,” and enforcing this behavioral standard in the form of community-wide rules (Irons, “Evolution” 388; “Evolve” 70). Rules clarify expectations and motivate those who may not be sufficiently motivated by moral sentiment alone to maintain a pattern of cooperation (Irons, “Evolution” 388). Not only are there rules, but also the rules have consequences. Alexander claims that humans police one another in order to reward good behavior and chastise bad behavior. Those who are observed to follow the rules (good reputation) are rewarded with praise, support, and help, while those who are witnessed to break the rules (bad reputation) are punished with criticism and ostracism (Irons, “Inquiry” 360). Appealing to Boyd and Richerson’s research, Irons points out that the strategy of punishing rule-breakers and those who fail to punish rule-breakers “can easily create strong pressure on people to conform to rules even when the rules are costly to follow” (“Hard-to-Fake” 297). In other words, observing the behavior of others and then “being nice to nice people and being nasty to nasty people” (Irons, “Inquiry” 360) motivates conformity to the moral rules.<sup>8</sup>

Irons summarizes his conception of morality in a three part definition: morality is 1) the biologically evolved human propensity to make judgments of what constitutes right behavior that is to be admired, rewarded, and imitated, and wrong behavior that is to be frowned upon, punished, and not imitated; 2) the rules and systems that codify and clarify these judgments; and 3) the tendency to be influenced by these rules and by others’ judgments of our conformity to these rules (“Evolve” 49-50; “Evolution” 375). The social strategies of indirect reciprocity and punishment are integral to solving the problem of free-riding in our evolutionary history: “This pattern of monitoring, rewarding, and punishing helps to prevent free riding and makes the formulation of large, well-united groups possible” (Irons, “Inquiry” 360).

Although moral monitoring and the rewards and punishments of morality help to form larger cohesive groups, Irons acknowledges that these social strategies serve their function within limits. Indirect reciprocity and punishment are not entirely effective at reducing within-group conflict because reputation-monitoring is imperfect and free-riders fake good reputations. Irons appeals to anthropologist Robert Dunbar’s argument that

there is a threshold of approximately 150 members to a group in which third-party observation can effectively identify and lead to the punishment of free-riders. Beyond 150 members, group members have difficulty monitoring one another even with networks of allied groups (Irons, “Inquiry” 362; “Co-Creator” 785). Independently from Dunbar, Alexander explains that large groups are difficult to hold together with indirect reciprocity because within-group selection for free-riding led to the evolution of methods to hide immorality when its detection is probable. Deception allows group members to fake good reputations and reap the rewards of gaining allies without incurring the actual costs of having a good reputation (Irons, “Evolve” 69): “When faced with the conditions of collective action, the incentive to falsely claim that one will cooperate is especially high since individuals can achieve their greatest gains by refraining from cooperation when others cooperate” (Sosis, “Hutterites” 93). Free-riding and deception undermined cooperation and precluded indirect reciprocity from operating optimally in our evolutionary history once group size superseded 150 members.

Irons appeals to economist Robert Frank’s discussion of commitments and hard-to-fake signals of commitment in game theory to explain how free-riding and deception can be prevented from hindering the benefits of indirect reciprocity, i. e., creating and maintaining cooperative groups of 150+ nonrelatives (Irons, “Inquiry” 358, 362). As signals of commitment play a central role in Irons’ theory of religion, I use the remainder of this section to outline the nature of commitments and signals of commitment: I look at what a commitment is, what effect a commitment has, why it is advertised, and how it is hard to fake.

Irons claims that Frank’s concept of commitment “corresponds to” Alexander’s notion of reputation (“Evolution” 385) such that, like a reputation, a commitment is a consistent behavioral tendency in social interaction. As mentioned above in the context of good reputations, although a commitment is a vow to act on the basis of a principle rather than calculations of net benefit and hence does not fit the opportunistic model, Frank emphasizes that it can serve one’s interests by influencing the behavior of others (Irons, “Inquiry” 363). Irons explains this paradoxical fact:

An inflexible commitment to behave in a particular way can serve one's interests in the majority of cases even if in particular situations the behavior is contrary to self-interest. Such a commitment does this by changing what others expect from us and thereby changing the way they behave toward us. ("Hard-to-Fake" 292)

The strong emotions that observers and gossipers attach to their moral judgments of others' commitments guide their behavior (Irons, "Inquiry" 361). Demonstrating that one's actions are not governed by cost-benefit calculations increases one's self-interest because it persuades third-party observers to enter into cooperative ventures ("Inquiry" 363). A promise to give aid indiscriminately or to abide consistently by a community's moral rules is a commitment that appeals to people who watch others, who are trying to avoid being cheated by free-riders, and who seek allies who can be trusted to engage honestly in cooperative exchanges. When members of a community share the same commitment, group and individual interests are not in conflict, trust reigns, and individual group members benefit from reciprocal interaction with one another ("Hard-to-Fake" 292). Although the costs of committing to cooperation in a group of cooperators are balanced by the benefits of sociality (Bulbulia, "Religious" 23), it is important to note that within a group of cooperators the social strategy that consistently brings the greatest benefits is defection, i. e., free-riding. In other words, the temptation to defect persists because it offers a greater immediate benefit than committing to cooperate.

In order to bring cooperative people together and prevent deception from disrupting trust and mutual aid within the group, commitments need to be easily observable and highly credible (Irons, "Co-Creator" 780-81). For commitments to entice others into cooperative exchanges, individuals must communicate them to fellow group members. Irons asserts that explicitly advertising one's commitment is an effective strategy in social communities where social reputations are monitored and rewarded or punished ("Hard-to-Fake" 297): "In a social environment in which indirect reciprocity is an important strategy, it would pay individuals not just to passively wait for other members of their group to discover that they are 'nice' people but rather to actively send out signals of their niceness" (Irons, "Co-Creator" 778). In fact, Irons claims that as the

social strategy of signaling commitment increases the size of cooperative groups, it is “a straight-forward extension of the idea of indirect reciprocity” (“Co-Creator” 778).

However, not any signal of commitment will motivate others to engage in social exchange if signals can be faked. When deception is prevalent or even possible in the social environment, then it is in the honest individual’s best interest to demonstrate that the signal is genuine: “If we are being continually observed and sized up and rewarded and punished for our behavior, what would make more sense than to send credible signals of commitments that will encourage others to act in our interests” (Irons, “Hard-to-Fake” 297)? Signals must be shown to be credible because free-riders are willing to pay the price to imitate signals of commitment to get into social exchanges when the costs of doing so are smaller than the benefits of a successful defection. The higher the costs of believably signaling the commitment, the less likely a free-rider would bother mimicking the signal to dupe a group member into a one-sided social exchange. Irons gives an example that explains this point: “if entering the group is allowed only after one does something costly such as putting up with a period of hazing or paying a large sum of money, the freeloader’s chances of recouping the costs of entrance before being discovered and ejected are slim” (“Co-Creator” 781). In order for the signals of commitment to be meaningful and useful, Irons concurs with Frank to claim that they must be “hard to fake” such that the costs are greater than the benefits. The costliness of a signal of commitment differentiates between sincere cooperators and deceptive free-riders. Irons concurs that, “the costlier the signal of commitment the less likely it is to be false” (Irons, “Hard-to-Fake” 298).<sup>9</sup> Generating and identifying hard-to-fake advertisements of commitment makes deception difficult, keeps free-riders out of the community, raises the level of trust among group members, and limits within-group conflict.

The propensities claimed by Irons to have evolved in the context of morality can be summarized as follows. Indirect reciprocity evolved because it functioned to limit internal group conflicts by instituting moral rules that define behaviors as right or wrong and encouraging group members to monitor one another for conformity to the community’s moral rules. Punishment added the element of consequence to one’s choice

regarding conformity to the moral code of the community: simply stated, reward for right behavior (good social reputation) and punishment of wrong behavior (bad reputation). Although punishment helped abate free-riding once it arose, Irons implies that the social strategy of costly signals of commitment served the preventative function of decreasing the incidents of free-riding in the first place. As free-riding is only “free” when there are no costs or when the benefits outweigh the costs, the requirement of costly communications of reputations made attempting a free-ride in a community based on indirect reciprocity and punishment not worth the trouble.

### The Evolution of Religion

Irons’ evolutionary explanation of religion is that, “religion facilitates intragroup cooperation by serving as a commitment and a hard-to-fake sign of commitment” (Irons, “Inquiry” 364). Religion is not just a hard-to-fake signal of commitment; it is the most “elaborate,” “effective,” and “powerful” cultural signal of commitment (Irons, “Hard-to-Fake” 293; “Inquiry” 366). Irons claims that religion was the functional solution to the fitness-relevant problem of within-group conflict stemming from deception in human evolutionary history. More specifically, religion increased within-group cooperation in groups of 150+ nonrelatives by clearly advertising and authenticating people’s reputations, thereby making the observation, reward, and punishment of reputations easier and more reliable.

As a definition is a “theory in miniature” (Saler, Conceptualizing 85), I present Irons’ definition of religion and examine its components to analyze his claim that religion is an adaptation. Irons offers a “working definition of religion for inquiries into the evolutionary origin and function of religion” that he claims is a composite of Clifford Geertz’s and William James’ definitions of religion and which he argues comes as close as possible to capturing the “central” elements of religion (“Hard-to-Fake” 296). In “Religion As a Hard-to-Fake Sign of Commitment” (2001), Irons asserts that, “*The common element of religion cross-culturally is a belief that the highest good is defined by an unseen order combined with an array of symbols that assist individuals and groups in ordering their lives in harmony with this order and an emotional commitment to*

*achieving that harmony*” (294). In “Why are Humans Religious: An Inquiry into the Evolutionary Origin of Religion” (2001) he states that, “religion is a commitment to bring one’s life into harmony with an unseen order, a commitment that defines and justifies the basic moral order of a community” (365).

Irons makes several claims about religion in these two expressions of his definition: in a religious worldview an unseen order defines the highest good, the human commitment to harmonize life with this unseen order is the supreme good, this commitment is expressed through a system of symbols, and this commitment defines and validates within-group morality. In this section, I explore how these claims support Irons’ functionalist argument that religion solves the problem of free-riding. I describe and analyze the elements of belief in an unseen order, the highest good, and ritual in light of Irons’ argument that, “religion is what has motivated basic morality” (Irons, “Inquiry” 365). I look at the role of the unseen order and the highest good in religious commitment under the subsection of “Belief” and I consider the nature and role of signaling religious commitment under the subsection of “Ritual.” In the course of discussing Irons’ account of religious belief and ritual, I address Irons’ explanations of what religion is a commitment to, how it signals this commitment in a hard-to-fake manner, and how religious commitment and signals of commitment are different from secular commitment and signals of commitment. As I find Irons’ theory lacks an explanation of the role of the unseen order in motivating costly signals of conformity to the moral standards of the community, I supplement it with elements from Joseph Bulbulia’s and Richard Sosis’ costly signaling theories of religion.

## Belief

Irons recognizes that belief plays an “essential” role in religion (“Inquiry” 364). According to Irons’ definition of religion, religious believers believe in four things: the existence of an unseen order, the unseen order as the source of moral rules, the unseen order as the source of the greatest good, and harmonizing one’s life with the unseen order as the achievement of one’s own greatest good.

Irons argues that in grounding a community's moral code on an unseen order, religion reinforces commitment to the moral code ("Inquiry" 365-66). For people who seek to form communities based on commitment, "basing our most sacred principles" on an unseen order is easy to "invent" and to "learn" ("Inquiry" 365; "Hard-to-Fake" 295). By "to base" I presume he means that the unseen order is seen as the source of the moral rules. Although Irons states that, "*somehow* people can best strengthen their commitments by appealing to something beyond the seen" ("Inquiry" 367; emphasis added), earlier in the same article he provides an explanation for why humans appeal to an unseen order: it "strengthen[s] the commitment that is religion by removing it from the world of empirical examination, which removes it further from the opportunistic model" ("Inquiry" 365). Irons elaborates that understanding an unseen order as the source of a community's moral code facilitates within-group cooperation by making moral rules absolute. Humans can only "rethink" elements of the world in light of experiences but an unseen order lies beyond human experience; hence the existence and nature of an unseen order is unquestionable. Irons acknowledges that people rethink religion in the face of negative experiences with the community and the moral code, but when religion is based on an unseen order that is not open to examination, it is easier to preserve and the moral rules become something that people "are simply committed to." The unseen order functions to identify the moral code and to remove it from rational, cost-benefit analyses. Appealing to an unseen order demonstrates one's abandonment of self-interest that in turn improves one's social reputation ("Inquiry" 365).

Religious commitments involve not only belief in an unseen order, but also a belief that the highest good is achieved by bringing one's life into accord with this unseen order (Irons, "Inquiry" 367). Irons explains that the idea that the unseen order is the source of the supreme good is central to religion ("Hard-to-Fake" 294). The second variation of Irons' definition of religion brings out this point: "religion is a commitment to bring one's life into harmony with an unseen order, a commitment that defines and justifies the basic moral order of a community" ("Inquiry" 365).<sup>10</sup> The subsequent stipulation that one's highest good is attained by attuning one's life with the unseen order implies that one's highest good lies in obeying the moral rules stipulated by the unseen

order. This added incentive not only further motivates people to comply with the moral code but it also justifies why one should do so: one should obey the moral code not only because the unseen order commands it and one cannot question the unseen order, but also because doing so puts one on the path to realizing her highest good. The unseen order both identifies and substantiates a community's moral code.

I have outlined Irons' understanding of the unseen order with no less detail than he presents it. In fact, my above exposition incorporates inferences on my part as to the relationship between the unseen order, the moral code, and the highest good (See "Inquiry" 364-65). Irons' minimalist depiction of the nature and role of the unseen order in his theory of religion does not convincingly explain why humans are motivated to commit to abide by costly moral rules and, as I discuss shortly, to signal this commitment in a costly manner. Irons claims that believers do not decide to commit to moral rules on the basis of whether the benefits outweigh the costs but instead simply commit to them because they lie beyond examination. Grounding the moral rules in an unseen order serves to strengthen religious commitment by removing it from the self-interest model. However, the notion of the highest good contradicts Irons' argument that the unseen order separates commitment from the opportunistic model. If the unseen order in religion prevents believers from calculating self-interest and choosing the actions that increase it, then what is the nature and function of believing in the highest good? Is the highest good positively correlated with self-interest, i. e., is it a benefit in the mind of the believer? If not, what is it and how does it motivate believers to make costly commitments? How does religion motivate morality if not through the promise of benefits, such as the highest good, that are real in the mind of the believer?

Joseph Bulbulia's assessment of the role of an unseen order<sup>11</sup> in his costly signaling theory of religion resolves the issue of motivation for costly commitment. Bulbulia contends that because there are high costs associated with humans policing, rewarding, and punishing one another to create and maintain within-group cooperative exchanges, but the efficiency of such external enforcement is low, free-riding remains unchecked.<sup>12</sup> However, Bulbulia argues that self-enforcement through supernatural cognition is not as costly as human external enforcement and is much more effective at

motivating cooperation. I contend that Bulbulia's specification that believers consider the unseen order to be a moral force that doles out rewards commensurate with cooperation in social exchange combines Irons' notions of a beyond-the-empirical, supremely good, and moral-rule-generating unseen order into a coherent argument that explains how belief in an unseen order motivates costly cooperative behavior ("Religious" 24).<sup>13</sup> This understanding of the unseen order as including an agent that scrutinizes human moral affairs alters the expected outcomes of social exchanges in accordance with cooperative interactions ("Religious" 24, 30).

The gods can dish out hurt – eternal damnation in hellfire, reincarnation as a garden shrub, bus terminal-like purgatories, and so on. But they frequently bring fortune to the good and righteous – lusty heavens, reincarnation as an emperor, release from the cycle of birth and re-birth, profound insight and protection from harms way. (Bulbulia, "Review" 665-67)

Byproduct theorist Scott Atran points out that it is the unseen order's characteristic of policing social exchange and rewarding cooperation and punishing free-riding that inspires believers to commit to it (Bulbulia, "Review" 663). Belief in supernatural causation motivates believers to cooperate by changing the strategy that renders the greatest benefit in the payoff matrix of social interaction from defection to cooperation when the other party cooperates.<sup>14</sup> With an unseen order ensuring perfect justice in the picture, cooperation rather than defection is the rational option.

The key to the theory of religion as an adaptation for social exchange is that all-seeing gods impinge on our lives to hold us morally accountable. The supernatural causation represented through religious conviction is one capable of solving prisoner's dilemmas between those who share similar religious outlooks. In an ordinary prisoner's dilemma, economic rationality favours defection. But religious persons views [sic] the world as bound by supernatural causation, one that alters the relevant payouts for exchange. (Bulbulia, "Review" 667)

Believers in supernatural causation may commit to an unseen order on the basis of zero empirical evidence, as Irons claims, but they do not commit to the cooperative moral rules without a form of rational calculation. Bulbulia contends that the imposition of a

novel payoff matrix in the mind of the believer renders the self-interested believer “befuddled” and “confused” such that the cooperative strategy in social exchange that was too costly is now on the benefit side of an analysis. Religious commitment is irrational to nonbelievers<sup>15</sup> because it is a commitment to cooperate and therefore a sacrifice of self-interest enhancing opportunities to free-ride when the partner cooperates or a sacrifice of self-preservation to defect when the partner defects. However, it is rational to a believer because within his belief system, the unseen order’s rewards increase self-interest and, as I discuss shortly, because reciprocal exchange brings about benefits. In the mind of the believer, the sacrifice of cooperating instead of defecting in a social exchange with a cooperator and thereby giving up the opportunity to free-ride is understood as an investment in which a supernatural agent ensures benefits through supernatural causation (Bulbulia, “Religious” 24-25). Rational calculation informs believers that the rewards and punishments meted out by this unseen causal order are too great to ignore: it is in the best interest of the believer to conform to the cooperative behavior that the unseen order commands. It is belief in a supernatural reciprocal justice that justifies the commitment to cooperate with cooperators, i. e., the conviction of supernatural justice certifies commitment to reciprocal exchange.

The cost of this internal enforcement of reward and punishment is cheap because an internalized belief system compels individual believers to police themselves; the fear of the unseen order’s retribution and the desire for the unseen order’s praise keep human behavior in check.<sup>16</sup> Anthropologist Richard Sosis adds that, even if groups of nonbelievers achieve similar levels of cooperation through external enforcement as groups of believers do through internal enforcement, groups of believers will have higher net gains than groups of nonbelievers because they pay lower costs to monitor and punish defectors (“Hutterites” 102). This is not to say that religious groups do not use external reinforcement and do not incur any costs of external monitoring and punishment. They do, but these are limited primarily to religious behaviors rather than all non-religious daily routines, and religious behaviors are much easier to scrutinize than other signals of commitment to the community because they are formal, conspicuous, repetitive, and public (Sosis, “Trust” 16; “Badges” 25-26).<sup>17</sup>

Clearly policing costs are substantially reduced in communities of prudent individuals who believe their transactions are perfectly policed by supernatural beings. All things equal then, members of religious coalitions are at an advantage over non-religious coalitions. They pay less to secure reciprocity. (Bulbulia, “Review” 668)

Internal enforcement is efficient because the believing observer finds that defectors are punished and cooperators are rewarded. Bulbulia qualifies that the unseen order does not need to be perfectly just in its rewards and punishments to alter expected utilities. For example, the unseen order’s reward may be a lesser infliction of pain for cooperators than defectors. However, it does need to bring better lives “on balance” to those who act righteously and worse lives “on balance” to those who act wickedly. Moreover, the rewards and punishments need not be visible, as in the case of rewards for the soul, or within this lifetime, as in the case of promised pleasures in the afterlife (“Religious” 30, 31). The notions of the soul and an afterlife increase the efficiency of religion as a system of cooperative enforcement by preventing seemingly unwarranted punishments of good people from disrupting belief in supernatural causation (“Religious” 30). It may seem to nonbelievers that the unseen order does not guarantee reciprocal justice but believers know that the unseen order’s rewards and punishments need not be observed to be real.<sup>18</sup>

Returning to Irons’ theory, although my description pinpointed the target of commitment as the unseen order, in several instances Irons claims that religion is a hard-to-fake signal of one’s commitment to be loyal to his group and to follow the society’s moral code: “Religion basically is a commitment to behave in certain ways without regard to self-interest. Most frequently religion entails a commitment to become a supportive member of a particular religious community and to adhere to its code of ethics” (“Hard-to-Fake” 293). This statement is consistent with his argument that religion is a commitment to an unseen order. As discussed above, Irons implies that to commit to an unseen order is to commit to the moral rules of one’s community: the unseen order is the source of the moral rules and abiding by the moral rules is how believers harmonize their lives with the unseen order and achieve their highest good.<sup>19</sup>

The notion of the unseen order is not only important to understand why believers are motivated to commit to their community, but also it may help to distinguish religion from other ideologies that involve hard-to-fake signals of commitment. Irons recognizes that religion is not the only hard-to-fake signal of commitment and it is not the only signal of commitment to appeal to an unseen order. Irons lists the following examples of cultural practices that, like religion, are hard-to-fake signals of social commitments: college degrees to denote intelligence and persistence, scars from fights to indicate that one does not back down physically, military punishment of disobedience to signal the seriousness of the military regarding conformity (“Hard-to-Fake” 298). Religion is different from these hard-to-fake “arrays of symbols” because additionally it appeals to an unseen order. However, Irons acknowledges that even modern secular political or moral philosophies appeal to an unseen order: secular ideologies “slide imperceptibly toward a tacit appeal to an unseen order” because the tendency to strengthen commitments by appealing to something beyond the seen is “built into our brains and is a product of past evolution” (“Hard-to-Fake” 295; “Inquiry” 367). Irons discusses Marxism as an example of a hard-to-fake signal of commitment that like religion involves rituals, symbols, strong commitments, and according to some, appeals to an unseen order (“Hard-to-Fake” 295). Irons’ contention that both religious and secular commitments appeal to an unseen order to gain strength prevents understanding belief in an unseen order as an element that distinguishes religious rituals from secular rituals.<sup>20</sup>

## Ritual

Irons considers belief to be an integral part of religion, but he argues that a definition of religion needs to be balanced with other elements of religion besides belief, namely rituals, sacred stories, and other symbols (“Hard-to-Fake” 294-95). In fact, he claims that ritual rather than belief is the “real core” of religion because rituals express hard-to-fake commitments (“Inquiry” 364).<sup>21</sup> In the following discussion of Irons’ evolutionary explanation of ritual, I look at how a religious signal is hard-to-fake and why believers are motivated to signal religious commitment.

