ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/321572176

The Evolution of Morality

Book · January 2016

DOI: 10.1007/978-3-319-19671-8

CITATIONS
17

READS
2,278

2 authors, including:

Todd K Shackelford
Oakland University
482 PUBLICATIONS **16,964** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

The Psychology of Conspiracy Theories View project

Study of homosexual and heterosexual men's ejaculate quality. Testing predictions from an alternative reproductive strategy hypothesis. View project

All content following this page was uploaded by Todd K Shackelford on 07 April 2020.

The user has requested enhancement of the downloaded file.

Evolutionary Psychology
Series Editors: Todd K. Shackelford · Viviana A. Weekes-Shackelford

Todd K. Shackelford
Ranald D. Hansen *Editors*

The Evolution of Morality

Springer

Evolutionary Psychology

Series Editor

Todd K. Shackelford
Rochester, MI, USA

Viviana A. Weekes-Shackelford
Rochester, MI, USA

For further volumes:
http://www.springer.com/series/10583

shackelf@oakland.edu

shackelf@oakland.edu

Todd K. Shackelford • Ranald D. Hansen
Editors

The Evolution of Morality

Springer

shackelf@oakland.edu

Editors
Todd K. Shackelford
Department of Psychology
Oakland University
Rochester, MI, USA

Ranald D. Hansen
Department of Psychology
Oakland University
Rochester, MI, USA

ISSN 2197-9898 ISSN 2197-9901 (electronic)
Evolutionary Psychology
ISBN 978-3-319-19670-1 ISBN 978-3-319-19671-8 (eBook)
DOI 10.1007/978-3-319-19671-8

Library of Congress Control Number: 2015946057

Springer Cham Heidelberg New York Dordrecht London
© Springer International Publishing Switzerland 2016
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

shackelf@oakland.edu

With deep sadness, this volume is dedicated to the memory of Randy Hansen, co-organizer of this and two previous conferences and much else at Oakland University, who died suddenly and unexpectedly in December 2014.

shackelf@oakland.edu

shackelf@oakland.edu

Preface

In March 2014, we welcomed dozens of scholars from North America, Europe, and Africa to join us at Oakland University in Rochester, Michigan, for a daylong interdisciplinary conference on "The Evolution of Morality." We invited as panelists some of the leading scholars in morality from many different disciplines, including psychology, neuroscience, criminology, biology, anthropology, archeology, law, philosophy, and medicine. Each of these scholars had conducted and published substantial work addressing morality from an evolutionary perspective. This volume showcases the groundbreaking empirical and theoretical work from several of these panelists and other distinguished conference guests.

The volume is presented in five parts. Part I includes three chapters that focus on psychological adaptation and developmental processes. In chapter "The Tripartite Theory of Machiavellian Morality: Judgment, Influence, and Conscience as Distinct Moral Adaptations," Kelly Asao and David M. Buss propose a tripartite theory of Machiavellian morality in which moral judgment, moral influence, and moral conscience are functionally distinct moral adaptations. The authors argue that moral judgment is an adaptation designed to determine how exploitative or benefit-bestowing a conspecific is and to use that information when selecting relationship partners. Moral influence is designed to identify cost-effective means of altering the behavior of others to be less cost-inflicting and more benefit-bestowing. Moral conscience is an adaptation designed to guide one's own behavior toward others to avoid ramifications from other's moral judgment and influence mechanisms. The authors illustrate the application of this tripartite framework of Machiavellian morality with two examples, sexual infidelity and property theft. Asao and Buss close with a thoughtful discussion of the potential for this framework to clarify some of the ambiguity in the morality literature and to refocus attention on novel areas of research.