Just as the case with morality, Irons claims that in religion believers must communicate their commitment in an easy-to-monitor and hard-to-fake manner. Although this brief discussion is repetitive, this unavoidable repetition reflects how Irons' theory of religion is no different from his theory of morality; very simply, religious ritual is the hard-to-fake signal of commitment that functions to facilitate intragroup cooperation.

Following Alexander, Irons contends that in our evolutionary history religious groups fared better and still fare better than nonreligious groups in warfare because religious people cooperate better with one another and are able to form larger, more cohesive groups. Irons, Bulbulia, and Sosis, all costly signaling theorists, agree that religious rituals signal the commitment to cooperate with fellow group members. Religious rituals communicate the believer's good reputation of commitment to the community's moral code.

Rituals must be hard-to-fake in order to prevent defectors from mimicking the rituals and invading groups of cooperating believers. As the advantages of successfully defecting in social exchanges with cooperators are substantial, nonbelievers could take it upon themselves to mimic rituals to infiltrate groups of cooperating believers. In other words, so long as individuals view the net benefits of joining a cooperative group as higher than those of joining an uncooperative group, whether or not they believe in the doctrines that promote cooperation, it is in their best interest to perform the ritual (Sosis, "Hutterites" 102-03). In order for rituals not only to help believers locate one another but also to communicate reliably that believers really are cooperators, rituals must be costly enough for nonbelievers to consider the costs of participating in rituals to be greater than the benefits of potential defection available once accepted by a community of believers.

Many rituals in various religions are costly to the individual carrying them out. They range from distortions of reality to tiresome tasks to life-threatening scenarios, but they all test whether the believer is willing "to put his money where his mouth is" (Bulbulia, "Religious" 27; see Bulbulia, "Religious" 29 for examples).<sup>22</sup> Irons explores the hard-to-fake religious practices of the Yomut Turkmen of Northern Iran to provide examples of religious hard-to-fake signals of commitment ("Hard-to-Fake" 299-301).

The rituals of the Yomut Muslims take the following forms: finding a clean place to perform ritual washings and praying five times a day, fasting from sunrise to sunset for one month per year, and giving a portion of their wealth, crop and stock to the poor. By participating in these rituals, they commit to the will of God as revealed by the prophet Muhammad, to adhere to the basic rules of morality, and to aid the less fortunate members of their community. The costs of the Yomut Turkmen's hard-to-fake rituals that signal commitment to the unseen order, the moral code, and the community are significant. Praying is time-consuming, costs the person opportunities to hunt, gather, breed, and rear, and distorts reality; fasting depletes metabolic reserves and could result in malnourishment; and donating money and goods incurs material cost (see Bulbulia, "Religious" 19-20).

Other evolutionary scientists studying religion concur with Irons' claim that religion is costly. Sosis discusses the significance of the dress of male *Haredim* in summertime Jerusalem as an example of a hard-to-fake religious signal of commitment to one's community:

By donning several layers of clothing and standing out in the mid-day sun, these men are signaling to others: "Hey, Look, I'm a *Haredi* Jew. If you are also a member of this group you can trust me because why else would I be dressed like this? Only a lunatic would spend their afternoon doing this unless they believed in the teachings of Ultra-Orthodox Judaism and were fully committed to its ideals and goals." Thus, the "quality" that these men are signaling is their level of commitment to a specific religious group. ("Badges" 10)

Cognitive psychologist Justin Barrett, who argues that religion is a byproduct, also concurs that by virtue of their costs religious acts provide concrete evidence of believable commitment:

Regular attendance at a long and perhaps tedious worship service costs time. Financially supporting the activities of a religious group may be costly. Agreeing to missionary service in faraway and hostile environments qualifies as costly religious behavior. Even publicly acknowledging belief in the face of ridicule counts as a cost. (Why 62)

Bulbulia lists several features characteristic of effective rituals: they are obligatory, highly emotional, minimally verbal, minimally practical, and ostentatiously public. Although Irons mentions that the “elaborateness” and time-consuming nature of rituals renders them hard-to-fake (“Hard-to-Fake” 293, 298), as Bulbulia adds to these features and elaborates on all of them, I briefly review his explanation of each of them. Bulbulia claims that if rituals are reliable gauges of religious commitment, then they must be obligatory and believer non-participation must be taken as a form of defection and punished (“Religious” 28). Regarding the quality of being expressed emotionally, like Irons, Bulbulia refers to Frank’s theory of emotion as an example of a hard-to-fake physiological response that clearly indicates the underlying motivation. However, Bulbulia further clarifies that a valuable signal of religious commitment handicaps<sup>23</sup> the signaler in a way that demonstrates the sincerity of the believer’s commitment to an unseen order of reciprocal justice: “the costs of religion are intrinsically connected to the meaning of the message conveyed” (“Religious” 26-27). The costs of rituals effectively deter those who do not believe in the religious teachings that underpin the rituals from imitating them (Sosis, “Badges” 10). With respect to the characteristic of being minimally verbal, Bulbulia argues that signals of religious commitment are not only conveyed with linguistic utterances because “words are cheap.” Declarations of faith and pious professions are poor vehicles for communicating commitment because non-religious defectors can easily mimic them and defect in social exchanges (Bulbulia, “Religious” 26, 28). Concerning the constraint of separation from everyday survival activities, Bulbulia makes it clear that rituals occur in a communal context that is distinct from practical dealings in the business of ordinary life (“Religious” 28). Although religious people display passion and confidence in their religious conviction, these illusory beliefs do not prevent them from following through with actions that are necessary to sustain their own lives: “The faithful believe the gods will provide, but still till fields, provide for children, arm themselves against attack, and seek medicine when ill” (Bulbulia, “Review” 680).<sup>24</sup> Finally, over and above Irons’ contention that religious commitments must be displayed in an easily observable manner, Bulbulia emphasizes that if their costs are to serve the strategic advantage of manipulating others into

cooperative exchanges, then they need to be put to the test on public occasions. Bulbulia elaborates that unlike private situations with family and close friends, public measures of sincerity are required between nonrelatives if they are to trust one another to the extent that both parties opt for the strategy of cooperation in social exchange (“Religious” 28).<sup>25</sup>

Thus far I have addressed Irons’ claims regarding the importance of believers signaling their commitment and of those signals being costly. Irons is much less clear, however, on what motivates believers in their own minds to perform costly signals of their commitment to the unseen order and the community’s moral code. Given that rituals are costly, an explanation of what motivates a believer to participate in them is vital to understanding their origin and function. Appeals to tradition cannot provide the motivation to make costly commitments and to participate in costly rituals because Irons’ evolutionary explanation is one of the origin and function of religion. Although tradition may explain a lot in ritual studies of actual religions, past or present, it could not have been a motivating factor when religion was first forming in our evolutionary history. Nor can appeals to a conscious desire to reduce conflict within the group explain the motivation to perform costly rituals. It is doubtful that believers choose to commit to costly moral rules and to signal this commitment through costly rituals with the conscious intention of finding fellow cooperators to form cooperative groups in order to reap the benefits of sociality. Moreover, there are many rituals of the utmost importance in religion that are not obviously or directly related to cooperation between individuals, the moral rules, or the well-being of the community. Irons’ reference to Alexander’s example of the Muslim prohibition of eating pork provides a useful illustration: Muslims do not eat pork even though this costly act is unrelated to settling conflicts of interest within the Muslim community (“Evolve” 73). If people are rational calculators by nature and therefore evaluate not only whether a commitment benefits their self-interest but also whether signaling this commitment is worthwhile, then for both the commitment and the signal to offer more benefits than they incur costs, there must be a guarantee of substantial benefits that outweigh these costs.

Although Irons should be commended for recognizing that belief in an unseen order plays a part in religion, his development of the notion of the unseen order as well as

the extent to which belief in it motivates cooperative behavior and its advertisement is minimal. Likewise, despite proposing a relationship between ritual and belief, Irons suggests a one-way relationship of ritual reinforcing belief in an unseen order but fails to consider how belief reinforces ritual.<sup>26</sup>

In the same manner in which Bulbulia explains that belief in a supernatural force of reciprocal justice motivates commitment, he explains that this same belief motivates costly signaling of commitments. For believers who genuinely accord a high probability to future supernatural judgment and reward, the cost of participation in religious ritual multiplied by the frequency of performing the ritual is less than the conditional probability of reward for having pleased the unseen order (Bulbulia, "Review" 670-71; "Religious" 29). For believers, rituals guarantee future advantages through supernatural channels such that the expected returns from the unseen order outweigh and therefore justify the costs of the rituals: "Only believers perceive the net gains of ritual performance and are therefore willing to pay the short-term costs in order to achieve the long-term benefits offered by religious communities" (Sosis and Alcorta, "Signaling" 267). Bulbulia claims that the logic is simple: if people believe that the unseen order will repay their sacrifice to cooperate, then they also believe that the unseen order will repay their ritual sacrifice (Bulbulia, "Religious" 29). For believers, the supernatural causal system that they consider to be reality accrues fitness benefits in terms of immeasurable rewards from the unseen order as well as actual rewards that are only available to individuals in cooperative societies. Sosis argues that the intensity of religious belief is negatively correlated with the perceived cost of religiously prescribed ritual actions such that the more one is committed to religion, the less one perceives the ritual to be costly ("Hutterites" 104). Belief in a cosmology reinforces ritual with meanings and rationalizations that provide "significance, insight, and appreciation for the performers" (Sosis, "Hutterites" 98) which lower the perceived costs of ritual performance for believer.

However, for defecting nonbelievers, the conditional probability of gain from cheating a believer by mimicking participation in ritual is less than the definite costs of the participation in a religious ritual multiplied by the frequency of performing the ritual

added to the possible cost of being found out as a cheater. For a nonbeliever or a weak believer, these same rituals are pointless and dangerous because they are seen to be based on an illusion that misrepresents the payoff matrix (Bulbulia, “Review” 670-71; “Religious” 29). Nonbelievers who are firm in their disbelief find the costs of performing a ritual high because in performing rituals they forgo alternative activities with high utility and benefits. Sosis explains that the cost-benefit analysis of a behavior is evaluated in terms of the set of feasible alternatives. He explains by reference to a religiously prescribed fast. For believers, the threat of hell-fire for eating during a religiously prescribed fast limits the option of going for high tea. In fact, with the promise of eternal damnation, there are no feasible behavioral options besides performing the fast and the costly behavior is eating during the fast, not the actual fast. For nonbelievers, the promise of eternal damnation holds no weight and such alternative options as snacking on a carrot while fasting, ending the fast early, or not fasting at all are not costly at all while fasting carries the cost of hunger (Sosis, “Hutterites” 103).

My intention in this chapter was to provide an exposition of Irons’ theory of morality and his theory of religion. I depicted the three propensities of indirect reciprocity, punishment, and commitment that Irons considers as having evolved in the context of morality to facilitate the formation of large, well-united groups of nonrelatives. I explored the central elements of Irons’ definition of religion to examine how religion solved the problem of free-riding. I summarized Irons’ argument that religion was selected for because as a hard-to-fake signal of commitment it defined, rationalized, and communicated moral behavior within a community. The discussion moved beyond summary into an analysis regarding Irons’ lack of explanation of why believers perform costly signals if natural selection has selected for rational calculators. I pointed out that although Irons fails to consider this issue of motivation in the mind of the believer for carrying out costly signals of already costly commitments, Bulbulia and Sosis supply the lack with their explanation of the relationship between belief and ritual.

The above exposition was also intended to set the stage for the next chapter in which I argue that Irons’ explanation of religion is not consistent with his adaptive

approach. The problem lies in his understanding of the relationship between the evolution of morality and the evolution of religion and his apparent unawareness of the criteria for adaptation in terms of evolved psychological mechanisms. Although Irons claims that his theory of religion is an extension of his theory of morality, I contend that the relationship is more aptly described as one of a superimposition of religion onto morality. An analysis of Irons' theory shows that religion does not merely draw on adaptations for morality while adding to the operation of morality with appeals to a religion-specific adaptation with a religion-specific function. Instead, religion is entirely based on adaptations that ground moral systems and that were selected for their non-religion-specific function in enhancing group unity. Irons implies that religion is not an extension of hard-to-fake signals of commitment, but rather one effective expression of the general capacity to generate and use hard-to-fake signals of commitment. In claiming both that the propensity for hard-to-fake signal of commitment was selected for because those signals facilitated indirect reciprocity and that religion is a hard-to-fake signal of commitment that facilitates indirect reciprocity, Irons' theory of the evolution of religion seems to make the same claims as his theory of the evolution of morality. As I demonstrate in the next chapter, religion cannot be an adaptation even if it is the most powerful costly signal of commitment if it relies entirely on adaptations evolved in the context of morality. Because Irons' understanding of the evolutionary foundation of religion is not clearly distinct from that of hard-to-fake signals of commitment in general, his theoretical evidence does not justify his claim that religion is an adaptation.

---

Chapter Two Notes

<sup>1</sup> Sosis and Bulbulia both credit Irons in the development of their costly signaling models of religion. Although Sosis' main contribution to Irons' theory is his empirical examination of the costly signaling model of religion, he does provide clarification on some points that Irons does not. I call upon Sosis' empirical testing of Irons' theory in the next chapter. Irons does not refer to Joseph Bulbulia's work as it was published subsequent to Irons' 2001 articles in which he focuses on the evolution of religion. However, like Sosis' model, Bulbulia's model of costly signaling supplements several areas where Irons' theory needs development.

<sup>2</sup> My exposition of these theories is based on Irons' understanding. I recount both of these theories as Irons presents them in his own work. It is not my intention to assess whether Irons' interpretation is accurate or whether the nature of the empirical support for these theories is valid and reliable. The point of this exposition is to show how Irons uses these pre-existing theories to conclude that religion is an individual-level adaptation and to comment on Irons' theory and method.

Irons refers to a third theory, anthropologist Robin Dunbar's theory that gossip in humans evolved to facilitate the formation of closer alliances and larger social groups ("Inquiry" 358). However, as Irons contends that Dunbar's theory is very similar to Alexander's theory (but uses different terms), I limit my discussion of Dunbar's arguments. However, I mention one aspect of his theory within the text of this chapter, namely, the group-size cap of 150 members, as Irons' evolutionary explanation of religion implicitly draws on it. My treatment of Dunbar's theory is relegated primarily to a note placed where Irons uses elements of it to develop his argument.

<sup>3</sup> Irons points out that Alexander calls this evolutionary explanation the "balance-of-power hypothesis" ("Shaped" 3).

<sup>4</sup> Although Irons claims that religion is adaptive, his incorporation of Alexander's theory challenges his claim that religion is adaptive at the individual level. To unpack my argument, a brief discussion of multilevel selection theory (MLS), including group selection theory, is in order. MLS holds that an entity at any biological level that interacts as a whole with the environment can be a unit of selection. In other words, natural selection can operate simultaneously at any biological levels at which entities differentially survive and reproduce because of variations in heritable traits (Gould, *Structure* 673-704). Alexander's theory contends that groups compete with other groups in warfare, such that some survive while others die out. Group selectionist David Sloan Wilson argues that this is a perfect example of the operation of group selection, but Alexander claims that it is an instance of individual selection (see Wilson's "A Critique of R.D. Alexander's Views on Group Selection," 1999). As Irons grounds his evolutionary explanation of religion at least partially on Alexander's identification of the selective pressure of group warfare in the evolution of sociality, Irons' evolutionary explanation too may be interpretable from a group selection perspective. Even Irons admits that a discussion of the "means of preventing the separate, and potentially competing, interests of unrelated individuals from making cooperation unfruitful for some members of the group, and thus leading to the breakup of cooperative groups [...] is the part of a theory of extensive reciprocity that has been hard for evolutionary biologists to explain without appeal to group selection" ("Evolution" 395). Although I do not focus on the group-selection adaptive approach to religion, I will point out that, like Irons, David Sloan Wilson argues that religion is an adaptation for unifying groups into cohesive units, but unlike Irons, he contends that it was selected for at the group level (see *Darwin's Cathedral*, 2002). I discuss D. S. Wilson's theory briefly in Chapter Three in the subsection concerning bio-cultural evolution.

<sup>5</sup> Irons claims that primary among the theoretical concepts that can explain larger cooperating groups are Alexander's indirect reciprocity, the game theory concept of commitment, and punishment. However, he adds that a complete model of how humans form large cooperative groups cannot ignore the role of political hierarchies that employ coercion to assure cooperation ("Co-Creator" 777). As Irons does not place much emphasis in coercion in his other works, I do not discuss it here.

<sup>6</sup> The rule of reciprocal altruism still holds in indirect reciprocity: natural selection favors 'altruistic' acts only if they increase the probability of the altruist receiving reciprocal aid in the future. However, in indirect reciprocity, to calculate the probability of reciprocity, one evaluates the observers rather than the direct benefactor. In indirect reciprocity, reciprocation comes from the indirect route of observers rather than solely from the benefactor of the altruistic act.

<sup>7</sup> Irons argues that Robert Dunbar's explanation of the evolution of close alliances reinforces Alexander's theory of morality (Irons, "Inquiry" 361). Dunbar asserts that gossip evolved to signal close social alliances among humans (Irons, "Inquiry" 361; "Shaped" 6; "Co-Creator" 784). He explains that grooming was the primary means of signaling alliance among our nonhuman primate ancestors (Irons, "Shaped" 6) where the longer the grooming time, the closer the relationship between the groomer and groomee (Irons, "Inquiry" 360). However, as these groups became larger and the grooming of more associates took too much time away from other basic survival and reproductive activities, grooming was no longer advantageous to fitness and the alternative behavior of conversation evolved to replace it (Irons, "Inquiry" 361). Conversation involves lower costs than grooming because several people can converse at once in short amounts of time whereas grooming is a one-on-one, time-consuming behavior (Irons, "Shaped" 6). The need of language for conversation was "the primary selective force for the evolution of linguistic abilities" (Irons, "Shaped" 6; "Co-Creator" 784). Irons states that for Dunbar, gossip was "the most common form of speech in evolving human populations" ("Inquiry" 361).

To show the parallel between Dunbar and Alexander's theories, Irons highlights Dunbar's recognition that gossip is moral judgment; he claims that Dunbar defines gossip as conversation that conveys information about other group members ("Inquiry" 361). Irons clarifies that gossip was the most useful form of conversation in the context of monitoring social reputations and communicating this information to allies ("Shaped" 7). He considers Dunbar's concept of gossip to be equivalent to Alexander's notion of morality because both involve the transmission of moral judgment and both have the same purpose: they are tools of indirect reciprocity that allow the formation of larger social groups ("Shaped" 7).

<sup>8</sup> Irons discusses punishment in the context of Alexander's theory of indirect reciprocity and specifies several kinds of strategies of punishment ("Co-Creator" 778). Punishment is another strategy that makes reciprocation more likely. For example, he states that Boyd and Richerson's "moralistic strategies" extend Alexander's notion of indirect reciprocity mediated by reputation. While indirect reciprocity is the social system based on individuals who reciprocate those observed to help others and punish those observed to hurt others, Boyd and Richerson introduce the additional strategy to reward and punish people according to how well they police one another. The heavy demand of "moralistic strategies" for conformity not only establishes the practice of reciprocity, but also establishes the regular practice of any costly act (Irons, "Evolution" 386-87; see "Co-Creator" 781-784 for Irons' discussion of different forms of punishment that maintain indirect reciprocity).

<sup>9</sup> Frank suggests that because humans have evolved not able to control the outward signs of their emotions such that they directly display the motivations underlying them, emotions are effective signals of commitment (Irons, "Inquiry" 363): "Were emotional displays easy to consciously manipulate, [they] would lose their value as signals. Were the displays not intrinsically linked to motivation, their information content would be uninteresting to observers" (Bulbulia, "Religious" 27).

<sup>10</sup> Irons acknowledges that he borrows the first part of this definition directly from James' definition of religion which he quotes as: "(R)eligion, in the broadest and most general terms possible, ... consists of the belief that there is an unseen order, and that our supreme good lies in harmoniously adjusting ourselves thereto" (James qtd. in Irons, "Hard-to-Fake" 294).

<sup>11</sup> Bulbulia is aware that the category "supernatural" is more inclusive than the category "god" but uses the terms interchangeably, "sacrificing accuracy for clarity" (see "Medicine" footnote 1). He holds that the "gods of fortune" can be supernatural agents or supernatural impersonal guiding forces but that they must

deliver reciprocal justice ("Religious" 31). In order to avoid the western connotations associated with the term "god," I refer to Bulbulia's "gods of fortune" as the supernatural. Although Bulbulia does not use Irons' term of the "unseen order," as I argue that his conceptualization of religion is like Irons' concept of the unseen order but given a moral aspect, I use the term "unseen order" in the context of Bulbulia's arguments.

<sup>12</sup> Bulbulia argues that, even with excessive rewards or punishments, the efficiency of police in enforcing cooperative exchange is limited for several reasons: they cannot assure accuracy in the observation of others' behavior, the reward of cooperation, or the punishment of defection; they succumb to the temptations of corruption (for example, bribery); occasionally they erroneously punish cooperators; and they may be rebelled against by coalitions disagreeing with their methods ("Religious" 24). The cost of external human enforcement is substantial as it requires rewarding people who police others as well as setting in place a system to police the police.

<sup>13</sup> Bulbulia argues that for this belief to evolve, individuals must have evolved several other cognitive constraints, including: "a system of projection and denial that generates supernatural commitment with zero empirical evidence [i. e., Irons' belief in the unseen order that escapes empirical inquiry]; the desire to seek out con-specifics who are of a similar mind about the gods; careful attention to displays that authenticate commitment; a willingness to publicly manifest and present evidence of god commitment; mistrust of heretics; and moralistic aggression against unbelievers where the costs of defection are high" ("Religious" 25). I discuss the remaining factors that I have not yet addressed throughout the rest of this section especially as they relate to the mechanics of indirect reciprocal exchange.

<sup>14</sup> Bulbulia frames this argument in more technical terms: for believers, the "strictly efficient" strategy is the "Nash equilibrium" ("Religious" 25). The change in the punishment structure equates the strictly efficient strategy where the greatest payoff is gained by interacting with the same strategy to the Nash equilibrium where the strategy with the greatest payoff is the one in which no player can benefit by changing his strategy while the other player maintains her strategy (see "Religious" 22-23 for a more thorough explanation of what Bulbulia means by a strictly efficient strategy and the Nash equilibrium). In a typical prisoner's dilemma situation with two players, each player is motivated to defect given that the best strategy of the other player is to defect. Defection is the Nash equilibrium because it is the best strategy no matter what the other player does. The strictly efficient strategy in this scenario would be for both players to cooperate – at least both gain something. However, if a player cannot be sure that his partner will cooperate and his partner's defection is at great cost to him, then he is not motivated to cooperate. Cooperation is the irrational strategy. Cooperation becomes a Nash equilibrium only if cooperators can secure encounters with fellow cooperators. An unseen order of reciprocal justice alongside hard-to-fake signaling of one's strategy ensures that cooperators find one another. I discuss Bulbulia's arguments concerning signaling later in this chapter.

It may be helpful here to relate the terms defection and free-riding in an attempt to clarify the relationship between them. When a player defects upon the knowledge that her partner intends to cooperate, she is free-riding. When a player defects upon the knowledge that his partner intends to defect, he is defecting but not free-riding because although he is not incurring any costs, he is not gaining anything for free either.

<sup>15</sup> Bulbulia's discussions pit "religious altruists" against "selfish atheists." I have called religious persons cooperating believers and nonreligious persons defecting nonbelievers. I have chosen to use the more inclusive term of nonbeliever rather than atheist because anyone who does not sincerely accept the reality and involvement in human affairs of supernatural causation will view the payoff in the usual way of defection, bringing about greater benefits regardless of the strategy of the other player.

<sup>16</sup> Bulbulia clarifies that religious individuals do pay a price for supernatural police forces with costly signals of religious commitment ("Review" 669). I discuss costly signaling, i. e., ritual, later in this chapter but my point here is that the costs of rituals are lower than those of secular policing and the believers

themselves do not consider rituals to be costly in any way because they are viewed as means to great supernatural rewards.

<sup>17</sup> In “Does Religion Promote Trust? The Role of Signaling, Reputation, and Punishment” (2005), Sosis contends that religious rituals reduce monitoring costs of collective pursuits but points out that religious groups must pay additional costs to monitor the performance of ritual obligations. However, because religious rituals are conspicuous, they are cheaper to monitor than other less observable activities. He specifies that monitoring costs are decreased in religious groups in three ways: 1) through making a greater investment in the religious as opposed to secular training of adolescents, 2) permanent stigmatizing markers, and 3) creating “cognitive dissonance” between religious beliefs and practices. Sosis makes it clear that the supernatural system of reward and punishment is an efficient complement to institutional structures using the mechanisms of reputation and punishment, but it does not replace it (16-17).

<sup>18</sup> Sosis’ discussion of the internalization of religious beliefs provides another reason why internal enforcement is more efficient than external enforcement: for individuals to remove the cognitive dissonance that they experience when they perform rituals but do not believe in the underpinning supernatural causation, they must either accept the beliefs or cease faking the rituals. In other words, even if nonbelievers attempt to invade groups of believers in order to free-ride, they end up converting, in which case they are now cooperating believers, or giving up the sham, in which case they no longer pose a threat as invading defectors (“Hutterites” 108).

<sup>19</sup> However, this statement is problematic in light of Bulbulia’s theory because it equates religion, more specifically belief in supernatural causation, with the abandonment of self-interest. As discussed above, Bulbulia argues that self-interest plays a major role in the decision to cooperate consistently rather than defect in social exchange. It is not that believers relinquish self-interest when they cooperate, but instead supernatural causation imposes a payoff matrix in which believers reason that their self-interest is increased by cooperation instead of defection. The factor of the highest good in Irons’ theory suggests that Irons too considers the increase of self-interest to motivate believers to conform to costly moral rules. I interpret one’s highest good to be positively correlated with one’s self interest. Believers cooperate rather than defect with cooperators because it is in their best interest, i. e., highest good, to do so.

<sup>20</sup> I discuss the implication of Irons’ understanding of belief in the unseen order in religion on his theory of the evolution of religion in greater detail in the next chapter.

<sup>21</sup> Although in his discussion of his theory of religion Irons makes ritual more prominent than belief, I present and analyze ritual only after a treatment of Irons’ explanation of the role of belief in an unseen order for several reasons. For one, I used Irons’ definition of religion as a starting point to analyze his evolutionary explanation of religion and this definition places an equal if not stronger emphasis on belief in the unseen order than it does on the signaling of commitment. Another reason is that an understanding of religious rituals as communicating commitment requires knowing what the commitment is. In other words, to understand the importance of religious ritual in facilitating indirect reciprocity, it is important to understand that it signals commitment to the unseen order as well as the community. As in the case of hard-to-fake signals of commitment in Irons’ theory of morality, religious ritual is the reliable, hard-to-fake signal of one’s good social reputation.

<sup>22</sup> Sosis specifies that religious communication takes the form of “the three B’s: religious behavior (ritual), badges (the physical manifestations of some ritual behaviors, such as tattoos or religious garments), and bans (behavioral restrictions known in anthropological circles as taboos)” (“Badges” 4). See Sosis, “Badges” 14-16 for a more detailed explanation of the differences among these terms. In keeping with Irons’ usage and for the sake of simplicity, I use the term “ritual” to encompass all of these forms of religious communication.

---

<sup>23</sup> Biologist Amotz Zahavi coined the term “handicap” to refer to costly-to-fake signals in which the costs to perform the signal are higher for “low quality” individuals and the benefits of performing the signal are higher for “high quality” individuals (Sosis, “Badges” 7). Low quality individuals are those who are “smaller, slower, weaker, sicker, and uglier than they want others to believe they are,” while high quality individuals are what they want others to believe they are. Irons, Bulbulia, and Sosis all credit Zahavi with introducing and defining costly signaling behavior in animals.

<sup>24</sup> Perhaps this is why the content of rituals is often irrelevant to resolving the practical problem of within-group conflict. For example, both Muslims and Jews do not eat pork and yet this ritual has nothing intrinsically to do with increasing within-group cooperation.

<sup>25</sup> Sosis points out that for the purpose of decreasing the costs of monitoring others’ commitments to the community, religious behaviors need to be public. However, he contends that private rituals are also important in the costly signaling explanation of religion. Private rituals convince the believer of his belief in doctrine, give meaning to public rituals, and help believers perceive greater benefits and lower costs of public rituals (“Hutterites” 109-10; “Badges” 26).

<sup>26</sup> Sosis explains how this relationship goes in both directions. Sosis claims that religious practices possess four characteristics that enable them to promote and internalize cosmological beliefs: they are physically displayed, publicly performed, formally presented (they lack ambiguity), and routinely practiced (continuous, daily, weekly, monthly, seasonally, or yearly). Notice the similarity between these features and Bulbulia’s list of characteristics of reliable signals of commitment. If the performer participates in the ritual behavior but is undecided about religious commitment or only weakly believes in its underpinning cosmology, then repeated performance of such a costly ritual will cause psychological discomfort, i. e., cognitive dissonance. This discomfort forces the performer to reconcile her attitude to be consistent with the behaviors involved in the ritual or to discontinue faking the ritual (Sosis, “Hutterites” 97-98; “Badges” 22-23). In this way, religious rituals maintain or transform attitudes to correspond with religious beliefs.

### Chapter Three: Irons' Theoretical Evidence for Religion as the Target of Selection

The result [of the selection for large, cooperative groups] is a strong propensity built into human psychology to learn culturally specific signals of commitment and to use these to create and maintain within-group cooperation. The most powerful cultural signals of commitment are religious ones, and thus evolution has built into human beings a strong propensity to seek a religious orientation toward life and to hold this orientation to be of the highest value.

Irons, "Hard-to-Fake" 293

In this chapter I consider whether Irons provides compelling theoretical evidence for his adaptive story of religion. I examine whether his adaptive story is sufficient to establish an adaptive evolutionary explanation. In other words, I assess whether his arguments theoretically justify his claim that religion is a biological adaptation. I identify two related arguments that do not support his claim that religion itself is a biological adaptation and I consider whether those arguments could be more coherently understood as arguments within different evolutionary approaches to religion.

This chapter is divided into two sections. In the first section, "Two Problems," I single out two arguments in Irons' explanation of religion that are inconsistent with an adaptive explanation of religion. Both of Irons' problematic arguments arise because of his understanding of the relationship between the evolution of morality and the evolution of religion: his theory of religion duplicates his theory of morality. One inconsistency concerns Irons' contention that the three non-religion-specific social strategies of indirect reciprocity, punishment, and signaling commitment solved the problem of free-riding in human evolutionary history and that religion is a hard-to-fake signal of commitment. My objection is that, if religion is built on ordinary<sup>1</sup> pre-existent adaptations and not directly selected for, then religion cannot itself be a biological adaptation. The second problem concerns Irons' contention that religion is the most effective hard-to-fake signal of commitment. This argument implies that religion is only one of many social control

systems that use those pre-existing adaptations for sociality. If religion is not distinguishable from ordinary social systems that use hard-to-fake signals of commitment by means of a religion-specific mechanism or function, then it cannot itself be a biological adaptation.

The conflict between these arguments and an adaptive approach to religion is rooted in the fact that Irons' two arguments clash with evolutionary psychology's criteria for adaptation in terms of psychological mechanisms. Irons opens the door for an analysis of his arguments from the perspective of evolutionary psychology because he explains that the evolution of religion is entirely grounded in propensities, i. e., psychological mechanisms, that evolved in the context of morality. Evolutionary psychologists, more specifically cognitive scientists of religion, argue that for religion to qualify as a biological adaptation, there must be evidence of the direct selection for a religion-specific psychological mechanism because it solved an adaptive problem at the time of its evolutionary origin (see Table 1 above, p. 20). Instead, Irons entirely grounds religion on adaptations evolved in the context of morality. As neither of the Irons' arguments I identified above conforms to this criterion, I show that Irons' theoretical evidence does not justify his adaptationist claim that religion was the target of biological selection.

In the second section, "Alternative Evolutionary Approaches," I establish a dialogue among the different evolutionary approaches to religion with respect to Irons' arguments. I consider whether Irons' claims make more sense within other evolutionary approaches. Also, I assess the impact of these different frameworks on the conceptualization and empirical testing of religion. I divide this section into three subsections based on the three alternative evolutionary approaches within which I consider Irons' arguments, namely, the byproduct, coopted adaptation, and bio-cultural. I argue that because all of these approaches take into account the evolution of human cognition, they can account for Irons' claim that religion is in some manner made possible by pre-existent, non-religion-specific adaptations. However, none but the bio-cultural approach is compatible with Irons' assertion that religion is the most powerful hard-to-fake signal of commitment because only the bio-cultural approach allows religion to have the same effect as the original function of the non-religion-specific adaptations.

## Two Problems

Because Irons studies the evolutionary nature of religion within the discipline of behavioral ecology, he focuses on examining how religious behavior provided the believer with a fitness advantage in past environments (“Inquiry” 357-58). In other words, he explains why religion exists in terms of how it benefited human ancestors. Irons argues that religion arose because it served as a hard-to-fake signal of commitment that reliably communicated individuals’ reputations to group members thereby facilitating cooperation among group members.

As discussed in Chapter Two, Irons’ theory of the origin and adaptive value of religion draws on several already existing theories of strategies evolved for sociality. Although Irons claims that religion is the evolved adaptation to promote intragroup cooperation, in his theory of morality, he lists several non-religion-specific behavioral strategies as the solutions to the problem of free-riding that hampered the formation and maintenance of large groups of nonrelatives (“Co-Creator” 777-78). Human sociality is founded upon reciprocity (“Inquiry” 359), but the behavioral strategy of indirect reciprocity (reciprocity + reputation + monitoring of reputation) as well as its extensions of punishment (ostracism of free-riders) and commitment (reputation + hard-to-fake signals of commitment) facilitate the establishment of cooperation in groups larger than those able to be held together by direct reciprocity (Irons, “Inquiry” 359; “Co-Creator” 777-78). The strategies of punishment and signaling commitment limit the extent to which free-riders can send deceptive signals of cooperative intentions in societies based on indirect reciprocity and undermine cooperation among group members (“Co-Creator” 778-79). In introducing these behavioral strategies as the mechanisms through which morality fulfils its function of facilitating within-group cooperation, Irons moves beyond an explanation of why morality evolved to one that invokes evolutionary psychology to account for the evolved morality-specific mental modules that generate morality. Irons concedes that the behavioral strategies of indirect reciprocity, punishment, and commitment are produced by psychological mechanisms when he states that they are “built into the human mind” (“Inquiry” 367).<sup>2</sup>

Although Irons contends that his theory of the evolution of religion is an extension of his theory of the evolution of morality, it is nothing more than a superimposition on it. In his earliest article addressing the evolution of religion, “Morality, Religion, and Human Evolution” (1996), Irons states that, “This paper presents a theory of the evolution of human morality by natural selection, and then expands the theory to formulate a parallel theory of religion” (375). Later in the same article he contends that, “Interpreting systems of religious belief as a way of communicating commitment is a logical extension of the theory of morality discussed above” (Irons, “Morality” 392). I understand ‘expansion’ and ‘extension’ to imply that the theory of religion takes the theory of morality as its starting point and makes claims that, though grounded in the theory of morality, move beyond it. Neither of these descriptors adequately depicts the relationship between Irons’ understanding of the evolution of religion and that of morality. With respect to the evolution of morality, Irons contends that the psychological mechanism for the reliable signaling of commitment was selected for in the context of morality because it solved a communication problem in an environment where free-riding threatened group unity. As Irons explains in his evolutionary account of religion that religion is as a hard-to-fake signal of commitment that solves the problem of unreliable social reputations in communities plagued by free-riders, his conceptualization of religion is not clearly distinct from his conceptualization of hard-to-fake signals of commitment in general. His account of religion does not extend or expand on his explanation of morality because religion and hard-to-fake signals are the same behavioral strategy that is based on the same psychological mechanism evolved in the context of morality for its function of solving the problem of free-riding that disrupted the formation and maintenance of large cooperative groups of nonrelatives.

This overlay of the evolution of religion onto the evolution of morality renders Irons’ explanation of religion at odds with his claim that religion is an adaptation. It results in the first problematic argument within Irons’ adaptive approach, namely, that religion is made possible by several pre-existing social adaptations and serves the same function as they do. Irons claims that the three behavioral propensities / psychological mechanisms of indirect reciprocity, punishment, and signaling commitment expanded the

size of cooperative human groups to a level that gave them an advantage in warfare among groups: “This pattern of monitoring, rewarding, and punishing helps to prevent free riding and makes the formulation of large, well-united groups possible” (“Inquiry” 360). Likewise, he claims that as a hard-to-fake signal of commitment, religion fulfils this same function. In other words, he recognizes that religion exists because it is grounded in these pre-existent ordinary psychological mechanisms and it serves the same function of controlling free-riding to enhance within-group cooperation as they do. If this is the case, religion is not a biological adaptation. Religion cannot be a biological adaptation without having been directly selected for to solve a problem at the time of its evolutionary origin (see Table 1 above, p. 20). In other words, religion cannot be an adaptation if it is built on pre-existing ordinary adaptations and serves the same role as they do.

In his discussion of religion, Irons does not mention any religion-specific psychological mechanisms or a religion-specific function. Although other evolutionary approaches can account for religion’s use of pre-existent psychological mechanisms, the adaptive approach to religion cannot. For religion to be an adaptation, at least one of the psychological mechanisms underlying it must be religion-specific. In other words, at minimum, one mental module must have evolved because of direct selection for its function of producing religious belief, experience, or behavior (Kirkpatrick 925). Irons does not claim that indirect reciprocity, punishment, or commitment are religion-specific psychological mechanisms. Instead he suggests that these social strategies pre-existed religion and are integral to any kind of communal existence, religious or not. In light of the cognitive science of religion’s criteria for adaptation in terms of religion-specific mechanisms, because Irons’ appeal to psychological mechanisms does not argue for any religion-specific mechanisms, he does not offer a coherent adaptive explanation of religion. As religion is built on pre-existing ordinary adaptations and not a single religion-specific adaptation, religion itself is not an adaptation.

Irons’ second argument that is not consistent within an adaptive approach stems from his contention that religion is not just a hard-to-fake signal of commitment, but more specifically that it is the most effective hard-to-fake signal of commitment. This claim implies that religion is not the only hard-to-fake signal of commitment but rather

one of many. Like religion, other manifestations of hard-to-fake signals of commitment communicate reputation to facilitate the monitoring and punishment of individuals with bad reputations in societies based on indirect reciprocity. Irons seems to argue that religion is a biological adaptation because, given all of the alternatives, the religious expression of hard-to-fake signals of commitment provided the most effective unifying system. However, for religion to be a biological adaptation, selection must have favored religious hard-to-fake signals of commitment in particular and not hard-to-fake signals of commitment in general. But Irons admits that religion as well as various other non-religious social control systems reliably unite large numbers of individuals into cohesive units: "In the formation of large cooperating groups, patriotic commitments to nations or commitments to adhere to the precepts of a particular religion can serve as means of creating cooperation" (Irons, "Co-Creator" 780). Since there is a plethora of non-religious hard-to-fake signals and many of them do a reasonable job at fulfilling their function of enabling cooperation among members of secular groups, it is probable that the propensity for hard-to-fake signals in general rather than religious hard-to-fake signals in particular was biologically selected for because it solved the recurrent problem of free-riding in any kind of social group. In this case, how is the unifying system of religion a biological adaptation while those of political organizations and soccer teams are not?

Although I argue that Irons' theoretical evidence does not support his overall claim that religion is an adaptation, it is important to consider briefly whether empirical evidence substantiates this claim. Irons' functionalist assessment of religion makes the key prediction that the degree of cooperation among group members should be greater in religious groups than nonreligious groups because religion is a more reliable, harder-to-fake signal of commitment than its secular counterparts. In order to distinguish his adaptive explanation from a "just-so story," Irons needs to show that empirical evidence substantiates this prediction as well as his wider claim that religion is an adaptation. Although he appeals to his own ethnographic study based on 30 months of personally conducted fieldwork among the Islamic Yomut Turkmen of northern Iran "to illustrate and clarify" how religion works as a commitment and signal of commitment, he

emphasizes that his discussion is “descriptive and suggestive” but cannot be taken as empirical support for his theory (“Hard-to-Fake” 299, 302).<sup>3</sup>

Both Irons and Joseph Bulbulia single out and commend anthropologist Richard Sosis as having done the most to elaborate and test the theory that religion enhances intragroup cooperation by serving as a hard-to-fake signal of commitment (Irons, “Inquiry” 367; “Hard-to-Fake” 302-03; “Co-Creator” 781; Bulbulia, “Review” 669). Sosis and his colleagues’ experimental designs are intended to test Irons’ claim that religion enhances intragroup cooperation.<sup>4</sup> The results of their analyses indicate that there is a positive correlation between religion and intragroup cooperation, costly rituals and intragroup cooperation, and frequent participation in collective rituals and individual levels of cooperation. However, their results cannot establish that religious hard-to-fake signals of commitment cause within-group cooperation or that intragroup cooperation was the selected function of religion. Sosis recognizes that such results can be coherently and consistently interpreted by alternative evolutionary approaches: “The results support any theory that posits that religion promoted intragroup cooperation” (Sosis, “Intragroup” 80-81). In Chapter One I discussed how traits need not be adaptations in order to have the adaptive value of being fitness-enhancing: coopted adaptations and coopted byproducts have fortuitous utility. For this reason, empirical tests that demonstrate a positive relationship between religion and within-group cooperation do not necessitate the interpretation that religion evolved because of this social function. Sosis admits that the burden is on future hypotheses to generate predictions that are unique to each of these alternative theories (Sosis, “Intragroup” 81).<sup>5</sup> At present, the empirical evidence for Irons’ costly signaling model of religion does not seem to be in much better shape than his theoretical evidence. Although tests and analyses support Irons’ claim that religion is positively correlated with intragroup cooperation, these claims may be shared by exaptive approaches and accounted for by adaptive or exaptive frameworks.

Putting aside the limitations of Sosis and his colleagues’ empirical evidence, their results do suggest that religion is positively related to within-group cooperativeness. Considering Irons’ theory, the relevant question is, if religion is based on ordinary social strategies and serves the same role that they do of facilitating sociality, then how is it

different from regular social systems and why is it better at serving this function than they are? Irons implies that religion is a souped-up moral system that is better than others at motivating its believers to cooperate. Sosis and his colleagues' tests too suggest that there is something about religion as opposed to secular systems that renders it better at unifying groups. This begs the question of what quality makes religion more effective. As many claim that passionate belief in the supernatural is an aspect of religion that seems to differentiate religion from other ideologies, I examine Irons' arguments to see whether his discussion of belief in the unseen order suggests that it is the religion-specific characteristic that would save his theory. If it is, then this aspect of religion must have evolved in the context of religion. In that case, religion would not necessarily be entirely explicable in terms of pre-existent social strategies and could be understood as a biological adaptation.