In chapter "Morality as Cooperation: A Problem-Centred Approach," Oliver Scott Curry presents a new theory of morality as cooperation. This theory uses the mathematics of cooperation to identify the many distinct problems of cooperation and their solutions; and it predicts that it is the solutions deployed by humans that constitute "morality." According to Curry, therefore, morality is a collection of

shackelf@oakland.edu

biological and cultural solutions to the problems of cooperation and conflict evolutionarily recurrent in human social life. This theory generates a comprehensive taxonomy of moral values—what Curry refers to as a Periodic Table of Ethics—that includes obligations to family, group loyalty, reciprocity, bravery, respect, fairness, and property rights. Curry argues that morality as cooperation generates predictions about the structure and content of human morality and that these predictions can be tested against those of rival theories, thereby revealing that the study of morality is simply another branch of science.

Chapter "An Evolving and Developing Field of Study: Prosocial Morality from a Biological, Cultural, and Developmental Perspective" rounds out the first part of the volume. In this chapter, developmental psychologist Gustavo Carlo and colleagues open by noting that most theoretical accounts of human morality emphasize the role of biological, psychological, environmental, or developmental processes. The authors concede that these theories have guided much research and have advanced our understanding of morality. However, theories and research examining the role of culture-related processes are less common, and, according to the authors, there is a need for integrative approaches. Carlo and colleagues briefly review biologically based and developmental research, discuss definitional issues, and present a model that incorporates culture-related processes. The model highlights biological, cultural, and environmental mechanisms, sociocognitive and socioemotive traits, and culture-related processes. The authors close the chapter with calls for research that addresses biology X environment interactions and refinements in conceptualizations of morality and moral behaviors.

Part II of the volume includes four chapters that broadly address philosophical and ethical perspectives on morality. In chapter "Evolutionary Awareness: A Metacognitive Framework for Ethics," Gregory Gorelik and Todd K. Shackelford advance the concept of "evolutionary awareness," a metacognitive framework that examines human thought and emotion from a naturalistic, evolutionary perspective. The authors begin by discussing the evolution and current functioning of the moral foundations on which their framework rests. Next, they address the possible applications of such an evolutionarily informed ethical framework to several domains of human behavior: sexual maturation, mate attraction, intrasexual competition, culture, and the separation between various academic disciplines. Gorelik and Shackelford close their chapter with a discussion of the ways in which an evolutionary awareness can inform our cross-generational activities—which they refer to as "intergenerational extended phenotypes"—by helping us to construct a better future for ourselves, for other sentient beings, and for our environment.

Tyler Millhouse and colleagues open chapter "The Containment Problem and the Evolutionary Debunking of Morality" by highlighting recent work arguing that existing evidence does not support the claim that *moral cognition*, understood as a specific form of normative cognition, is a product of evolution. The authors of this recent work suggest, instead, that the evidence only supports the claim that a general capacity for *normative cognition* evolved. These authors argue that if this is the case, then the prospects for evolutionary debunking arguments of morality are bleak: A debunking argument which relied on the fact that normative cognition in

shackelf@oakland.edu

general evolved seems like it would debunk *all* areas of normative belief, including the epistemic norms upon which the argument relies. Millhouse and colleagues accept, for the sake of argument, the claim that specifically moral cognition did not evolve. However, they reject the contention that this critically undermines evolutionary debunking arguments of morality. A number of strategies are available to solve what Millhouse and colleagues refer to as the "containment problem" of how to effectively debunk morality without *thereby* debunking normative cognition. Furthermore, and according to Millhouse and colleagues, the debunking argument need not rely even on the claim that normative cognition in general evolved. So long as at least some aspects of moral cognition have evolved, this may be sufficient to support an evolutionary debunking argument against many of our moral beliefs. Thus, according to Millhouse and colleagues, even if these previous authors are correct that specifically moral cognition did not evolve, research in evolutionary psychology may have important implications for moral philosophy.

The claim that "life is good" is a popular mantra among the cheery and those aspiring to such "positive thinking." In opposition to this optimism, David Benatar argues in chapter "Life Is Not Good" that while some lives are better than others, no life is good enough to count as non-comparatively good. Benatar concedes that this conclusion will strike many as outrageous, and thus he also addresses good reasons why we should distrust positive assessments of the quality of life. Benatar considers and rejects various "secular theodicies"—attempts to reconcile the vast amount of evil in life with the claim that "life is good." Benatar closes his chapter with a careful consideration of what does and what does not follow from the grim view that he defends.