Irons' explanation of the unseen order is minimal. Although he considers the unseen order sufficiently important in religion to include reference to it in his definition of religion, his explanation of the origin and function of the belief in an unseen order is underdeveloped. Belief in the unseen order strengthens commitments by removing them from empirical examination: as the unseen order is the source of the moral code and the unseen order is beyond question, one cannot query the content, purpose, fairness, etc. of the moral code. In this manner the unseen order justifies the community's behavioral standards and Irons claims that group members simply conform to them. Irons' equation of the highest human good with harmonizing one's life with the unseen order further justifies the moral code: one should obey the moral rules not only because they are beyond investigation, but also because doing so leads to one's highest good. Irons' argument is clear that belief in the unseen order plays a role in motivating group members to abide by the moral standards of the group.

However, he is ambiguous about whether the propensity for belief in the unseen order is grounded in psychological mechanisms that evolved in the particular context of religion. Is the belief in an unseen order a religion-specific characteristic that differentiates religious social systems from secular systems and explains why religious systems are more efficient at facilitating within-group cooperation? Irons' discussion

reveals that the appeal to an unseen order is not unique to religious ideologies: any commitment is strengthened when it appeals to something beyond the seen. As appeals to an unseen order are not limited to religion, the question arises whether the ease with which humans invent, learn, and use the unseen order was selected for its function of strengthening religious commitment or any commitments, religious or non-religious. For the concept of the unseen order to be religion-specific, it would have to have evolved in the peculiar context of religious commitments. Irons qualifies that religious ideologies call upon the unseen order “more strongly” than non-religious ideologies: “Whether any particular set of symbols is a religion or not depends on the extent to which it appeals to an unseen order. Some arrays of symbols fall at the center of our definition of religion and others fall on the fuzzy boundaries” (“Hard-to-Fake” 296). However, this qualification does not help to determine whether the concept of the unseen order arose in the context of religious hard-to-fake signals of commitment. Moreover, Irons’ claim that religion appeals more strongly to an unseen order, like his argument that religion comprises the hardest-to-fake signals of commitment, does not clearly differentiate religion from non-religious ideological systems. It does not clarify the issue when he calls Marxism the “secular analog of religion” because like religion it is an array of motivating symbols that appeals to an unseen order but unlike religion its appeal to an unseen order is not as strong (“Hard-to-Fake” 295-96).<sup>6</sup> Without both reliable and valid quantitative measures of the strength of an appeal to an unseen order or the costliness of a signal, and the quantification of a minimum threshold for a religious commitment, in the realm of science the differentiating criteria of “more strongly” or “hardest-to-fake” is meaningless.

Although Irons’ attempt to distinguish religion from secular ideologies and philosophies by the degree to which they call upon an unseen order is unsuccessful, his idea of the unseen order as the key to differentiation is worthy of further scientific examination. Richard Sosis and Eric Bressler’s analysis of Irons’ claim that costly rituals promote within-group cooperation reveals that there is something about religious rituals in particular, as opposed to costly rituals in general, that enhances cooperation within groups. Sosis and Bressler reviewed data of imposed ritual requirements from records of 19<sup>th</sup> and 20<sup>th</sup> century American religious and secular communes to evaluate whether the

difference in the number of costly requirements of group members increases the group's ability to overcome the problems of collective action (operationally defined by commune duration). Their results indicate that costliness is not the only feature of rituals to promote group-member solidarity because an increase of costly requirements is positively correlated with commune longevity in religious communes but not in secular communes. If only the costliness of signals increased cooperation, then the level of within-group cooperation in both secular and religious communes should have increased to the same extent given the same degree of costly requirements. Sosis and Bressler state that their results indicate a "shortcoming" of the costly signaling theory of religion: it fails to capture the critical elements of religious belief that distinguish it from belief in a secular ideology (227). Religion is not distinguishable from and more efficient than non-religious systems because it is harder to fake.

Sosis and Bressler appeal to the work of economist Roy Rappaport to suggest that the difference between the communicative abilities of religious and secular rituals may lie in reference to a supernatural concept: religious rituals refer to supernatural entities while secular ones do not (227-28; see also Sosis and Alcorta, "Adaptive" 729). Moreover, Rappaport contends that religious rituals sanctify propositions that have no material referents and render them unquestionable true to the believer. Rappaport's suggestion is similar to Irons' in that both appeal to something beyond the seen to differentiate between religious and secular rituals. However, Rappaport's distinction criteria of supernatural entities and the sacred are much clearer than and seem to be quite different from Irons' contention that religion appeals more strongly to an unseen order. Nevertheless, Irons makes one brief statement that suggests that the unseen order may be sacred in addition to being outside of examination: Irons refers once to Rappaport and claims that participation in religious ritual invokes a "sense of the sacred or numinous" that is central to religion ("Hard-to-Fake" 295).

However, not all appeals to an unseen order are religious in nature, as Irons himself acknowledges. The existence of hard-to-fake signals of commitment that appeal to an unseen order but are not usually categorized as religious raises issues relevant to the scientific study of religion. How is religion categorized if like other social systems it is

grounded in ordinary social strategies and appeals to an unseen order? How can scientists study the evolutionary origin and function of religion if they cannot identify what differentiates religion from other social systems? The distinction of belief in the supernatural might fit the bill, but because Irons does not explicitly equate his conception of the unseen order with Bulbulia's or Rappaport's more possibly religion-specific understandings and he does not claim that only religious commitments appeal to an unseen order, it is not religion-specific.

Defining religion is a challenge for any scholar of religion, evolutionary scientist or not. A definition can either specify a necessary and sufficient common core or essence to the category (monothetic definition) or it can hold that the category has no single or conjunctive set of distinguishing features but instead is composed of overlapping features that criss-cross to form a collective such that all instances of the category do not necessarily share a single feature in common (family-resemblance definition) (Saler, Conceptualizing 79-81, 160). A definition of religion must be general enough to encompass the diversity of religious expression across history and cultures but narrow enough not to be vague and to exclude beliefs and actions that are typically classified as non-religious. It must be stated clearly such as to avoid ambiguity and alternative interpretations that can lead to borderline cases. It must not be stated in ethnocentric terms, it must not obscure some aspects of the category in favor of others, and it must not take superficial features as an indication of an underlying commonality (Boyer, Naturalness 29-30).

Irons uses a substantive and functional approach to defining religion: the "common thread" that runs through all religions is that religion is a hard-to-fake signal of commitment that appeals to an unseen order and functions to facilitate the communication of social reputations among group members such that larger more cohesive groups can form. The advantage of a monothetic definition is that it identifies a necessary common core that, in theory, should assist in the identification of the phenomena that fall into this category. However, even Irons admits that such is not the case with his monothetic definition of religion. Irons acknowledges that religion is a "fuzzy" category such that any cross-culturally valid definition of religion, including his

own, will not have sharp boundaries (“Hard-to-Fake” 296). His recognition that religion is a fuzzy category is supported by his arguments that religion is grounded in the same ordinary psychological mechanisms as all other moral systems. This along with Irons’ contention that all strong commitments appeal to an unseen order implies that religion is not clearly distinguishable from other social systems in general. Employing a monothetic definition for a fuzzy category is a contradiction that creates problems with the theoretical conceptualization and empirical examination of religion: a definition that employs clear criteria and boundaries for category inclusion cannot adequately depict a category that has no sharp boundaries. It is difficult to concur with Irons’ reasoning that the best that can be done with a fuzzy category is to recognize its central elements and “not get too exercised” about not being able to draw a clear boundary (“Hard-to-Fake” 295). Science is the very embodiment of “getting exercised” about clarity, validity, and reliability. Without some criteria to differentiate religion from other unifying systems it is unclear whether Irons is explaining the evolutionary foundation of unifying systems in general or the particular unifying system of religion.

Religion is a fuzzy category because it is multifaceted and complex. Irons’ essentialist definition of religion does not appreciate this many-sided nature of religion. Religion is difficult to define because it manifests itself in diverse ways. Defining religion as having a necessary and sufficient core increases the likelihood that the definition will not encompass religion’s many aspects. In other words, using an essentialist definition misses elements that arguably fall into the category of religion. For example, in defining religion as a hard-to-fake communication of one’s commitment to the group, Irons’ theory seems unable to account for the importance of private prayer to god(s). If the reason religious rituals are effective is that they are costly and easily observable, then why are some important costly religious rituals performed when no one can observe them? I question the implication of Irons’ definition that all that religion ever has been, is, and will be across cultures and time can be validly explained as a hard-to-fake signal of commitment that appeals strongly to an unseen order. Such a definition does not reduce religion’s manifest diversity in a justifiable manner but instead loses the ability to account for some important aspects of religious life.

However, swinging the pendulum in the opposite direction does no more to render a definition of religion capable of categorizing religion for the purpose of scientific study. Defining religion with more general terms risks rendering the category of religion so vague that it will include various ideologies and philosophies that are questionably religious. Irons' essentialist definition of religion cannot distinguish religion from Marxism and various other secular ideologies and philosophies. As Irons himself acknowledges, hard-to-fake signals of commitment are not unique to religion. Although Irons claims that religion is a hard-to-fake signal of commitment that appeals to an unseen order and serves to promote within-group cooperation, he does not consider all hard-to-fake signals of commitment that appeal to an unseen order and serve to promote within-group cooperation to be religious. As discussed above, Sosis and Bressler's study suggests that it is not just the hard-to-fake signaling of commitment that is responsible for religious groups having greater levels of within-group cooperation. Although belief in an unseen order may explain why costly signals are more effective at facilitating within-group cooperation in religious communes, I have demonstrated that Irons' undeveloped invocation of the unseen order does not establish that belief in the unseen order is the religion-specific mechanism.

Although I have placed a considerable amount of importance on Irons' understanding of belief in the unseen order as the key to rendering his theory coherent with the adaptive approach given the criteria of cognitive psychologists, I contend that even if Irons did establish belief in the unseen order as present only in religion, his theory still would not be fixed. In other words, the tendency for belief in the unseen order even if limited to religion may be explicable without the postulation of a religion-specific mechanism. Given the research by byproduct theorists that I discussed below, namely, that among many other findings, at minimum, it can explain why supernatural forces are typically agents and why they are minimally counterintuitive in terms of ordinary psychological mechanisms, it is likely that an evolutionary account of this belief in the unseen order aspect of religion could be explained entirely by appeal to ordinary psychological mechanisms and not have to posit any religion-specific modules. The determination of whether the belief in the unseen order is religion-specific or a byproduct

of an ordinary adaptation or an exaptation of ordinary adaptations or byproducts remains to be supported by empirical tests, but current theoretical evidence (for example, Irons) and empirical evidence (byproduct theory) suggests that it is not religion-specific.

Evolutionary psychologist Lee Kirkpatrick argues that it is unlikely that religion is an adaptation precisely because it is doubtful that religion, referring to such a diverse assembly of beliefs and practices, is the product of a unitary adaptation with a single identifiable function (Kirkpatrick 926): “what is commonly known as ‘religion’ necessitates the cooperation of so many different mechanisms that it is quite impossible to view this ephemeral conglomeration as such as an adaptation” (Pyysiainen, “Bridge” 340-41). Sosis and Bulbulia concur: “the multiple roles and complex functions of religion render it difficult to capture within a single theoretical approach” (Sosis, “Adaptive” 264) and “It is doubtful that diverse ways of thinking and acting that we describe as ‘religious,’ (for example: religious awe, ritual performance, belief in the unseen, duties to gods, concepts of supernatural causation, hope for an afterlife, fear of divine retribution, desire for redemption, piety, and the rest) can be made out in terms of a single mechanism” (Bulbulia, Eden 83). The cognitive evolutionary approach states that there are numerous psychological mechanisms involved in all human activities (Symons 147), including religion. For religion to be a biological adaptation, at least one of the converging psychological mechanisms active in religious behavior would have to have evolved for a religion-specific function. Cognitive scientist of religion Ilkka Pyysiainen argues that even with this tailored understanding, it is unlikely that religion is a biological adaptation: “it is very difficult to show that some given feature of human cognition has evolved only for religion even in this narrow sense” (“Bridge” 341). As my analysis of Irons reveals, he does not argue that any religion-specific mechanisms exist.

### Alternative Evolutionary Approaches

#### Do Irons’ Arguments Support a Byproduct Explanation?

Irons’ explanation of religion as a biological adaptation is not the only evolutionary theory of religion. In the first chapter, I introduced the byproduct approach as the main category of evolutionary explanations of religion that opposes the adaptive

approach. As the byproduct approach to religion grounds a fruitful research program with a network of evidence that is exponentially increasing in its “evidentiary breadth and depth,” to understand the significant challenge that it presents to the adaptive approach in general and Irons’ adaptive explanation in particular, I examine the its key principles and provide brief examples of byproduct explanations of religion.<sup>7</sup> The question, then, is whether the basic elements of Irons’ theory could be captured by a theory formulated within the byproduct framework. The conclusion I reach is that they cannot. Although the byproduct approach can explain religion’s reliance on ordinary, pre-existent psychological modules, it cannot account for Irons’ argument that religion is the most effective hard-to-fake signal of commitment that serves the same function as mundane costly signals of commitment.

The cognitive science of religion adopts a core principle of evolutionary psychology that cultural phenomena are the products of domain-specific psychological mechanisms (Tooby and Cosmides 24). Evolved human cognitive architecture is functionally organized into a collection of domain-specific information processing programs that each respond to different information from the environment to solve particular long-enduring adaptive problems characteristic of our hunter-gatherer past (Tooby and Cosmides 24, 49-50, 63-64). This position challenges the traditional view in the social sciences that the evolved human mind is a general-purpose and content-independent computer that requires external cultural input, via the processes of learning and socialization, for functional organization (Tooby and Cosmides 23).

When Irons considers social behavioral propensities and the psychological mechanisms that underlie them, he is clear that they are specific to the domain of social function and the problem of intragroup conflict. As he argues that punishment and commitment are extensions of indirect reciprocity, he suggests that these related social strategies serve slightly different functions in the context of social interaction. Indirect reciprocity’s introduction of monitoring, reputation, and observer coalition-formation solved the problem of the benefactor not paying back the altruistic act in direct reciprocity, punishment solved the lack of costly consequences for free-riding, and costly

signals of commitment solved the problem of deceptive reputation as a tool for free-riding.

As my previous discussions have implied, cognitive scientists of religion abide by the “naturalness-of-religion thesis”: they claim that the cognitive resources active in religious concepts, beliefs, and practices are the same as those employed by natural concepts, beliefs, and practices (Barrett, “Exploring” 29).<sup>8</sup> Religious belief and practice are “cultural manipulations of ordinary psychological processes of categorization, reasoning, and remembering” (Atran and Norenzayan 714); religious representations are not produced by specialized religious cognitive mechanisms that involve a special religious way of thinking (Boyer qtd. in Pyysiainen, Lindeman, and Honkela, “Counterintuitiveness” 343). Religion is the naturally communicated and regulated product of ordinary, mostly unconscious, cognitive and perceptive processes. Religion exists not because it was the target of selection but because it is a convergent byproduct of several non-religion-specific mechanisms that were themselves targets of selection (Atran and Norenzayan 714).

Positing that religious beliefs and behaviors are mediated by the same cognitive mechanisms as all other human thought and behavior contests the tradition in theology and religious studies that emphasizes the *sui generis* nature of religion (Pyysiainen, “Gods” 175). The *sui generis* perspective is incompatible with the cognitive science of religion because it holds that religion is rooted in an independent human faculty that is uniquely transcendent and irreducible to psychology. The cognitive science viewpoint likewise questions scientific endeavors that focus on distinguishing religion from ordinary life by investigating special ecstatic experiences, peculiar brain states, and uncommon emotional commitments (Barrett, “Exploring” 29). This includes the adaptive approach to religion. Because the adaptive approach argues that religion was directly selected for, it holds that religion cannot be broken down any further as somehow evolved from pre-existing evolutionary adaptations or byproducts. As an adaptation, religion evolved because it solved a problem at the time of its evolutionary origin. The adaptive framework contends that religion cannot be understood as a byproduct of ordinary adaptations or a product of the cooption of previous non-religion-specific

adaptations or byproducts because it has its own evolutionary mechanism that originated in the context of religion; adaptive reasoning implies that religion is *sui generis*.<sup>9</sup>

Given the principles of the cognitive science of religion, byproduct theorists Stewart Guthrie, Pascal Boyer, Lee Kirkpatrick, Scott Atran, and Ara Norenzayan maintain that religion is an incidental consequence of various psychological mechanisms, none of which were selected for religious cognition and behavior and none of which are activated solely in religious belief and practice. These underlying mental systems accommodate and constrain religious belief and practice but they did not evolve to process them for fitness advantages. Instead, they evolved as non-religion-specific adaptations to ordinary, recurrent fitness problems encountered by our ancestors.

To give a better sense of byproduct theories of religion, I present brief synopses of Stewart Guthrie's evolutionary explanation of religious agent concepts as a spin-off of agency-detection and Pascal Boyer's evolutionary explanation of religious supernatural agent concepts as a side-effect of intuitive ontology. Guthrie argues that religion, more specifically belief in purposeful, social, supernatural agents, is the byproduct of an adaptive perceptual strategy to interpret ambiguous thoughts, actions, events, etc. as animate, and more specifically, anthropomorphic. Humans involuntarily but systematically anthropomorphize in the face of perceptual uncertainty because this strategic interpretation affords rich inferences. Because humans are complex and generate a wide variety of phenomena, they explain a lot about the world. Thus, anthropomorphic expectations allow the interpreter to take precautionary measures to avoid harm and to take advantage of fitness opportunities (*Faces* 36). Natural selection has selected for a sensitive agency-detection mechanism because the strategy to anthropomorphize is an intuitively good bet: even if the majority of anthropomorphic guesses are proven incorrect in hindsight, on the rare occasion that the perceptual stimulus was correctly attributable to intentional thought or action, the interpreter enjoyed a fitness advantage. The tendency for religious concepts to involve supernatural agents stems from the ordinary evolved mental system of agency-detection.

There are at least two reasons to assume that other non-religion-specific mental systems are involved in religion. The first is that agent concepts in religion tend to be

intentional and human-like but with a supernatural twist (see Bulbulia, Eden 128-129). Guthrie does not account for why the agents of religion tend to have fantastic characteristics: “If we follow Guthrie’s reasoning to the letter then we should expect, in ambiguous perceptual circumstances, inferences to human beings, *real* flesh and blood persons, not gods. If it is important to find other people, why do we not infer to other people? That is, why do we infer to agents with supernatural natures and capacities?” (Bulbulia, Eden 129-130). The second reason why psychological mechanisms other than agency-detection must be involved in religion is that these superhuman agents tend to persist even when hindsight suggests a natural non-anthropomorphic explanation for an initially ambiguous occurrence (see Bulbulia, Eden 124-125). Guthrie does not explain why some humans detect, believe in, and respond to, the same extraordinary human-like agents over and over even when our senses confirm that a human, animal, or object caused the initially vague event and that such incredible human-like agents do not exist. Bulbulia explains that, “religious believers do not just accept certain supernatural facts as true, but also adopt certain behaviors as a result of those beliefs and harbor a full complement of powerful emotions, attitudes, and motivations” (Eden 130). Why are religious agents so plausible and people so passionate about them if anthropomorphism is by definition a mistake?

Cognitive scientist Pascal Boyer’s work provides one evolutionary reason why supernatural religious agent concepts persist. Boyer argues that supernatural concepts are easily acquired and transmitted because they are an incidental consequence of the ordinary evolved mechanism of intuitive ontology. A determining characteristic of supernatural concepts is that they are “minimally counterintuitive” (Barrett, “Exploring” 30, Barrett Why 22): they involve an explicit violation of an intuitive expectation for a given ontological category (Boyer, “Gods” 71-72). A crying statue of the Virgin Mary is an example of a minimally counterintuitive concept: the statue has the expected properties of an artifact such as it is solid, unable to talk, and incapable of self-propelled movement, but it violates the category of artifact with the transference of the biological property of crying which is expected of the ontological category of persons but not artifacts. It is this violated expectation that renders supernatural concepts attention-

arresting and memorable: Boyer argues that, “certain combinations of intuitive and counterintuitive claims constitute a cognitive optimum, in which a concept is both learnable and nonnatural” (Naturalness 121). Religious concepts of supernatural agents are more comprehensible upon combining Guthrie’s agency-detection and Boyer’s minimally counterintuitive concepts.

With this basic background in the principles of the cognitive science of religion as well as several examples of byproduct explanations of the belief in supernatural forces aspect of religion, what remains is to consider Irons’ problematic claims within the byproduct approach. Irons’ indirect argument that religion is built on pre-existent psychological mechanisms is consistent with the byproduct approach to religion. Grounded in the principles of the cognitive science of religion, the byproduct approach recognizes that ordinary mental modules are involved in generating religion. This approach specifies that these mundane psychological adaptations are involved because religion is an incidental consequence of them. As Irons’ implicit argument that religion is built on pre-existent adaptations for morality does not specify how religion is grounded in previous adaptations in the sense of specifying whether it is a byproduct of them, a cooption of byproducts of them, or a cooption of them, it is possible to interpret religion as being based on other adaptations because it is a byproduct of them. Both the byproduct framework and Irons’ implicit claim that religion is built on pre-existent adaptations for morality hold in common the idea that religion is made possible by ordinary pre-existing adaptations.

Although Iron’s indirect contention that religion is grounded in pre-existing psychological mechanisms is attuned with the byproduct approach, his argument that religion is a powerful hard-to-fake signal of commitment is not. If religion is a byproduct of the psychological adaptation for hard-to-fake signals of commitment, then it could not itself be a hard-to-fake signal of commitment, and what is more, a very effective one. As side-effects of adaptations, not only are byproducts not the same as the adaptations of which they are byproducts, but also they are traits that have no relation to fitness (see Table 1 above, p. 20). In the above examples of byproduct explanations of the supernatural concepts in religion, religious supernatural agent concepts are incidental

outcomes of the adaptive strategy to assume the presence of intentional agents and the adaptive ability to categorize external stimuli to form expectations and as incidental byproducts they do not enhance fitness.

However, Irons does not claim that religion is an incidental but inevitable, useless consequence of the evolution of several ordinary psychological mechanisms. Instead, religion is a powerful hard-to-fake signal of commitment that serves the same function, though enhanced, as other hard-to-fake signals. Neither the costliness nor the communication value of religious signals is a side-effect of the selected psychological mechanism of signaling commitment. Instead, these qualities of hard-to-fake signals of commitment are precisely why Irons claims that religion is the most effective type of hard-to-fake signal of commitment for promoting within-group cooperation. Although George Williams argued that adaptive traits must show design specificity, including the quality of being economic such that the benefits of the trait outweigh the cost to the bearer of having it, Irons' theory holds that costliness is not some uneconomic aspect of religion that begs for a non-functional evolutionary explanation. Instead, the costs of religious signals of commitment are balanced not only by often-immeasurable psychological benefits but also by real individual-level fitness advantages brought about by cooperative social exchange within a group. Individuals who incur costs to belong to religious groups that are more cooperative than secular groups live longer and reproduce more successfully. The costs of religion bring about the fitness benefits of religion that far outweigh the original costs. Bulbulia responds to byproduct theorist arguments concerning the costs of religion in the following manner:

If religion is cognitive noise, its sounds seem to be reproductively deafening. It is critical not to lose sight of the broad spectrum of human investment that lies behind a god-centered reality, one extreme of which is occupied by convulsing shamans, celibate priests and suicide bombers. It seems selection should have placed mufflers over the relevant cognitive systems that produce such understandings and practices as by-products. On the contrary, selection seems only to have amplified religious distortions with powerful emotional responses and motivations. (Bulbulia, "Religious" 21)

In other words, if religion in all its costly glory was a byproduct, then its destructive-to-one's-fitness quality should have been selected against and religiosity's costliness should have toned down or religion should have become extinct.

In comparison to adaptive explanations of religion, byproduct explanations of religion conceptualize and define religion in a manner that reflects the complex and multifaceted nature of religion and this in turn facilitates the empirical examination of the evolutionary landscape of religion. Unlike adaptive forays into religion that operate with the preunderstanding that there is a substantial and functional essential core to religion, byproduct explanations of religion recognize that religion includes too diverse a convergence of elements to be reduced to one necessary and sufficient essence and function. Irons claims that religion is a hard-to-fake signal of commitment that functions to promote within-group cooperation. If religion were no longer reduced to the substance of a hard-to-fake signal of commitment, then it could be considered a composite of several different elements, including hard-to-fake signals of commitment.<sup>10</sup> It is not that Irons is wrong to argue that religion is a costly signal of commitment but that he does not stipulate that this is only one aspect of religion. Moreover, if Irons heeded the arguments of byproduct theorists that ordinary adaptations explain religion, he would recognize that the propensity for costly signaling of commitment in general rather than religious costly signaling in particular was the adaptation, and could explain why there are numerous non-religious hard-to-fake signals that facilitate solidarity among group members. If the fitness-enhancing usages of religion were no longer limited to the one function of promoting intragroup cooperation, then religion could be shown to benefit its believers in various ways, including social cohesion. Without the blinders of an approach that demands identifying one function, I hazard that Irons would be more apt to research religion's various fitness-enhancing effects, none of which were selected for, but all of which would further our understanding of the nature of religion. Religion may play a role in facilitating and maintaining sociality, as Irons argues, but it may also allay anxiety and contribute to its believers' fitness in various other ways.

Byproduct explanations tend to use family-resemblance definitions of religion more so than adaptive explanations. As mentioned previously, family-resemblance

definitions of religion hold that religion is not one (or several) necessary and sufficient elements but rather a criss-cross of several non-necessary and non-sufficient factors. Such definitions cover a wider range of religious beliefs, practices, and experiences because they recognize that not all religions have exactly the same features but that most of them will have several in common. For example, Atran and Norenzayan posit that religion is a composite of four converging factors: 1) counterintuition: counterfactual and counterintuitive beliefs in supernatural agents, 2) commitment: hard-to-fake public expressions (for example, offerings and sacrifices) of costly material commitments to supernatural agents, 3) compassion: the mastering of our existential anxieties by supernatural agents, and 4) communion: ritualized, rhythmic sensory coordination of the previous three factors. When taken together, these four factors provide the following depiction of religion: religion is “passionate communal displays of costly commitments to counterintuitive worlds governed by supernatural agents” (713). When considered independently of his main contention that religion is an adaptation and instead within a byproduct framework, Irons’ implication that religion is founded on the pre-existent mental module of signaling commitment could serve as evidence of Atran and Norenzayan’s second factor, namely religion’s activation of the psychological mechanism of commitment. This possibility highlights that Irons’ adaptive explanation of religion only addresses one out of many possible aspects of religion that each may have their own evolutionary history. Atran and Norenzayan’s study of the role of religious supernatural concepts in alleviating existential anxiety shows that religion may have several fortuitous effects, ones that Irons’ adaptive approach to religion precludes him from considering.<sup>11</sup> Although it is difficult for any one evolutionary theory of religion to explain everything about all religious expressions across all cultures and historical times, with a family-resemblance conceptualization of religion, several evolutionary theories can be employed to explain several different aspects of religion, at least some of which are pertinent to a diversity of religions.

Byproduct approaches to religion also offer benefits in terms of empirical testing. As mentioned in Chapter One, it is easier to show the adaptive value of mundane non-religion-specific psychological mechanisms than of complex, multifaceted behavioral

propensities. By definition, psychological mechanisms are domain-specific and therefore more manageable in terms of testing than diversely expressed behavioral traits. Byproduct family-resemblance definitions of religion can attest to the argument that religion is composed of too many aspects to belong to one domain. By condensing religion to the one domain of hard-to-fake signals of commitment, Irons neglects a sophisticated evolutionary explanation of the belief aspect of religion, which might be further broken down into domains such as Guthrie's agency-detection and Boyer's violations of intuitive ontology. When evolutionary scientists study one of these factors at a time, it is easier for them to conceptualize the religious factor clearly and to design valid and reliable empirical tests.<sup>12</sup> The byproduct approach to religion is a productive research program with an ever-expanding network of evidence comprised of high-quality, diverse psychological experiments testing different aspects of religion

#### Do Irons' Arguments Support a Coopted Adaptive Explanation?

Candace Alcorta and Richard Sosis contend that religion is not an adaptation or a byproduct but a coopted adaptation (see Table 1 above, p. 20). In this subsection, I present their theory of religion and I identify the similarities and differences between the evolutionary approaches of coopted adaptation and adaptation to examine whether Irons' two problematic claims are more coherent within the coopted adaptive approach. My analysis of Irons' indirect argument that religion is built on pre-existing mental modules for social interaction is consistent with the coopted adaptive approach while his argument that religion is the most powerful costly signal of commitment is not.

Adaptations and coopted adaptations belong to different evolutionary categories, although both enhance fitness (see Table 1 above, p. 20). However, when applied to religion, the key difference between them concerns whether the cognitive and emotional structures that generate religion evolved because of their fitness-enhancing role in religion.<sup>13</sup> The adaptive approach should hold that the psychological mechanisms involved are religion-specific and were selected for because of their function in religion. The coopted adaptive approach maintains that certain psychological mechanisms evolved prior to religion as ordinary adaptations with non-religion-specific functions but were

coopted for a novel usage in religion. This difference can be understood in terms of the historical genesis of evolutionary mechanisms: adaptations are evolutionary mechanisms that originated because they solved a fitness problem whereas coopted adaptations are pre-existing adaptations that have been selected for a new effect. The adaptive approach argues that religion-specific mechanisms were selected for in our evolutionary history because they solved a problem in the past. Religion-specific mechanisms are adaptations with functions that evolved in the context of religion. By contrast, the coopted adaptive approach contends that natural selection selected for ordinary adaptations that solved mundane problems in our evolutionary history and later coopted these already present ordinary mechanisms for their utility in religion. Ordinary mechanisms coopted for a fitness advantage in religion were and still may be functional adaptations in the ordinary contexts in which they first evolved, but they are coopted adaptations with different effects in the context of religion.<sup>14</sup>

Alcorta and Sosis' argument that religion is the product of the cooption of various mundane psychological mechanisms has several implications for the conceptualization and empirical examination of religion. This modified adaptive account of religion incorporates research from the cognitive science of religion regarding the activation of non-religion-specific mental modules in religious beliefs and practices while maintaining Irons' basic position that religion promotes within-group cooperation. Instead of arguing for a religion-specific mechanism, the coopted adaptive approach contends that religion is the product of the convergence of several ordinary psychological mechanisms selected for a novel fitness-enhancing effect. Like byproduct theorists Atran and Norenzayan, coopted adaptive theorists Alcorta and Sosis adopt a four-part definition of religion that includes the cognitive aspect of belief in supernatural agents as well as the behavioral aspect of participation in costly, communal ritual.<sup>15</sup> Alcorta and Sosis' composite understanding of religion not only recognizes that religion is a heterogeneous trait composed of different aspects that can be studied independently and in relation to one another, but also facilitates the empirical examination of religion by making variables simpler to define operationally and experimental methods more straightforward to design. Moreover, Sosis and Alcorta claim that religion's reliance on pre-existing non-religion-

specific psychological mechanisms is “evolutionarily parsimonious and parallels numerous other adaptations” (“Adaptive” 750). Although they do not develop this statement, I speculate that they concur with the cognitive scientists of religion that it is difficult to demonstrate a religion-specific mechanism and that previous ordinary adaptations can go a long way towards explaining current fitness-enhancing traits. However, unlike byproduct theorists, coopted adaptive theorists argue that religion is not a side-effect, but rather a quirky shift of the usage of pre-existing adaptations to produce a new effect. One drawback of the coopted adaptive explanations relative to adaptive claims is that they carry “the additional evidentiary burden” of having to document theoretically and empirically that the original function of the adaptation is distinct from the effect of the coopted adaptation (see Buss et al. 546).

Although a trait cannot be classified as an adaptation if it relies on pre-existing adaptations, Alcorta and Sosis contend that a trait can have adaptive value regardless of whether or not it is built on previous adaptations. The determination of religion’s adaptive value rests not on whether religion incorporates pre-existent mental modules but rather on whether there exists evidence of adaptation of those modules to solve novel ecological challenges (326-27).<sup>16</sup> They claim that selection has coopted mental modules that evolved for non-religion-specific tasks for religion “to solve an ecological problem by promoting group communication and cooperation across space and time” (349). For example, with respect to religious belief, Alcorta and Sosis build on Boyer’s work and point out that religious supernatural agents are not merely natural category agents but rather “full access strategic agents” that have unlimited access to socially wrong public and private behaviors (327). Although supernatural agent concepts exist because of selection for agency-detection (Guthrie) and/or the intuitive ontological category of persons (Boyer, Atran and Norenzayan), in religious belief these supernatural agent concepts serve to enhance within-group cooperation. In the case of religious ritual, Alcorta and Sosis claim that selection acting on animal ritualized display resulted in a non-religion-specific adaptation for intra- and inter-specific communication (330). Although non-human ritualized displays and religious rituals share the qualities of formality, patterning, repetition, and rhythm (330), unlike animal rituals, religious rituals

incorporate music, chanting, and dancing to intensify neurophysiological priming and extend the impact of ritual beyond dyadic interactions (349) and they create and define the sacred and separate it from the profane (332). In these examples, agency-detection, intuitive ontology, and ritualized display are non-religion-specific adaptations but they have been coopted for a new role in religion. Thus, while religious belief in supernatural agents as well as religious ritual practice are not adaptations, they do have adaptive value: they are coopted adaptations that serve the novel fitness-enhancing effect of within-group cooperation.

The above discussion demonstrates that Alcorta and Sosis succeed in establishing that several ordinary psychological mechanisms selected for ordinary functions have been coopted for novel effects that are distinct from the original functions. They argue that religious agent concepts cannot be byproducts of ordinary psychological mechanisms for reasoning about natural agents or intentionality because in religion these concepts are not random or interchangeable: “Anthropological and psychological evidence, however, suggests that supernatural agents of religious belief systems not only engage, but also modify evolved mental modules. Moreover, they do so in socioecologically specific and developmentally patterned ways” and they “share common structural elements that maximize retention, transmission and affective engagement” (326, 348-49).<sup>17</sup>

While Alcorta and Sosis successfully establish that religious beliefs are the products of the cooption for intragroup cooperation of mundane adaptations, their same reasoning does not hold for the ritual aspect of their theory of religion. Like Irons, they claim that one aspect of religion is costly signaling of commitment and that the fitness-enhancing role of religion is within-group cooperation. However, as evident in the above discussion, Alcorta and Sosis do not demonstrate that the function of ritual is sufficiently different from the effect of religious ritual for cooption to have taken place. They contend that religious ritual is not the same content-wise as animal ritualized behavior and that religious rituals, unlike pre-existing hard-to-fake signals of commitment, use “emotionally-charged symbols” (325). However, even with a different content, the problem remains that religious rituals and animal rituals evolved to solve the same problem of social cooperation. Their own explanation of the ritual aspect of religion

seems to contradict the basic principle of the coopted adaptive approach that requires a difference between the original function and the coopted effect (see Table 1 above, p. 20).

Irons' contention that religion is the most effective hard-to-fake signal is incoherent with the coopted adaptive approach for the same reason that Alcorta and Sosis are not successful in providing a coopted adaptive explanation of the ritual aspect of religion: religious ritual cannot be a coopted adaptation if it serves the same fitness-enhancing role as the adaptations on which it is based. Religion cannot be the product of the cooption of the adaptation for the propensity to reliably signal commitment because religion is a hard-to-fake signal of commitment that like all other hard-to-fake signals of commitment promotes intragroup cooperation. A coopted adaptation is by definition, the result of the cooption of a previously existent adaptation to serve a new fitness-enhancing task different from the original adaptation's function. This novel effect can be in addition to or instead of the function for which the original adaptive mental module evolved, but the new effect must be distinguishable from the original function. Religion cannot be interpreted as a coopted adaptation of an ordinary adaptation for hard-to-fake signals because in this case both the function responsible for its original selection and the coopted utility solve the problem of free-riding.

Although Irons' theory posits that religion exists because of its function and that religious beliefs or practices were never, at any point, epiphenomenal byproducts of non-religion-specific adaptations, for the sake of a complete treatment of his theory in relation to different evolutionary approaches, I consider whether Irons' arguments are theoretically consistent with a coopted byproduct framework of religion. Because the coopted byproduct approach is based on the principles of the cognitive science of religion, like its byproduct and coopted adaptive rivals, it can account for Irons' implicit contention that religion is built on ordinary psychological adaptations. However, Irons' argument that religion is the most effective hard-to-fake signal of commitment is more supportive of a coopted byproduct than a simple byproduct framework because a coopted byproduct perspective would consider religion to be fitness-enhancing. Unlike the simple byproduct approach that claims that religion is not directly related to fitness, a coopted

byproduct framework would hold that religion was coopted from being useless to having an effect. Nevertheless, Irons' argument conflicts with a coopted byproduct explanation for the same reason that it is incoherent with a coopted adaptation: Irons claims that religion is a hard-to-fake signal of commitment and hence serves the same function as a hard-to-fake signal of commitment, but indeed more efficiently. It does not make sense to say that religion is a powerful hard-to-fake signal of commitment and that it originally arose as a byproduct of a hard-to-fake signal of commitment.<sup>18</sup>

It is worth pointing out that Sosis, the evolutionary scientist who was an advocate of Irons' costly signaling theory and carried out empirical tests of Irons' adaptive explanation, has modified Irons' costly signaling model in his work with Candace Alcorta from an adaptive explanation to a coopted adaptive explanation of religion. At first this seems like an odd move as his empirical test-results of a positive correlation between religion and within-group cooperation do not support his new coopted adaptive explanation any more conclusively than they did his original adaptive explanation. However, this revision does present several benefits for the empirical study of religion. At minimum it takes into account the findings of cognitive scientists of religion on the role of ordinary mental modules in religion and it recognizes that religion has more than one aspect. Theoretically, at least one of these newly considered aspects explains a finding in Sosis and Bressler's results that they could not account for in terms of Irons' original theory: in suggesting that one aspect of religion is that it defines and separates the sacred from the profane, it explains Sosis and Bressler's "puzzling" results that increasing costly requirements of group members only increased within-group cooperation in religious groups. Religion plays a fitness-enhancing role in sociality not only because it is a hard-to-fake signal of commitment but also because it marks out the sacred.<sup>19</sup>

One drawback to consider regarding Alcorta and Sosis' coopted adaptive theory of religion is that although the theoretical difference between a coopted adaptation and a coopted byproduct is clear, it seems as though it would be difficult to provide empirical evidence that a trait belongs to one of these evolutionary categories rather than the other. For example, how can a scientist empirically differentiate between the belief aspect of

religion as coopted agency-detection or as a coopted side-effect of agency-detection? Is the fact that the agents in religious belief systems are supernatural and not ordinary agent concepts evidence that they are coopted from ordinary reasoning about agents for an effect or that they are incidental consequences of ordinary cognition about agents that have been coopted for an effect? Does the shift of concepts from natural to supernatural need to be explained as a byproduct or simply as a result of cooption? Both Alcorta and Sosis' coopted adaptive explanation and Atran and Norenzayan's coopted byproduct explanations of supernatural concepts demonstrate the problem. As mentioned above, Alcorta and Sosis contend that the capacities for agency detection and intentionality have been coopted for the effect of enhancing within-group cooperation in religion. Atran and Norenzayan claim that supernatural concepts are the byproduct of agency-detection and intuitive ontology that have been coopted to relieve existential anxiety. Although both of these teams of evolutionary scientists contend that supernatural concepts have been coopted (for which effect is irrelevant to my argument), they disagree on the issue of whether the supernatural concepts became supernatural as a side-effect of natural concepts and then the process of cooption gave them an effect, or whether they were coopted directly from natural concepts and reshaped as supernatural for an effect. It is difficult to conceive of a test to differentiate between these two explanations as both conclude that supernatural agents have an effect in religion.

To summarize, the entirety of Irons' costly signaling theory cannot be coherently interpreted as a coopted adaptive explanation nor can Alcorta and Sosis' coopted adaptive explanation be construed as a fluid extension of Irons' theory. Although coopted adaptive explanations and adaptive explanations claim that religion has adaptive value, they are very different from one another, primarily because the coopted adaptive approach dismisses the key adaptive argument that religion originated because of its functionality in the environment in which humans evolved. Despite indirectly arguing for religion's reliance on non-religion-specific mental modules for sociality, Irons maintains that religion is an adaptation to the human ancestral environment: "It has been suggested that religions and religious rituals have served as hard-to-fake signals of commitment in ancestral human environments and that our propensity to be religious is an evolved trait

for this reason” (“Co-Creator” 781). In explaining his theory, Irons focuses on demonstrating that religion was positively correlated with fitness in the past, arising as a solution to the recurrent problem of free-riding during the period of its evolution. In other words, his overall theory precludes understanding religion as the product of the cooption of any pre-existing adaptations or byproducts. Although religion’s current utility is not a measure of adaptation, Irons does argue that religion still serves its function in the current environment: in addition to claiming that, “throughout most of human history and prehistory, human beings have been religious, and religion has defined and rationalized the principles governing behavior,” he holds that, “religion still does this, in our present-day world, with much greater force than moral philosophy” (“Inquiry” 365).

#### Do Irons’ Arguments Support a Bio-Cultural Explanation?

Although I have not yet mentioned cultural evolution, in this last subsection I consider whether Irons’ problematic arguments might be rendered coherent within a bio-cultural approach to the evolutionary study of religion. So far my discussions of evolution and adaptation have been confined to biology because the nature of cultural evolution and its relationship to biological evolution remain contentious issues in the scientific community. Also, I have been reluctant to present a bio-cultural explanation of religion because the one evolutionary biologist who argues that religion is both a biological and cultural adaptation, David Sloan Wilson, sketchily develops his theory within the contested adaptive approach of group selection and provides questionable theoretical and empirical evidence to support his adaptive claim. However, I include an analysis of Irons’ arguments within the bio-cultural framework because the cultural adaptation aspect of this approach does not stipulate that religion cannot serve the same role as the function of the propensity in which it is grounded. As this was the main problem in interpreting Irons’ claims as consistent with all of the alternative evolutionary approaches considered thus far, I include this bio-cultural approach within the dialogue among current evolutionary approaches to religion. In the following pages I briefly outline the possible ways to understand cultural evolution, present D. S. Wilson’s basic idea of how religion can be both a biological and cultural adaptation, and examine whether Irons’ two

problematic arguments might support this kind of evolutionary understanding of religion. I conclude that Irons' arguments do indeed find a home in a bio-cultural approach to religion.

Currently there is no consensus regarding cultural evolution's precise nature, importance, or relationship to biological evolution (D. S. Wilson, Cathedral 28). Cultural evolution is like biological evolution in that it involves "(i) a set of units, (ii) changes that produce different variants of those units, [and] (iii) a mechanism of transmission that chooses between variants" (Boyer, Explained 34). In biological evolution the units of inheritance are genes whereas in cultural evolution the units of inheritance/transmission are cultural. The cultural unit capable of differential transmission has been labeled in various ways: for example, Charles Lumsden and Edward O. Wilson call it a "culturgen" (qtd. in Boyer, Naturalness 268), Boyd and Richerson call it a "cultural trait" (qtd. in Boyer, Naturalness 270), and Richard Dawkins calls it a "meme" (192).

Although Dawkins has been known to lobby against a straightforward analogy between the replication of genes and that of memes (Boyer, Naturalness 283), he claims that, "Just as genes propagate themselves in the gene pool by leaping from body to body via sperms or eggs, so memes propagate themselves in the meme pool by leaping from brain to brain via a process which, in the broad sense, can be called imitation" (Dawkins 192). Simply stated, a meme is a distinct memorable unit of culture that competes with other memes for space in the limited capacity of human minds and, if successful, parasitizes the brain and turns it into a vehicle of propagation (Dawkins 192; Dennett 349). A meme is the essential information or idea behind any belief or practice that so replicates. It can be transmitted through any medium: a tune, catch phrase, fashion expression, way of making pots, picture, book, tool etc., (Dawkins 192; Dennett 345, 347; Guthrie, Faces 199).