In the closing chapter of Part II, Sarah Perry addresses antinatalism—the view that it is morally wrong to procreate. Perry considers the history of human fertility, in particular the modern fertility transition of the past two centuries, in light of the ethical beliefs inferable from fertility practices. Perry argues that the patterns of change in fertility suggest that parents, faced with a version of Derek Parfit's "mere addition paradox," increasingly reject the "Repugnant Conclusion" and have fewer children for the good of these children, ignoring the purported benefit that existence would give to children they never have. Perry closes her chapter addressing the other end of life, offering a thoughtful analysis of the social and cultural evolution of suicide.

Part III includes two chapters that address morality in nonhumans. Katie Hall and Sarah F. Brosnan argue in chapter "A Comparative Perspective on the Evolution of Moral Behavior" that humans are not alone in the animal kingdom in displaying moral behavior. According to the authors, precursors to moral behavior exist in rudimentary form in many species, including our closest phylogenetic relatives, the nonhuman primates, and have evolved into the more complex moral behavior seen in humans. Hall and Brosnan argue that moral behavior functions to reduce tension in social groups to thereby afford cooperative, peaceful interactions that are essential for groups to exist. The authors consider four of these behaviors: conflict resolution, reciprocity, reactions to inequity, and empathy. Hall and Brosnan conclude that by studying these behaviors in other species, we may gain insight into the evolution

shackelf@oakland.edu

of moral behaviors, what the mechanisms are that produce these behaviors, how these behaviors develop in the individual, and, through a comparative approach in particular, the function of these behaviors.

In chapter "Helping Another in Distress: Lessons from Rats," Peggy Mason argues that, in mammalian communities, affective communication and pro-social acts support social cohesion, which in turn ancestrally increased an individual's chances of survival and reproductive success. Mason reviews her own and others' research indicating that pro-social behavior occurs in rodents as well as in nonhuman primates, reflecting the value of social cohesion and affective communication to mammals of all ages and both sexes. Mason's own groundbreaking research documents that, given a rat-appropriate challenge, adult rats help another rat in distress by freeing it from a restraining tube. Rats perform this pro-social act repeatedly and at shorter and shorter latencies, acting consistently and intentionally. This helping behavior occurs even if social contact between the helper and the recipient is prevented, by having the trapped rat released into a separate space. This result documents that the helper rat helps independent of earning an immediate social reward. As in humans, rats help strangers as well as individuals with whom they are familiar. In the case of rats, help is extended to unfamiliar rats but only if those rats are of a familiar type, even if the type is not the same as their own biological type. Mason reviews research in her lab documenting that cohousing with a single rat of a different stock is enough to confer familiarity to all rats of that stock. Mason suggests that the helping behavior test in rats might be applied analogously in humans to disambiguate cultural and biological influences on human social behavior.

Part IV of the volume includes two chapters that address work at the interface of evolutionary psychology and religious beliefs and behavior. Yael Sela and colleagues address religiously motivated violence as a downstream consequence of processes of sexual selection.

The authors open the chapter by noting that relying on religion as the basis of one's morality can be problematic. Although religion can motivate positive behaviors and cooperation, it also motivates and exacerbates violence in particular contexts. Seal and colleagues first provide a brief overview of human sexual selection from an evolutionary psychological perspective. They next discuss how and why an evolutionary perspective and, in particular, the concepts of intersexual and intrasexual competition may be useful in understanding religiously motivated violence. The authors then present an overview of the research addressing several types of religiously motivated violence, such as mate guarding and controlling behaviors, wife-beating and uxoricide, "honor" killing, child abuse and filicide, male and female genital mutilation, suicide, group violence and war, and terrorism, including suicide terrorism. Sela and colleagues close the chapter by highlighting the potential advantages that religiously motivated violence may have provided ancestrally within a sexual selection framework, and they conclude with suggestions for future research.