Several evolutionary theorists point out that genes and memes have important differences and that the processes of biological evolution and cultural evolution differ in several key ways (see Sober's discussion of Cavalli-Sforza and Feldman and Boyd and Richerson in "Models"). On the one hand, in biological evolution an organism must have an offspring for a phenotypic trait to spread: in humans the transmission of genes is

vertical from parent to child and requires one generation measured in tens of years (Sober, “Models” 490). Random mutations occur rarely and when they do introduce variation it is most often disadvantageous to fitness and selected against, less often neutral to the individual’s fitness and invisible to selection, and very rarely advantageous to the individual’s fitness and selected for. Biological adaptations depend on the accidental availability of favorable mutations when the need arises (Ayala 515) and do not occur over time scales that matter to living organisms struggling with their problems (D. S. Wilson, Cathedral 35).

On the other hand, in cultural evolution, transmission of acquired cultural units is multidirectional (horizontal and vertical) among any numbers of related or unrelated people and is certainly possible in time spans of under one generation. Memes are not perfectly replicated and mutations commonly arise because the human mind does not transfer content from one brain to another like a photocopying machine but instead it transforms, invents, interpolates, sensors, mixes, and adjusts inputs and outputs (Dennett 354-55; Boyer, Naturalness 283, 285). A fit meme is one that spreads relatively unchanged through a population and becomes fixed in the population’s cultural repertoire.

The primary issue of contention regarding the relationship between biological and cultural evolution concerns the fitness of a cultural trait and its effect on the bearer’s biological fitness. Do fit cultural traits increase their bearer’s survival and reproductive success? Can cultural units with neutral or even negative effects on the carrier’s biological fitness spread successfully within populations? On the one hand, scholars such as Irons and D. S. Wilson agree that relatively successful ideas are associated with biological fitness advantages. According to Irons, the psychological mechanisms that evolved for the human capacity and propensity to absorb culture tended to keep culturally influenced behavior directed toward reproductive goals (“Evolve” 60).<sup>20</sup> D. S. Wilson agrees that, “What will *not* be observed, or seldom observed, are major beliefs [. . .] that actually handicap the believer by motivating dysfunctional behaviors” (Cathedral 156). I contend that both would agree with E. O. Wilson’s statement that genes hold culture on a

leash; they provide some leeway but constrain cultural units, such as ethical values, to accord with their effects on the gene pool (Human 167).

On the other hand, Boyer argues that one of the main features characterizing the current “coevolutionary paradigm,” i. e., the theoretical frameworks describing meme-gene evolution and transmission, is the “reject[ion of] the assumption of early sociobiological models, following which the recurrence of cultural traits is a direct consequence of their inclusive fitness value for the organisms that carry them” (Naturalness 267). In support of Cavalli-Sforza and Feldman, Sober explains that although the mind is a product of biological selection, it is more than a device for generating behaviors that biological selection favors because genetic selection has given our species the ability to engage in social learning (“Models” 487). The mind is the basis of a selection process with its own fitness and heritability measures that can counteract the pressures of biological selection and allow for the evolution of cultural traits that otherwise could not have evolved (see D. S. Wilson, Cathedral 34). A cultural “trait can evolve in spite of its Darwinian disutility” (Sober, “Models” 484) so long as it is advantageous to itself (Dawkins 200).<sup>21</sup>

This debate regarding the relationship between the fitness of the cultural trait and its effect on the fitness of its carrier becomes clearer upon distinguishing between “evoked culture” and “acquired culture.”<sup>22</sup> Evoked culture consists of evolved psychological and behavioral responses of mechanisms adapted to certain incoming stimuli. The psychological mechanisms that reliably generate ideas and actions were selected for in our evolutionary history because they offered their carriers fitness advantages. Stewart Guthrie’s detection of agents is an example of evoked culture: anthropomorphic ideas are the standard result of ambiguity in the environment pushing the button of the adaptive strategy to detect agents when in doubt. In evoked culture, cultural ideas and behaviors recur because of their direct effect on the bearer’s adaptive fitness (Boyer, “Evolutionary” 994). In contrast to evoked culture, acquired culture is composed of ideas and behaviors that are ‘attractive’ to their bearers but need not positively influence their biological fitness. Acquired culture is “the recurrence of particular [mental] representations in a group [. . .] mainly caused by *communication*

between minds and by the susceptibility of minds to be affected by particular kinds of representations” (Boyer, “Evolutionary” 990).

Cognitive scientists explain that not all acquired ideas and behaviors become stable in populations. Different cultural traits have different structures, some of which activate pre-existing domain-specific psychological adaptations that render them comparatively easier to acquire, remember, and communicate. For example, supernatural concepts recur in populations because, as minimally counterintuitive concepts, in activating and violating in a minimal way the psychological mechanisms for intuitive ontology they are attention-arresting and memorable. Although cultural traits may seem to spread because they are ‘attractive’ to the carrier for a multitude of idiosyncratic reasons, the latent reason why they are successfully transmitted is that they play upon evolved cognitive mechanisms. In this sense, biological and cultural evolution cannot be completely independent of one another because both evoked and acquired culture, although to varying degrees, appeal to the evolved cognitive abilities of the people who construct them (Kirkpatrick 924).

Like William Irons, David Sloan Wilson argues that religion was selected for its function of preventing within-group cheating, thereby uniting groups into cohesive units. However, unlike Irons, he claims that religion is a biological and cultural adaptation at the group level.<sup>23</sup> D. S. Wilson argues that human nature cannot be explained entirely in terms of biological evolution because then it would have to have been set in stone during the Stone Age (Cathedral 2). Although he agrees with evolutionary psychologists that the mind is a collection of adapted, domain-specific mental modules, he argues that this picture of cognition is incomplete because “it seems to deny learning, development, and cultural change as open-ended processes” (Cathedral 28-29) and it does not account for how humans were able to survive in the varied human ancestral environment (Cathedral 31). He argues that biological evolution led to a modularity that is evolutionary in its own structure and capable of providing new solutions to new problems. Although these modules are genetically evolved and highly specialized to lead to biologically adaptive outcomes, they remain open-ended processes (Cathedral 31). This open-ended process is cultural evolution: without any evolution taking place at the biological level, cultural

variation can result in cultural adaptations with important functional consequences (Cathedral 32).<sup>24</sup>

Like Irons, D. S. Wilson builds his theory of religion on top of his theory of morality. He argues that, “human moral systems have both a genetically evolved component and an open-ended cultural component. An innate psychological architecture is required to have a moral system, but the specific contents can vary and therefore adapt to recent environments.” Given that moral communities larger than a few hundred people did not exist prior to the advent of agriculture and biological evolution slowed to a negligible pace after this point in human evolutionary history, the evolved mind was originally designed to work in small hunter-gatherer groups. A cultural component is required to explain modern moral communities the size of nations (Cathedral 119).

For D. S. Wilson, religion is one cultural expression of a unifying system that unified human groups of sizes well beyond what could be held together by biological adaptations alone. Religion is a product of cultural evolution made possible by evolved psychology:

innate psychological mechanisms do not necessary [sic] limit cultural evolution but rather provide the building blocks that cultural evolution uses to create innumerable forms. Forgiveness and faith are two examples of capacities that are part of the psychological toolkit of all normal humans and that have obvious function outside the context of religion, which are put to new use by a culturally constructed religious belief system. (Cathedral 120)

Built on ordinary psychological mechanisms, religion is a culturally-evolved system of sacred symbols that represents and puts into action a biologically-evolved moral system that motivates its believers toward actions that are adaptive in the real world (Cathedral 227, 228).

D. S. Wilson holds that religion is a biological adaptation and a cultural adaptation. He refers to it as a biological adaptation because it depends on innate psychology. It is important to emphasize that D. S. Wilson’s understanding of biological adaptation is problematic in the same manner as William Irons’, namely he considers religion to be an adaptation despite acknowledging that religion is built on previous

adaptations. As discussed earlier and shown in Table 1 above (p. 20), religion cannot be an adaptation if it was coopted from ordinary adaptations. To clarify D. S. Wilson's argument in light of the terms used in this thesis, like the cognitive scientists adopting a coopted adaptive perspective, he seems to argue that religion is built on pre-existing non-religion-specific mental modules that were coopted by natural selection for the effect of coordinating groups well beyond the size the human mind was biologically constructed for. Religion is also a cultural adaptation because the outputs of these biological mental modules are molded into a particular cultural manifestation that best meets the demands of the environment.

Although D. S. Wilson's basic idea of religion as a bio-cultural adaptation seems similar to Alcorta and Sosis' coopted adaptive explanation of religion, without using these terms D. S. Wilson emphasizes that the process of cooption is cultural selection rather than biological selection. In other words, the biology of ordinary psychological mechanisms activated in religion - for example, the capacity for symbolic thought - is not modified or reshaped in any way by biology but instead their natural outputs are "tinkered with" to create the cultural expression of religion that best unifies group members in a specific environment. This is why religion, although playing upon the same innate psychology and functioning to facilitate within-group cooperation, can be expressed so differently across cultures.

The above discussion addresses how D. S. Wilson's basic idea of a bio-cultural evolutionary approach is attuned with the contention implicit in Irons' argument that religion is built on pre-existing non-religion-specific adaptations. This argument comprises the biological aspect of a bio-cultural explanation. However, I suggest that both D. S. Wilson and Irons should not call religion a biological adaptation because this implies that religion was directly selected for, but instead call it a coopted adaptation because it relies on previous mundane adaptations.

Irons' assertion that religion is the most powerful hard-to-fake signal of commitment too is consistent with a bio-cultural explanation of religion. As this argument is incompatible with all of the other evolutionary approaches that I have considered thus far, I explain this coherency in greater detail. Irons' argument implies

that religion is only one of many different kinds of hard-to-fake signals of commitment that function to facilitate the reliability and communication of social reputations among group members. Although it is difficult to explain how religion, as a hard-to-fake signal or even the hardest-to-fake signal, can be a biological adaptation given that there are numerous manifestations of hard-to-fake signals, this is not a problem for bio-cultural evolution. Within evolutionary psychology, the field that has provided a model of the evolved human mind, the consensus seems to be that it is not the particular cultural expression of a belief or behavior that is the biological adaptation, but rather the propensity or tendency for it that is. In other words, it is not the Calvinist belief in God or even a religious belief in supernatural agents that is a biological adaptation, but rather the ordinary capacity for belief in intentional agent concepts. Likewise, to use an example relevant to Irons' theory, it is not the Jewish ritual of circumcision or even religious bodily mutilation in general that is a biological adaptation, but rather the mundane capacity to form, communicate, and interpret hard-to-fake signals of commitment to one's group. Irons' argument can be interpreted as the claim that of all of the possible cultural manifestations of the biological adaptation for hard-to-fake signals of commitment, religious costly rituals seem to be the most effective at coordinating group behavior in general. Religion is biologically grounded because it is made possible by the ordinary psychological mechanism for hard-to-fake signals of commitment but it is a cultural adaptation because it is a specific culturally-selected expression of a costly signal and is particularly effective at coordinating group members.

Although the coopted adaptive approach is similar to the bio-cultural approach, only the bio-cultural approach can explain a cultural adaptation that has the same function as the innate psychological mechanism that makes it possible. I argued that Irons' argument does not support Alcorta and Sosis' coopted adaptive explanation because that framework demands that the effect of the coopted adaptation must be different from the function of the original adaptation. However, Irons claims that the role of religion (the possible coopted adaptation) is the same as the function of any hard-to-fake signal (the original adaptation). However, the bio-cultural approach does not require the function of the cultural adaptation to be different from the function of the biological

adaptation. The capacity to employ hard-to-fake signals of commitment was molded into human cognitive architecture because it reduced free-riding thereby reducing within-group conflict. The particular religious content of hard-to-fake signals of commitment was shaped by cultural evolution because this cultural expression happened to be more powerful at serving the function of enhancing within-group cooperation.

Several of Irons' statements suggest that he would not be completely opposed to a bio-cultural explanation of religion. For example, he states that,

I have laid out a theory of the behavioral propensities built into the human mind that make it easy for human beings to create and maintain strong commitments. The strongest of these commitments appeal to an unseen order and to a belief that the highest good is achieved by bringing our lives into harmony with this unseen order. *Actual religions are cultural traditions built upon these evolved behavioral propensities.* Religions are, I suggest, built up over many generations and consist of numerous symbols. ("Inquiry" 367; emphasis added)

This statement evokes a bio-cultural explanation of religion: the psychological mechanism for hard-to-fake signals of commitment is biologically evolved while the religious expression of the capacity for hard-to-fake signals of commitment is culturally evolved. Again, Irons suggests that religious beliefs and practices are the most powerful and hardest-to-fake signals of commitment because they appeal to belief in an unseen order that is intimately tied to the achievement of one's highest good. Although I have argued that Irons is unclear whether the propensity to appeal to an unseen order is religion-specific, in the context of a bio-cultural explanation this statement about the unseen order is more comprehensible. Just as religion can be the hardest-to-fake signal of commitment, it can make the strongest appeal to an unseen order. More so than biological evolution, cultural evolution seems to operate on the basis of a continuum where all cultural expressions of hard-to-fake signals of commitment employ costly signals, but some are relatively more costly.

To summarize, in this chapter I argued that Irons' conceptualization of the relationship between morality and religion proved to be problematic for his adaptive

explanation of religion. In his theory of morality he provided compelling theoretical evidence that the propensity to signal commitments in a hard-to-fake manner was selected for because it solved the problem of deceptive free-riding in an environment where large, well-united groups has a fitness advantage. In viewing religion as a hard-to-fake signal of commitment, Irons could not demonstrate that religion was a biological adaptation. Likewise, in claiming that religion is the strongest hard-to-fake signal of commitment, Irons could not show that religion was a biological adaptation. I identified these two particular assertions to be inconsistent with an adaptive perspective and claimed that they resulted in his theoretical evidence not being able to justify his basic contention that religion was the target of biological selection. Although Irons' implicit appeal to pre-existent adaptations fits within all of the alternative evolutionary frameworks, his argument that religion is the most powerful hard-to-fake signal of commitment is inconsistent with all of them except for the bio-cultural approach. Both of Irons' claims are only sound within a bio-cultural approach.

## Chapter Three Notes

<sup>1</sup> By ordinary or mundane I mean non-religion-specific.

<sup>2</sup> Moreover, in his first published article concerning the evolution of religion, “Morality, Religion, and Human Evolution” (1996), referring specifically to Frank’s picture of morality, Irons clarifies that Frank discusses the “actual psychological phenomena” of commitment (380). In one of his latest publications, “An Evolutionary Critique of the Created Co-Creator Concept” (2005), Irons explicitly refers to these behavioral propensities as psychological mechanisms: “The hypothesis that the psychological mechanisms underlying religion are evolved adaptations with the function of enhancing within-group cooperation has been tested by Richard Sosis [...]” (781).

<sup>3</sup> Irons’ conception of empirical support for a theory versus description for the purpose of clarifying a theory is brought to light in the context of his reference to his fieldwork with the Yomut Turkmen. It is clear that he does not consider qualitative interpretation to count as scientific evidence for a model; empirical evidence demands quantitative experimental methods. Irons’ approach to testing his model lies in stark contrast to David Sloan Wilson’s approach to testing his group selection theory of religion. I discuss Wilson’s questionable “empirical” examination of his group selection adaptive theory of religion in a later note in this chapter.

<sup>4</sup> See Sosis’ “Religion and Intragroup Cooperation: Preliminary Results of a Comparative Analysis of Utopian Communities” (2000), Sosis and Bressler’s “Cooperation and Commune Longevity: A Test of the Costly Signaling Theory of Religion” (2003), and Sosis and Ruffle’s “Religious Ritual and Cooperation: Testing for a Relationship on Israeli Religious and Secular Kibbutzim” (2003).

<sup>5</sup> In the section below on coopted adaptations, I further expand on the fact that fitness-enhancement does not guarantee that a trait is a functional adaptation and I discuss how Sosis and his colleague’s analytical results that establish a positive correlation between religion and within-group cooperation cannot differentiate whether religion is an adaptation, coopted adaptation, or coopted byproduct.

<sup>6</sup> Irons explains that the secular ideology of Marxist Leninism calls upon an unseen order in the following manner: people believed that Marx’s and Lenin’s writings could infallibly predict the events of the next century – a feat impossible in the real world.

<sup>7</sup> See Schmitt and Pilcher for a discussion of networks of evidence as well as the means to measure whether they have exemplary breadth and depth.

<sup>8</sup> Alcorta and Sosis state that cognitive scientists and evolutionary psychologists primarily study religion by trying to uncover the psychological mechanisms that produce belief in supernatural agents across cultures. Although most have concluded that religion is a byproduct of ordinary adaptations, Alcorta and Sosis point out that Joseph Bulbulia and Jesse Bering are notable exceptions who argue that religion is not a byproduct. I have already discussed Bulbulia’s costly signaling model of religion. Jesse Bering argues that the propensity to infer supernatural agents’ communicative intent behind natural events may be an adaptation (see Bering and Shackelford 732).

<sup>9</sup> Although claiming that religion is an adaptation goes hand in hand with a *sui generis* understanding of religion, it is not clear whether Irons contends that religion is *sui generis*. Irons’ adaptive explanation of religion invokes several non-religion-specific psychological mechanisms evolved for social functions and even though he argues that religion is more effective than they are at promoting within-group cooperation, religion carries out the same function as these ordinary mental modules. Nevertheless, in arguing that religion itself is an adaptation, his theory implies that there must be something unique to religion that distinguishes religion from other hard-to-fake signals of commitment. However, as discussed above, he does not put his finger on exactly what this unique, *sui generis* quality is.

<sup>10</sup> I clarify that as an essence need not be one element but can encompass multiple factors, the problem with the idea of an essence of religion is that it claims that the list of factors is necessary and sufficient. This leads to the difficulty discussed above that the category of religion becomes narrow and various religious expressions that do not fit the standard mold are left out.

<sup>11</sup> Although Atran and Norenzayan claim that supernatural concepts arose as incidental consequences of such ordinary adaptive mental modules as agency-detection and intuitive ontology, they claim that commitment to supernatural concepts is at least partially sustained because of its role in “the relieving of pervasive existential anxieties” such as “death and deception that forever threaten human life everywhere” (726). They are adamant that the human propensity for supernatural concepts did not arise in the context of a religious function of allaying anxiety caused by the knowledge of death: “All of this isn’t to say that *the* function of religion is to promise resolution of all outstanding existential anxieties any more than *the* function of religion is to neutralize moral relativity and establish social order, to give meaning to an otherwise arbitrary existence, to explain the unobservable origins of things, and so forth. Religion has no evolutionary function *per se*. It is rather that existential anxieties and moral sentiments constitute – by virtue of evolution – ineluctable elements of the human condition; and that the cognitive invention, cultural selection, and historical survival of religious beliefs have resulted, in part, from success in accommodating these elements. There are other factors in this success, involving naturally selected elements of human cognition, such as the inherent susceptibility of religious beliefs to modularized (innate and universal) conceptual and mnemonic processing” (728). In other words, these concepts have been coopted for this usage, which, in part, explains why they are prevalent. Although Atran and Norenzayan do not use the precise term of coopted byproduct, their arguments suggest that they consider the anxiety-alleviating aspect of religious supernatural concepts to be an instance of one.

<sup>12</sup> Sometimes the empirical tests that address one particular factor of a family resemblance definition of religion can be quite far removed from religion. Take, for example, Atran and Norenzayan’s test to determine whether minimally counterintuitive concepts (the category of concepts that includes religious supernatural concepts), are more likely to reach a higher cultural level of distribution than intuitive, bizarre, and maximally counterintuitive concepts because of their mnemonic power (721-22). The statements that they use to exemplify degrees of intuitiveness are mundane and seem incomparable to religious concepts, which evoke meaning and emotion: “giggling seaweed” seems to lack the connotations associated with Lazarus rising from the dead, although both are minimally counterintuitive because they involve one violation of the category of biology. Moreover, religion seems to involve highly counterintuitive concepts: Jesus is not only the Son of God but performed many miracles and rose from the dead.

Barrett and Keil’s research on the difference between theological and religious concepts in terms of degree of counterintuitiveness addresses my second concern that psychological experiments such as Atran and Norenzayan’s do not acknowledge that many concepts in religion are more than minimally counterintuitive. They suggest that our affinity for minimally counterintuitive concepts explains why religious concepts occupying the thoughts of everyday religious people are watered-down versions of cognitively complex theological concepts (see Barrett and Keil’s “Conceptualizing a Nonnatural Entity: Anthropomorphism in God Concepts” (1996) for a detailed discussion; Barrett, “Exploring” 29-30). Barrett and Keil explain that people appear to have at least two parallel but often incompatible God concepts that are used in different contexts – the everyday, “implicit,” minimally counterintuitive, religious God concepts and the abstract, “official,” counterintuitive, theological God concepts (240; Boyer, *Explained* 88). The main difference between the two is that religious concepts are minimally counterintuitive while theological concepts are counterintuitive and do not conform to the violation + default expectations model (Barrett, “Exploring” 30; Boyer and Ramble 538). For example, “theologians might fully appreciate that the god Shiva knows their every thought before conceived, but will still intuitively feel it necessary to make Shiva aware of their thoughts through prayer” (Barrett, “Exploring” 30). Experimentation supports the prediction that religious God concepts are immediately available whereas theological God concepts require conscious reflection. For example, in tasks that allow for conscious monitoring, people say explicitly that God can attend to all sorts of things at the same time but when the task requires fast access and a spontaneous

representation of God, they construe God like any other standard agent who attends to one thing after another (Boyer, *Explained* 88-89). Stewart Guthrie agrees with Barrett that the “more abstruse versions” of religion are reserved for theologians while “[m]ost of us stick closer to home, where matter remains solid and gods are neither totally Other nor totally disembodied” (“Animal” 43-44).

<sup>13</sup> As discussed above, the adaptive approach is typically concerned with why the behavioral trait of religion exists in terms of its function and not in terms of the evolution of the mental modules that generate the various aspects of religious belief and practice. Irons’ adaptive explanation of religion does consider the evolution of human cognition in relation to the evolution of religious behavior for the promotion of intragroup cooperation. I call his theory ‘cognition-sensitive’ in that it posits psychological mechanisms upon which religion is built. For this reason, my comparison and contrast of the adaptive approach and the coopted adaptive approach incorporates a discussion of psychological mechanisms.

<sup>14</sup> Alcorta and Sosis use the term ‘function’ to refer to the fitness-enhancing quality of a trait with adaptive value. They do not differentiate, like Williams, Gould, and myself between ‘function’ and ‘effect’. For the sake of consistency, I employ the term ‘effect’ where appropriate regardless of whether they use the term ‘function.’ For example, even though they contend that religion is a coopted adaptation, in which case its fitness-enhancing role should appropriately be labeled an ‘effect,’ they argue that, “religion’s ability to promote cooperation is its evolutionary function” (329).

<sup>15</sup> I claim that Alcorta and Sosis’ definition of religion is not a true family-resemblance definition. For instance, their statements that four traits “constitute the basic elements of religion” and are “integral components” of the evolved complex of traits that is religion (325) imply that these four factors are necessary and sufficient to a comprehensive understanding of religion.

<sup>16</sup> It is important to note that Alcorta and Sosis do not differentiate between ‘adaptation’ and ‘adaptive value.’ In other words, they contend that religion can be an evolved adaptation even if religion engages ordinary evolved mental modules (326). However, to remain consistent with Williams’ and Gould’s criteria that I adopt in this thesis, as mentioned in previous notes, I hold ‘adaptation’ to refer to a trait that has its origin in the selection for its function, and ‘adaptive value’ to refer to the fitness-enhancing nature of a trait (see Table 1 above, p. 20).

<sup>17</sup> Although Alcorta and Sosis claim that, “neither the content and structure of religious belief systems supports the assertion that such beliefs constitute epiphenomenal by-products” (328) and that, “the ubiquity and ritual commonality of religions across cultures indicate that religion is more than a mere by-product” (“Adaptive” 750), it is impossible to interpret these arguments as supportive of understanding religion as a coopted byproduct. Unlike a byproduct, a coopted byproduct was selected for a fitness-enhancing effect. Like a coopted adaptation, a coopted byproduct modifies the evolutionary mechanism that it coopts (see above Table 1, p. 20). Alcorta and Sosis’ compelling justifications for why religion is not a byproduct do not justify why religion is not a coopted byproduct. This is a possible avenue for future theoretical and empirical research.

<sup>18</sup> This discussion of which evolutionary approach is consistent with Irons’ claim that religion is the most powerful signal of commitment highlights a problem with the state of the empirical testing of Irons’ adaptive explanation of religion. An adaptive theory, coopted adaptive theory, and coopted byproduct theory of religion can each account for Sosis and his colleague’s results that costly rituals are positively correlated with religious commune longevity. Establishing a positive correlation between religion and intragroup cooperation does not establish that the selected function of religion is the enhancement of within-group cooperation. All that Sosis’ and his colleague’s analyses show is that religion is fitness-enhancing (if fitness is defined in terms of cooperation among group members and intragroup cooperation is operationally defined in terms of group longevity). However, this fitness-enhancing quality of religion can be an effect of religion as a coopted adaptation or coopted byproduct, or a function of religion as an adaptation.

<sup>19</sup> See Alcorta and Sosis 332-340 for a detailed explanation of the role of “the sacred” in religion.

<sup>20</sup> Irons does discuss coevolutionary views in his 1991 article “How Did Morality Evolve?” However, this discussion is in the context of a summary of the basics of evolutionary theory in a short section addressing recent theoretical research on humans. His main point in introducing cultural evolution is to dispute the idea that it creates behaviors that hamper the reproductive interests of individuals. He rejects such assumptions that argue that because of culture, the theory of evolution “will not work for human beings as well as it has for animals” (“Evolve” 60). In other words, Irons’ reference to cultural evolution is underdeveloped and not considered in the context of evolutionary theories of religion. However, I mention it because Irons apparently holds that biological and cultural evolution operate in concert with one another.

<sup>21</sup> D. S. Wilson cites Dawkins as subscribing to the category of evolutionary theories of religion that hypothesizes religion to be “a cultural parasite that spreads at the expense of both human individuals and groups” (*Cathedral* 44).

<sup>22</sup> Another way to untangle the relationship between the fitness of cultural units and the carrier’s biological fitness is to understand the types of possible interactions between the two. According to Durham the connections between genetic and cultural inheritance systems can be described in terms of direct interaction and indirect or comparative interaction. In the case of the two types of *interactive modes*, the fitness of memes and genes directly affect one another. On the one hand, with *genetic mediation*, cultural fitness is genetically mediated. On the other hand, with *cultural mediation*, the genetic fitness value of certain practices can be mediated by cultural memes. In the case of the three types of *comparative modes*, gene and meme selection processes have no direct effect on each other. First, in the *enhancement mode* gene and meme selection processes converge to favor certain types of behaviors. Second, in the *neutral mode*, cultural selection strongly favors certain features that do not affect the inclusive fitness of their carriers. Third, in the *opposition mode*, cultural selection strongly favors behaviors that threaten the inclusive fitness of the actors (qtd. in Boyer, *Naturalness* 273-74).

<sup>23</sup> I do not afford D. S. Wilson’s group-level adaptive explanation of religion as much consideration as individual-level adaptive, coopted adaptive, byproduct and coopted byproduct explanations for several significant reasons. One is that Wilson’s evolutionary theory has a minimal network of theoretical and empirical evidence. He is the only evolutionary scientist pursuing a group-level explanation of religion and he has not conducted any quantitative empirical experiments to support it. To test his theory, Wilson looks at whether his predictions can account for the nature of past and present religions as recorded by scholars of religion. He justifies his non-quantitative, instrument-free, and statistic-free method in the following way: if Darwin could establish his theory of evolution on the basis of naturalists’ descriptive information about plants and animals, then he can establish his group-level adaptation hypothesis of religion on the basis of “traditional religious scholarship” about religion (*Cathedral* 87; “Testing” 386).

Although Wilson holds that, “Quantitative methods *refine* but do not *define* scientific inquiry” (“Testing” 386), I contend that quantitative analyses serve a greater function in scientific research than mere refinement in cases when the qualitative information is contradictory or uncertain. Wilson’s functionalist analysis of Calvinism as well as three other religions in *Darwin’s Cathedral* (2002) is unscientific precisely because, as he admits, it is descriptive and not quantitative (*Cathedral* 117). Unlike Irons who actually conducted his own extensive fieldwork with the Yomut Turkmen of Northern Iran but refrained from deeming his observations to be empirical evidence for his theory, Wilson considers his arm-chair speculation based on others’ ethnographic reports to support empirically his group-level explanation of religion. He considers quantifying a descriptive study to have little effect in strengthening the conclusions of the study: in the case of his descriptive study of Calvinism, he contends that, “There is such a consensus among historians that the Catholic Church in Geneva was corrupt and beholden to vested interests outside Geneva in comparison to Calvinism that there seems little point in toting up the numbers, any more than we need to quantify the function of a can opener” (*Cathedral* 117-18). I disagree with Wilson: one must quantify the function of a can opener in order to know what kind of cans the opener will actually open.

Conducting controlled studies, toting up numbers, and running statistical analyses provides evidence of relationships between variables as well as measures of the probability that the findings are due to chance or error. Without controls and quantitative measures of variables, scientists cannot know which variable is responsible for a given effect or lack of effect. Although qualitative assessments may inspire avenues for quantitative analyses, quantitative analyses calculate whether one's observations are statistically relevant.

I contend that Wilson's empirical examination of religion is problematic in another way that further limits the strength of his evolutionary explanation of religion. He admits that "idealized" religion is "more 'purely'" associated with the welfare of the group, ("Testing" 385) and that religion in its practiced form "often deviates" from its ideal form (*Cathedral* 240), but nevertheless focuses his empirical analysis of religion, particularly Calvinism, on its idealized form. "Idealized religion" refers to religious beliefs and actions as written and commanded in official doctrine, while "practiced religion" is religion as it is lived by religious people – it is the way that religious practitioners translate their understanding of their idealized religion into their real-world, everyday beliefs and actions. Wilson recognizes that, even though religions are ideally designed to benefit groups, all actual social systems, religious and secular, are subject to group-disruptive effects of passive free-loading and active exploitation between group members ("Testing" 396). In other words, in practiced religion, among-group selection contends with and not always prevails against within-group selection, but in idealized religion group-level selection operates without competition (*Cathedral* 46). I argue that Wilson's focus on idealized religion despite recognizing that it becomes modified, or "corrupted" ("Testing" 385), when practiced is a serious problem. An *evolutionary* explanation of religion must explain religion as it is actually practiced. Evolution by natural selection is based on differential fitness caused by the *actual*, not the *ideal*, interaction between individuals with different variants of a specific heritable trait and the local environment. Wilson himself argues that the core of evolutionary science is a detailed understanding of organisms in relation to their actual local environments (*Cathedral* 87). When evolutionary science is applied to religion, although an understanding of religious doctrine is important, it is more important to assess the extent to which the religious community follows these required behaviors.

Wilson's questionable methodology and empirical evidence only adds to the general dissatisfaction with his theory. His theoretical evidence rests on an evolutionary approach that is not widely accepted within the community of evolutionary scientists. Even the evolutionary scholars who agree that group selection is possible in theory, contend that it is such a weak evolutionary force in comparison to that of individual level selection that group selection is irrelevant (Williams). In light of the above discussion, if *the* one group-level selectionist considers a group-level adaptation hypothesis of religion difficult to support through a study of practiced religion and instead turns to an analysis of idealized religion, this may confirm the majority of the scientific community's suspicions that group-level selection is possible but too weak to be a significant evolutionary force in real scenarios. Wilson colorfully describes the attitude of evolutionary scientists in the 70s and 80s towards the group selection view of groups as adaptive units like a fart at a cocktail party – something that distinguished people just do not do (Wilson, "Review"). I contend that current theories that posit groups as adaptive units are still like farts at a cocktail party.

That said and keeping in mind that competition among groups and the survival of groups are concerns of group selectionists, like Wilson, I speculate that Irons' theory, especially in its reliance on Alexander's theory of the selection pressure of group warfare, contains group selectionist arguments. Wilson claims that the selected function of religion is to unify groups with the result of more success in warfare. Irons makes the same claim, but it seems out of place in an individual-level selection theory. Moreover, Sosis and Sosis and Bressler's quantitative studies of Irons' theory offer support for Wilson's theory as much so or even more than they do for Irons' theory. They show that there is a positive correlation between religion, more specifically costly rituals, and commune longevity. The dependent variable of commune longevity is more rightly an indicator of group survival than individual survival. Only Sosis and Ruffle's study looks at the effect of ritual participation on individual cooperativeness.

<sup>24</sup> Although D. S. Wilson does not use the terms of 'evoked culture' and 'acquired culture,' in his explanation of why cultural evolution is required to complement biological evolution, he seems to categorize evoked culture as a product of biological evolution and acquired culture as a product of the on-going open-ended process of cultural evolution made possible by innate psychology.

## Conclusion

In this thesis I analyzed William Irons' explanation of religion as an adaptation. I demonstrated that Irons does not justify his claim that religion is an adaptation because several of his arguments are inconsistent within the adaptive framework. For religion to be an adaptation, it must have been directly selected for because it solved a problem at the time of its origin. I contended that Irons does not establish that religion was directly selected for because he superimposes his theory of the evolution of religion onto his theory of the evolution of morality and does not propose a religion-specific mechanism. In his theory of morality, Irons' presents the propensities for indirect sociality, punishment, and commitment and signals of commitment as the solutions to the problem of free-riding in our evolutionary history. However, in his theory of religion, Irons' posits that religion is a hard-to-fake signal that fulfils the same role as any hard-to-fake signal of commitment, namely, to solve the problem of deceptive communications of commitments by increasing their reliability. Because Irons argues that religion is a hard-to-fake signal of commitment, he implies that religion was not directly selected for but instead is grounded in the previous ordinary adaptation for the ability to employ costly and easy-to-monitor signals of commitment. In failing to consider the importance of the evolutionary origin of the mechanisms/structures that are responsible for religion and its fitness-enhancing propensity, Irons appears not to recognize that religion cannot be an adaptation if it is built on previous non-religion-specific adaptations. In arguing that religion is the most effective hard-to-fake signal of commitment, Irons implies that religion is one manifestation of the ordinary propensity to employ costly communications. As religion is built on adaptations for morality and serves the same evolutionary use as a hard-to-fake signal of commitment, religion itself is not a biological adaptation. I argued that these problems render Irons' theoretical evidence incoherent within an adaptive approach and unpersuasive of his overall claim that religion is a biological adaptation for social exchange.

Irons' inability to support his contention that religion was biologically selected for does not preclude his arguments from being coherent with an evolutionary explanation of

religion. In examining Irons' claims for compatibility with alternative evolutionary frameworks, I showed how his appeal to pre-existing adaptations is consistent with all of the other evolutionary approaches that I considered because they bear in mind human cognition. However, his contention that religion is the most powerful hard-to-fake signal of commitment is only consistent with a perspective that considers cultural evolution. Although Irons appears to be Panglossian in that he argues that religion has a particular function, even if religion is not an adaptation *per se*, it can still be fitness-enhancing and thereby have adaptive value. Even if religion was not selected for its function of facilitating the formation and maintenance of large, well-united groups of non-relatives, it can still serve this role in human history, but as an effect rather than a function. Understanding this effect as a result of the cooption of previous adaptations for social exchange or previous byproducts of adaptations for social exchange is not possible because the original function of the original adaptation is the same as its later effect. However, it may be an effect of the cooption of other non-social-exchange-specific adaptations or byproducts: for example, as discussed in Chapter Three in the context of Candace Alcorta and Richard Sosis' belief in supernatural agents aspect of religion, the adaptations for agency detection and intuitive ontology have been coopted for the enhancement of within-group cooperation.

I stated in the introduction that the underlying intention of this thesis was to create a platform for dialogue among evolutionary scientists studying religion and among those evolutionary scientists and religious studies scholars. I briefly consider issues that I have addressed within this thesis that I argue could serve to facilitate an interdisciplinary study of the evolutionary foundation of religion. I divide my key contributions on the basis of those that could enhance discussions among evolutionary scientists examining religion and those that could help establish collaborative efforts among evolutionary scientists and religious studies scholars.

As mentioned in Chapter One, several evolutionary scientists have pointed out that evolutionary explanations are more complete when they consider both why religion exists and how evolved mechanisms generate religious beliefs and practices (Mayr, "Cause" 1503; Irons, "Adaptively" 198; D. S. Wilson, Cathedral 170-177; D. S. Wilson,

“Testing” 392). The integration of the basic principles of behavioral ecologists, who focus on determining the origin of religion, with those of evolutionary psychologists, who concentrate on pinpointing the evolutionary history and current effects of mental modules, would realize a framework from which to propose such comprehensive evolutionary assessments of traits. One reason why Irons’ theoretical evidence does not support his adaptive explanation is that he is a behavioral ecologist interested in explaining the origin of religion in terms of its fitness consequences but who also incorporates into his explanation an appeal to the mechanisms that carry out religion’s function. Irons’ claim that religion exists because it was the solution to the recurrent fitness-relevant problem of free-riding in an environment where large, cohesive groups had the advantage is sound within his adaptive approach. However, his introduction of ordinary social strategies to explain what religion is and how it carries out its function renders his explanation of religion inconsistent with his adaptive approach. Irons seems not to be aware of the stipulation in the cognitive science of religion that for religion to be an adaptation at least one mechanism must have been selected for the function of producing religious belief, experience, or behavior. As he claims that hard-to-fake signals are widespread because of their function of facilitating indirect reciprocity and suggests that the appeal to the unseen order is used to strengthen ordinary commitments, he does not propose any religion-specific mechanisms.<sup>1</sup> Although Irons should be applauded for attempting to explain both the why and how of religion’s evolution, he fails to do so within an adaptive approach because of an apparent lack of understanding of the criteria for adaptation within the evolutionary psychological study of religion. As Sosis and Alcorta point out, “Integrating cognitive and behavioral approaches to the evolutionary study of religion is vital to our progress in understanding religious behaviors and supernatural beliefs” (“Adaptive” 749). The flaws in Irons’ theory give witness to the problems that arise in evolutionary explanations that do not appreciate the principles of evolutionary psychology as well as behavioral ecology.

Scott Atran and Ara Norenzayan’s explanation of religion appears to take a step in the right direction of integration. They contend that neither commitment theories nor cognitive theories alone are adequate to explain religion. “Commitment theories” refers

to adaptive explanations of religion like William Irons' and David Sloan Wilson's that focus on explaining the apparent altruism in religion in terms of long-term fitness benefits to the organism or group. Atran and Norenzayan deem these theories inadequate because they "underplay or disregard" the causal role of psychological mechanisms in their evolutionary appraisals of religion; they claim commitment theories are "mind-blind" to the cognitive constraints on religious beliefs and practices (714). In such theories, religion is attributed to biological adaptations for religious belief and practice but how genes relate to the mental and public representations that actually produce religious belief and practice is left unexplained (Atran, *Gods* 202). They argue that in not accounting for the cognitive peculiarities of religion, commitment theorists "cannot distinguish Marxism from monotheism, or secular ideologies from religious belief." My discussion of Irons' lack of a thorough explanation of belief in an unseen order and his appraisal of Marxism support Atran and Norenzayan's contention. Likewise they argue that cognitive theories that attempt to explain religion as cultural manipulations of the ordinary psychological processes of categorization, reasoning, and remembering too are inadequate. Referring to Pascal Boyer's early works among others, they state that such explanations are "often short on motive and are unable to distinguish Mickey Mouse from Moses, cartoon fantasy from religious belief." A complete theory of the evolutionary foundation requires both an explanation of the cognitive structures that make religious belief and practice possible as well as an explanation of why it is that the costlier the signal of commitment, the more likely it is genuine and the person is trustworthy (714).<sup>2</sup>

To facilitate discussions among evolutionary scientists and between evolutionary scientists and religious studies scholars, a consistent evolutionary terminology needs to be established. As a non-scientist, my examination of current evolutionary theories of religion was challenged by the inconsistent use of evolutionary terms among the scientists. Without a consensus on the meaning of terms, evolutionary scientists are talking past one another. Moreover, humanities scholars are either too confused to participate in the discussion or concentrate on identifying the theoretical inconsistencies of evolutionary terms and frameworks from their hermeneutically-inclined vantage point.

Without standard terms and principles in evolutionary theory, non-scientists may find it difficult to move past considerations of evolutionary terminology and frameworks to consider the actual evolutionary explanations of religion.

This thesis touches upon the philosophy of science by recognizing areas within evolutionary science where a consensus of terms has not yet been reached among the different disciplines and scientists. Although I have used the terms ‘function’ and ‘effect’ in the manner suggested by George Williams, such that ‘function’ is restricted to adaptations while ‘effect’ refers to the fitness-enhancing quality of all of the evolutionary categories other than adaptation, many evolutionary theorists do not abide by this rule and consider any fitness-enhancing role to be a ‘function’ of the trait in question (for example, see Buss et al.). As mentioned above, even Stephen Jay Gould, a stickler for the proper use of terms, is inconsistent with his criteria for what evolutionary categories can have a function. Some evolutionary theorists appear not to differentiate between adaptations and exaptations and call any trait or mechanism that is fitness-enhancing an adaptation and consider its role in increasing fitness to be the function for which it was selected (for example, Irons and D. S. Wilson). However, as discussed in Chapter One, this conflates the historical origins of a trait or mechanism with its current utility. Following Gould’s suggestion, I distinguish adaptations from exaptations on the basis of whether the mechanism generating the trait evolved for that function (adaptation) or was coopted for that effect (exaptation) (see Table 1 above, p. 20). The relationship too between the usage of a trait, whether function or effect, and its adaptive value is not set in stone. I have explored evolutionary theories of religion with the understanding that so long as a trait is fitness-enhancing, it has adaptive value. However, a trait is only an adaptation if it has adaptive value because its fitness-enhancing role is the function for which it was originally selected. The confusion associated with an exaptation having adaptive value could be clarified considering Gould and Vrba’s suggested term of “aptation” which refers to a feature with current utility but makes no claim about its source of origin (Gould, *Structure* 1051). Instead of ‘adaptive value,’ evolutionary scientists could speak of ‘aptive value’ that acknowledges that the trait is fitness-enhancing but remains agnostic whether the mechanism generating the trait originated by

direct selection for that function or originated by direct selection for a different function and was coopted for this novel effect.

Although I am not a scientist, I took several precautions to establish explicitly my understanding of evolutionary theory and evolutionary terms such that my readers would not have to guess the reasons behind my various arguments. The intention of Chapter One was to provide a common background of the basic elements of evolutionary theory and evolutionary explanation while Table 1 was meant to establish a reliable set of criteria for basic evolutionary categories of adaptations, exaptations, and byproducts (see page 19 above). A more comprehensive treatment of evolutionary theories of religion should also consider the theoretical debates concerning the units of selection as well as the relationship between biological and cultural evolution. Most of my discussions of the adaptive framework have been limited to Irons' theory that religion is an adaptation at the individual level. As mentioned in Chapter Three in the context of bio-cultural approaches, David Sloan Wilson offers an adaptive explanation of religion that is similar to Irons', but he claims that religion is an adaptation at the group level. Although explanation by reference to group selection is frowned upon within evolutionary science, given my call for dialogue among all perspectives within evolutionary science, D. S. Wilson's idea of religion as a group-level adaptation for unifying groups may deserve more attention than I have allotted it. Likewise, a more comprehensive treatment of the role of cultural evolution in religion is warranted. As shown in this thesis, one of Irons' problematic arguments is only consistent with an explanation that considers cultural evolution.