James R. Liddle begins chapter "Disentangling Religion and Morality: An Analysis of Religiosity in the United States" noting that religion is considered by many practitioners to form the foundation of morality. However, religiosity varies

shackelf@oakland.edu

substantially at the individual and societal level. According to Liddle, understanding this variation from an evolutionary perspective can aid in disentangling religion and morality. Liddle summarizes his recent research designed to replicate and extend previous findings regarding the "Secure Society Theory" of religiosity, which states that religiosity varies with the extent to which one feels secure in one's environment. The relationship between individual perceptions of societal security—as opposed to national indicators—and religiosity has yet to be tested. Liddle's research addressed this by analyzing US data from the General Social Survey, supplemented by the Federal Bureau of Investigation and Census data. Liddle reviews the results of this research, which indicates that the extent to which one feels safe walking around one's neighborhood at night predicts religiosity, even when crime rate, poverty rate, age, sex, and race are controlled statistically. Additionally, time series analyses of data from 1980 to 2012 provide partial support for Secure Society Theory, with neighborhood fear and poverty predicting future religiosity.

The final three chapters comprise Part V of the volume, and these chapters address the evolution of morality in the contexts of politics, the law, and game theory. In chapter "The Evolved Functions of Procedural Fairness: An Adaptation for Politics," Bøggild and Petersen review recent research on the evolved functions of procedural fairness, making a clear case for the existence of adaptations for politics. According to the authors, politics is the process of determining resource allocations within and between groups. The authors argue that group life has constituted an enduring feature of human evolutionary history and we should expect the human mind to contain psychological adaptations for dealing with political problems. Bøggild and Petersen note that previous research has focused on adaptations designed to produce moral evaluations of political outcomes: Is the allocation of resources fair? They argue that people are not only concerned about *outcomes*. They also readily produce moral evaluations of the political *processes* that shape these outcomes. In short, the authors argue that people have a sense of procedural fairness. Bøggild and Petersen argue that intuitions about procedural fairness evolved to deal with adaptive problems related to the delegation of leadership and, specifically, to identify and counteract exploitative leaders. The authors first introduce the concept of procedural fairness, review psychological theories, and make the case for why an evolutionary approach is necessary. Next, they discuss the evolved functions of procedural fairness and review previous research through the lens of evolutionary psychology. Finally, the authors discuss how environmental mismatches between ancestral and modern politics make procedural fairness considerations more potent in modern politics, often generating powerful sources of moral outrage.

Jeffrey Evans Stake makes a convincing case in chapter "Property Law Reflections of a Sense of Right and Wrong" that an evolutionary perspective on human morality may help us understand and critique the law. Stake examines three areas of American property law. In two of the three areas, title by first possession and title by adverse possession, the pieces of legal doctrine fit together when seen through an evolutionary lens. In the third area of law, compensation for eminent domain, Stake argues that the inconsistency between the legal doctrine and evolved psychology suggests why governmental takings of property raise public ire.

shackelf@oakland.edu

Stake closes the chapter by highlighting suggestions for what can be done to make the law less offensive to evolved sensibilities.

In the final chapter, Hoffman and colleagues address the strategic logic of moral intuitions from the perspective of game theory. The authors present an analysis of the Nash Equilibria of a series of simple games to reframe and explain many puzzling aspects of human morality. These include why we have a sense of rights, why we give to charity in odd and often inefficient ways, why we admire principled people, why we distinguish between transgressions of omission and commission, and several other themes in the literature on moral psychology. Hoffman and colleagues also enumerate several novel predictions and policy prescriptions. The arguments in this chapter suggest that supposedly a priori arguments for why our moral sentiments exist (descriptive ethics) and why they should be followed (prescriptive ethics) presented by philosophers such as Aristotle and Kant may be more about folk intuitions. According to Hoffman and colleagues, it should not surprise us that these philosophers generated explanations for our moral intuitions that strike us as post hoc, folk psychology. Hoffman and colleagues concede that although their own analysis might strike some readers as similarly folk-informed, these claims are based in rigorous modeling of the underlying dynamics using population level analyses of equilibria that are difficult to intuitively grasp. Hoffman and colleagues suggest that their own arguments question the notion that morality can be justified based on any a priori logic, at least one that does not account for individual incentives within one's lifetime. The authors also question the notion of "moral truths" other than if such truths are the moral intuitions that emerge from Nash Equilibria. These premises underlie much of moral philosophy and thus lead the authors to question the methodology commonly employed within the field, which often relies on psychological explanations that are little more than folk intuitions, neuropsychological description, or on unverifiable evolutionary processes with superfluous predictions. Hoffman and colleagues contend that their argument applies not only to esoteric philosophical debates but should also make us doubt the logic we give for our own morality, such as when we have political debates with our friends. And, it similarly draws into question the premise that moral progress is driven by reason.