Discussions among evolutionary scientists and religious studies scholars are important not only because the evolutionary frameworks deserve philosophical attention, but also because evolutionary conceptualizations of religion require consideration. Religious studies scholars are sensitive to the fact that the preunderstanding of religion affects the framework from which religion will be studied and the approach used to study religion affects the conceptualization of religion. As discussed above, Irons' adaptive reasoning leads him to employ an essentialist definition. Despite recognizing that religion is a "fuzzy" category, Irons defines religion in terms of a necessary and sufficient

substance and function. The structure of this definition prevents his theory of religion from accounting for the aspects of religion that do not concern hard-to-fake signaling or a social function. Also, it averts those who empirically test his costly signaling model of religion from analyzing different facets of religion that may have an independent evolutionary history. Finally, it hampers the scientists conducting the empirical examinations from explaining various results: for example, Richard Sosis and Eric Bressler could not explain why the level of hard-to-fake signaling in religious and secular communities did not have the same effect on within-group cooperation and had to go outside Irons' costly signaling model to considerations of a relationship between religion and the sacred to offer an explanation of the finding.

As discussed in the previous chapter, a family resemblance approach to defining religion seems to circumvent many of the above-listed problems. Such a definition better accommodates the multifaceted nature of religion and appreciates that religious beliefs, practices, and experiences can be expressed very differently across cultures and time and yet still be encompassed within one meaningful category. As I showed, several evolutionary scientists that are sensitive to the evolution of the mind have recently presented models of religion adopting such definitions (see Atran and Norenzayan; Alcorta and Sosis).<sup>3</sup>

The evolutionary study of religion is comparatively new to the evolutionary study of physiology and even social behavior. As much work remains to be done before any evolutionary explanations of religion gather definitive networks of evidence, this is the time for religious studies scholars to join the project and contribute their expertise on the subject of religion.

---

Conclusion Notes

<sup>1</sup> As mentioned in Chapter Three, even if Irons tried to establish belief in the unseen order as grounded in a religion-specific mechanism, such an argument would be significantly challenged by the research of byproduct theorists that suggests that belief in supernatural agents is explicable in terms of ordinary mental modules.

<sup>2</sup> Like Atran and Norenzayan, comparative scholar of religion Ilkka Pyysiainen too claims that he is working on an evolutionary model of that combines the perspectives of cognitive theorists, including Boyer and Atran, with commitment theorists such as D. S. Wilson and Sosis (personal email, 8 Feb. 2005).

<sup>3</sup> Alcorta and Sosis adopt a four-part definition of religion. However, I contend that Alcorta and Sosis' definition of religion is not a true family resemblance definition because of several statements that they make in explaining it. They state that there are four traits that "constitute the basic elements of religion" and are "integral components" of the evolved complex of traits that is religion (325). These statements imply that these four factors are necessary and sufficient to a comprehensive understanding of religion. Their bold claim that their definition's synthesis of previously independent conceptualizations of religion not only "encompasses religion's cross-culturally recurrent features," but also "captures that which differentiates the religious from the secular" (324) seals the case that their definition is essentialist. Nevertheless, at least they recognize that there is more to religion than a costly signal of commitment and their definition of religion better accounts for the multifaceted and complex nature of religion than Irons' 'one substance and one function' definition.

## Bibliography

- Alexander, R. D. The Biology of Moral Systems. New York: Aldine de Gruyter, 1987.
- Anttonen, Veiko. "Identifying the Generative Mechanisms of Religion: The Issue of Origin Revisited." Current Approaches in the Cognitive Science of Religion. Pyysiäinen and Anttonen 14-37.
- Alcorta, Candace S. and Richard Sosis. "Ritual, Emotion, and Sacred Symbols: The Evolution of Religion as an Adaptive Complex." Human Nature 16 (2005): 323-359.
- Atran, Scott. "Adaptationism for Human Cognition: Strong, Spurious or Weak?" Mind and Language 20 (2005): 39-67.
- . In Gods We Trust: The Evolutionary Landscape of Religion. Oxford: Oxford UP, 2002.
- Atran, Scott, and Ara Norenzayan. "Religion's Evolutionary Landscape: Counterintuition, Commitment, Compassion, Communion." Behavioral and Brain Sciences 27 (2004): 713-770.
- . "Why Minds Create Gods: Devotion, Deception, Death and Arational Decision Making." Behavioral and Brain Sciences 27 (2004): 754-770.
- Ayala, F. J. "Human Nature: Biology Precedes, Culture Transcends: An Evolutionist's View of Human Nature." Zygon 33 (1998): 507-523.
- Barbour, Ian G. Religion and Science: Historical and Contemporary Issues. New York: HarperCollins, 1997.
- . When Science Meets Religion: Enemies, Strangers, or Partners? New York: HarperSanFrancisco, 2000.
- Barrett, Justin L. "Exploring the Natural Foundations of Religion." Trends in Cognitive Sciences 4 (2000): 29-34.
- . Why Would Anyone Believe in God? Walnut Creek, CA: AltaMira Press, 2004.
- Barrett, Justin L., and Frank C. Keil. "Conceptualizing a Nonnatural Entity: Anthropomorphism in God Concepts". Cognitive Psychology 31 (1996): 219-247.
- Bering, Jesse M., and Todd K. Schackelford. "Supernatural Agents May Have Provided Adaptive Social Information." Behavioral and Brain Sciences 27 (2004): 732-733.

- Boehm, Christopher. "Conflict and the Evolution of Social Control." Journal of Consciousness Studies 7.1-2 (2000): 79-101.
- . "Impact of the Human Egalitarian Syndrome on Darwinian Selection Mechanics." American Naturalist 150 (1997): S100-S121.
- . "The Origin of Morality as Social Control – Response to Commentary Discussion." Journal of Consciousness Studies 7.1-2 (2000): 149-183.
- Boyer, Pascal. "Evolutionary Psychology and Cultural Transmission." The American Behavioral Scientist 43 (2000): 987-1000.
- . "Functional Origins of Religious Concepts: Ontological and Strategic Selection in Evolved Minds." Journal of the Royal Anthropological Institute 6 (2000): 195-214.
- . Religion Explained: The Evolutionary Origins of Religious Thought. New York: Basic Books, 2001.
- . "Religious Thought and Behaviour As By-Products of Brain Function." Trends in Cognitive Sciences 7 (2003): 119-124.
- . "Religion, Evolution, and Cognition." Rev. of In God We Trust: The Evolutionary Landscape of Religion, by Scott Atran, and Darwin's Cathedral: Evolution, Religion, and The Nature of Society, by David Sloan Wilson. Current Anthropology 45 (2004): 430-433.
- . The Naturalness of Religious Ideas: A Cognitive Theory of Religion. Berkeley: U of California P, 1994.
- . "Why Do Gods and Spirits Matter at All?" Current Approaches in the Cognitive Science of Religion. Pyysiäinen and Anttonen 68-92.
- . "Why is Religion Natural?" Skeptical Inquirer 28.2 (2004): 25-32.
- Boyer, Pascal, and Charles Ramble. "Cognitive Templates for Religious Concepts: Cross-cultural Evidence for Recall of Counter-intuitive Representations." Cognitive Science 25 (2001): 535-564.
- Bulbulia, Joseph. "Are There Any Religions? An Evolutionary Exploration." Method and Theory in Study of Religion 17 (2005): 71-100.
- . "Area Review: The Cognitive and Evolutionary Psychology of Religion." Biology and Philosophy 19 (2004): 655-686.

- . Before Eden: Religion and The Evolved Mind. Diss. Princeton U, 2001. Ann Arbor: UMI, 2001. ATT 3007198.
- . "Nature's Medicine: Religiosity As An Adaptation for Health and Co-operation." Evolutionary Approaches: Evolution and God: How Evolution Shaped the Religious Brain. Vol. 1 of Where God and Man Meet. How the Brain and Evolutionary Sciences are Revolutionizing Our Understanding of Religion and Spirituality. Ed. Patrick McNamara. 3 vols. West Port, CT: Praeger, forthcoming. Joseph Bulbulia home page. Victoria University of Wellington School of Art History, Classics, and Religious Studies. 5 June 2006 <[http://www.vuw.ac.nz/religion/staff/joseph\\_bulbulia/index.html](http://www.vuw.ac.nz/religion/staff/joseph_bulbulia/index.html)>.
- . "Religious Costs as Adaptations that Signal Altruistic Intention." Evolution and Cognition 10 (2004): 19-42.
- Buss, David M., et al. "Adaptations, Exaptations, and Spandrels." American Psychologist 53 (1998): 533-548.
- Cosmides, Leda, and John Tooby. "Cognitive Adaptations for Social Exchange." The Adapted Mind: Evolutionary Psychology and the Generation of Culture. Ed. Jerome H. Barkow, Leda Cosmides, and John Tooby. New York: Oxford UP, 1992. 163-228.
- Cronk, Lee. "Evolutionary Theories of Morality and the Manipulative Use of Signals." Zygon 29 (1994): 81-101.
- . "The Use of Moralistic Statements in Social Manipulation: A Reply to Roy A. Rappaport." Zygon 29 (1994): 351-355.
- Darwin, Charles. The Descent of Man. 1874. Amherst, NY: Prometheus, 1998.
- . The Origin of Species: By Means of Natural Selection. 1859. Ed. J. W. Burrows. London: Penguin, 1985.
- Dawkins, Richard. The Selfish Gene. 1976. Oxford: Oxford UP, 1989.
- Dennett, Daniel C. Darwin's Dangerous Idea. New York: Simon and Schuster, 1995.
- Evolutionary Neurobehavior Laboratory. Dept. home page. Boston University Medical Campus. 18 Apr. 2005. <<http://www.bumc.bu.edu/Dept/Content.aspx?DepartmentID=481&PageID=9974>>.
- Gould, Stephen Jay. "Exaptation: A Crucial Tool for Evolutionary Psychology." Journal of Social Issues 47 (1991): 43-65.

- . The Structure of Evolutionary Theory. Cambridge: Harvard UP, 2002.
- Gould, Stephen Jay, and Richard C. Lewontin. "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme." Proceedings of the Royal Society of London 205 (1978): 581-598. Rpt. in Conceptual Issues in Evolutionary Biology. Ed. Elliott Sober. Cambridge: MIT P, 1994. 73-90.
- Gould, Stephen Jay, and Elisabeth Vrba. "Exaptation – A Missing Term in the Science of Form." Paleobiology 8.1 (1982): 4-15. Rpt. in The Philosophy of Biology. Ed. David Hull and Michael Ruse. Oxford: Oxford UP, 1998. 52-71.
- Guthrie, Stewart E. "Animal Animism: Evolutionary Roots of Religious Cognition." Current Approaches in the Cognitive Science of Religion. Pyysiäinen and Anttonen 38-67.
- . Faces in the Clouds: A New Theory of Religion. New York: Oxford UP, 1993.
- Hamilton, William D. "The Genetical Evolution of Social Behavior." Journal of Theoretical Biology 37 (1964): 1-52.
- Hefner, Philip. "Theological Perspectives on Morality and Human Evolution." Richardson and Wildman 401-423.
- Iannaccone, Laurence. "Sacrifice and Stigma: Reducing Free-Riding in Cults, Communes, and Other Collectives." Journal of Political Economy 100 (1992): 271-291.
- Irons, William. "Adaptively Relevant Environments Versus the Environment of Evolutionary Adaptedness." Evolutionary Anthropology 6 (1998): 194-204.
- . "An Evolutionary Critique of the Created Co-Creator Concept." Zygon 39 (2004): 773-790.
- . "How Did Morality Evolve?" Zygon 26 (1991): 49-89.
- . "How Has Evolution Shaped Human Behavior? Richard Alexander's Contribution to an Important Question." Evolution and Human Behavior 26 (2005): 1-9.
- . "Morality, Religion, and Human Evolution." Richardson and Wildman 375-399.
- . "Religion As a Hard-to-Fake Sign of Commitment." Evolution and the Capacity for Commitment. Ed. R. M. Nesse. New York: Russell Sage Foundation, 2001. 292-309.
- . "Why are Humans Religious: An Inquiry into the Evolutionary Origin of Religion." Currents in Theology and Mission 28 (2001): 357-368.
- The Institute of Cognition and Culture. Institute home page. Queen's U, Belfast. 2 Mar. 2005. <<http://www.qub.ac.uk/icc/>>.

- Kamppinen, Matti. "Explaining Religion: Cognitive and Evolutionary Mechanisms." Current Approaches in the Cognitive Science of Religion. Pyysiäinen and Anttonen 260-270.
- Kirkpatrick, Lee. "Toward an Evolutionary Psychology of Religion and Personality." Journal of Personality 67 (1999): 921-951.
- Lloyd, Elizabeth A., and Marcus W. Feldman. "Evolutionary Psychology: A View From Evolutionary Biology." Psychological Inquiry 13 (2002): 150-156.
- Mayr, Ernst. "Cause and Effect in Biology." Science 134 (1961): 1501-1506.
- . "How to Carry Out the Adaptationist Program?" American Naturalist 121 (1983): 324-334.
- McGrath, Alister E. Science and Religion: An Introduction. Oxford: Blackwell, 1999.
- Nesse, Randolph M., ed. Evolution and the Capacity for Commitment. New York: Russell Sage Press, 2001.
- Pals, Daniel L. "Is Religion a Sui Generis Phenomenon?" JAAR LV (1987): 259-284.
- Peacocke, Arthur. "A Map of Scientific Knowledge: Genetics, Evolution, and Theology." Ted Peters 189-210.
- Peters, Karl E. "Pluralism and Ambivalence in the Evolution of Morality." Zygon 38 (2003): 333-354.
- Peters, Ted, ed. Science and Theology: The New Consonance. Boulder, CO: Westview P, 1998.
- Peterson, Gregory R. Rev. of Darwin's Cathedral: Evolution, Religion, and The Nature of Society, by David Sloan Wilson. Christian Century 120.13 (2003): 41.
- Pigliucci, Massimo, and Jonathan Kaplan. "The Fall and Rise of Dr. Pangloss: Adaptationism and the *Spandrels* Paper 20 Years Later." Trends in Ecology and Evolution 15 (2000): 66-70.
- Pyysiäinen, Ilkka. "Bridge Over Troubled Water: Crossing Disciplinary Boundaries." Method and Theory in the Study of Religion 16 (2004): 336-347.
- . "Buddhism, Religion, and the Concept of 'God'." Numen 50 (2003): 147-171.
- . Email by the author. 8 Feb. 2005.
- . "Gods, Genes, and Passions." Rev. of In God We Trust: The Evolutionary Landscape of Religion, by Scott Atran. Journal of Cognition and Culture 3 (2003): 175-186.
- . How Religion Works: Towards a New Cognitive Science of Religion. Leiden: Brill, 2003.

- . "Introduction: Cognition and Culture in the Construction of Religion." Current Approaches in the Cognitive Science of Religion. Pyysiäinen and Anttonen 1-13.
- . "Religion and the Counter-Intuitive." Current Approaches in the Cognitive Science of Religion. Pyysiäinen and Anttonen 110-132.
- Pyysiäinen, Ilkka, and Veikko Anttonen, eds. Current Approaches in the Cognitive Science of Religion. London: Continuum, 2002.
- Pyysiäinen, Ilkka, Marjaana Lindeman, and Timo Honkela. "Counterintuitiveness as the Hallmark of Religiosity." Religion 33 (2004): 341-355.
- Richardson, Mark, and Wesley J. Wildman, eds. Religion and Science: History, Method, Dialogue. New York: Routledge, 1996.
- Rue, Loyal. "Religion Generalized and Naturalized." Zygon 35 (2000): 587-602.
- Salter, Benson. "Biology and Religion: On Establishing a Problematic." Method & Theory in the Study of Religion 11 (1999): 386-395.
- . Conceptualizing Religion: Immanent Anthropologists, Transcendent Natives, and Unbounded Categories. Leiden: Brill, 1993.
- Schmitt, David P., and June J. Pilcher. "Evaluating Evidence of Psychological Adaptation: How Do We Know One When We See One?" Psychological Science 15 (2004): 643-649.
- Sesardic, Neven. "Review Article: Altruism." Rev. of Unto Others: The Evolution and Psychology of Unselfish Behavior, by Sober, Elliott, and David Sloan Wilson. British Journal for the Philosophy of Science 50 (1999): 457-466.
- Smith, Eric A., Monique Borgerhoff Mulder, and Kim Hill. "Controversies in the Evolutionary Social Sciences: A Guide for the Perplexed." Trends in Ecology and Evolution 16 (2001): 128-135.
- Sober, Elliott. "Models of Cultural Evolution." Trees of Life: Essays in the Philosophy of Biology. Ed. P. Griffiths. Kluwer, 1991. 17-38. Rpt. in Conceptual Issues in Evolutionary Biology. Ed. Elliott Sober. Cambridge: MIT P, 1994. 477-492.
- . "Six Sayings About Adaptationism." Philosophical Exchange 26 (1996): 41-57. Rpt. in The Philosophy of Biology. Ed. David Hull and Michael Ruse. Oxford: Oxford UP, 1998. 72-86.

- . "What is Evolutionary Altruism?" Canadian Journal of Philosophy 14 (1988): 75-99. Rpt. in The Philosophy of Biology. Ed. David Hull and Michael Ruse. Oxford: Oxford UP, 1998. 459-478.
- Sober, Elliott, and David Sloan Wilson. "A Critical Review of Philosophical Work on the Units of Selection Problem." Philosophy of Science 61 (1994): 534-555. Rpt. in The Philosophy of Biology. Ed. David Hull and Michael Ruse. Oxford: Oxford UP, 1998. 198-220.
- . Unto Others: The Evolution and Psychology of Unselfish Behavior. Cambridge: Harvard UP, 1998.
- Sosis, Richard. "The Adaptive Value of Religious Ritual." American Scientist 92 (2004): 166-172.
- . "Religion and Intragroup Cooperation: Preliminary Results of a Comparative Analysis of Utopian Communities." Cross-Cultural Research 34 (2000): 70-87.
- . "Religious Behaviors, Badges, and Bans: Signaling Theory and the Evolution of Religion." Evolutionary Approaches: Evolution and God: How Evolution Shaped the Religious Brain. Vol. 1 of Where God and Man Meet. How the Brain and Evolutionary Sciences are Revolutionizing Our Understanding of Religion and Spirituality. Ed. Patrick McNamara. 3 vols. West Port, CT: Praeger, forthcoming. Richard Sosis home page. University of Connecticut Department of Anthropology. 5 June 2006 <<http://www.anth.uconn.edu/faculty/sosis/publications/index.htm>>.
- . Rev. of Darwin's Cathedral: Evolution, Religion, and The Nature of Society, by David Sloan Wilson. Evolution and Human Behavior 24 (2003): 137-143.
- . "Why Aren't We All Hutterities? Costly Signaling Theory and Religious Behavior." Human Nature 14 (2003): 91-127.
- Sosis, Richard, and Candace Alcorta. "Is Religion Adaptive?" Behavioral and Brain Sciences 27 (2004): 749-750.
- . "Signaling, Solidarity, and the Sacred: The Evolution of Religious Behavior." Evolutionary Anthropology 12 (2003): 264-267.
- Sosis, Richard, and Eric R. Bressler. "Cooperation and Commune Longevity: A Test of the Costly Signaling Theory of Religion." Cross-Cultural Research 37 (2003): 211-239.

- Sosis, Richard, and Bradley J. Ruffle. "Religious Ritual and Cooperation: Testing for a Relationship on Israeli Religious and Secular Kibbutzim." Current Anthropology 44 (2003): 713-722.
- Symons, Donald. "On the Use and Misuse of Darwinism in the Study of Human Behavior." The Adapted Mind: Evolutionary Psychology and the Generation of Culture. Ed. Jerome H. Barkow, Leda Cosmides, and John Tooby. New York: Oxford UP, 1992. 137-162.
- Tooby, John, and Leda Cosmides. "The Psychological Foundations of Culture." The Adapted Mind: Evolutionary Psychology and the Generation of Culture. Ed. Jerome H. Barkow, Leda Cosmides, and John Tooby. New York: Oxford UP, 1992. 19-136.
- Trivers, Robert L. "The Evolution of Reciprocal Altruism." Quarterly Review of Biology 46 (1971): 35-57.
- Understanding Evolution: An Evolution Website for Teachers. Ed. Roy Caldwell et. al. 2005. U of California Museum of Paleontology, Berkeley. 15 Feb. 2005 <<http://evolution.berkeley.edu/evosite/evohome.html>>.
- Williams, George C. Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought. Princeton: Princeton UP, 1966.
- Wilson, David Sloan. "A Critique of R.D. Alexander's Views on Group Selection." Biology and Philosophy 14 (1999): 431-449.
- . "Altruism and Organism: Disentangling the Themes of Multilevel Selection Theory." American Naturalist 150 (1997): S122-S134.
- . Darwin's Cathedral: Evolution, Religion, and The Nature of Society. Chicago: U of Chicago P, 2002.
- . "On the Relationship between Evolutionary and Psychological Definitions of Altruism and Selfishness." Biology and Philosophy 7 (1992): 61-68. Rpt. in The Philosophy of Biology. Ed. David Hull and Michael Ruse. Oxford: Oxford UP, 1998. 479-487.
- . "Religious Groups as Adaptive Units." History and Philosophy of the Life Sciences 23 (2001): 467-503.
- . Rev. of Nonzero: The Logic of Human Destiny, by Robert Wright. The Skeptics Society and Skeptic Magazine. Home page. 2004. 17 May 2005 <<http://www.skeptic.com/>>.

----. "Testing Major Evolutionary Hypotheses About Religion with a Random Sample." Human Nature 16 (2005): 382-409.

Wilson, David Sloan, and Elliott Sober. "More on Group Selection and Human Behavior." Behavioral and Brain Sciences 19 (1996): 782-787.

---. "Reintroducing Group Selection to the Human Behavioral Sciences." Behavioral and Brain Sciences 17 (1994): 585-654.

Wilson, Edward O. Consilience: The Unity of Knowledge. New York: Knopf, 1998.

---. On Human Nature. Cambridge: Harvard UP, 1978.

---. Sociobiology: The Abridged Edition. 1975. Cambridge: Harvard UP, 1980.