The Evolution of Morality showcases the profound and wide-ranging intellectual value of an interdisciplinary approach to human psychology and behavior. Guided by Darwin's insights, the contributions to this wide-ranging volume provide a compelling case for an evolutionary analysis of morality.

Rochester, MI, USA

Todd K. Shackelford
Ranald D. Hansen

shackelf@oakland.edu

Contents

shackelf@oakland.edu

shackelf@oakland.edu

Part I
Psychological Design and Development

shackelf@oakland.edu

The Tripartite Theory of Machiavellian Morality: Judgment, Influence, and Conscience as Distinct Moral Adaptations

Kelly Asao and David M. Buss

Introduction

Debates surrounding morality dominate much of people's everyday conversations (Dunbar, 2004). Humans are fascinated by uncovering, discussing, and analyzing the moral decisions of themselves and others. People adamantly believe that their opinions on the subject of morality are objective and impartial. Can they be correct in these beliefs?

Consider a recent media explosion, the fallout after the U.S. government ordered drone strikes that killed Anwar al-Awlaki, a U.S. citizen and al-Qaida propagandist in Yemen. Many Americans were outraged that the target of the strike was a fellow citizen despite his affiliation with al-Qaida. They felt numerous morally charged emotions, such as disgust, fear, anger, and confusion. Additionally, people experienced a multitude of diverse cognitions. Some people believed that the strike was justified because of al-Awlaki's terrorist affiliations. Others were concerned that an American citizen, a member of their in-group, was killed without due process of law. This event brought attention to the legality of the targeted killings of the U.S. drone program. Despite the fact that drone bombings have killed several innocent Yemeni citizens, moral outrage was only sparked after the strike was ordered on an al-Qaida member who also happened to be an American citizen.

The authors thank Dan Conroy-Beam, Frank Mann, Todd Shackelford, and Joy Wyckoff for valuable suggestions on an earlier version of this chapter.

K. Asao (✉)
Department of Psychology, The University of Texas at Austin,
1 University Station, Austin, TX 78712-0187, USA
e-mail: kellyasao@gmail.com

D.M. Buss
The University of Texas at Austin, Austin, USA

© Springer International Publishing Switzerland 2016

T.K. Shackelford, R.D. Hansen (eds.), *The Evolution of Morality*,
Evolutionary Psychology, DOI 10.1007/978-3-319-19671-8_1

shackelf@oakland.edu

Group membership appears to dramatically alter people's feelings, beliefs, and behaviors when it comes to moral outrage. This demonstrates, to the surprise of few scientists, that moral intuitions are rarely dispassionate and objective. However, the key point from this example is that real-life situations perceived as moral violations evoke a wide array of emotional, cognitive, and behavioral phenomena. A deeper understanding of the mechanisms responsible for moral intuitions and actions could help elucidate such complicated social issues.

Some moral theorists have treated morality as a unitary phenomenon, treating distinct components of moral reasoning and behavior as singular in nature, and hence amenable to a unitary explanatory framework. Perhaps most frequently, scientists treat morality as more or less synonymous with altruism or cooperation (Gintis, Bowles, Boyd, & Fehr, 2003; Wilson, 2012). Studies demonstrating helping behavior in nonhuman primates and which provide the foundation for the argument that cooperation is the evolutionary root of morality focus solely on one aspect of moral behavior, namely, conscience (Tomasello & Vaish, 2013). Similarly, research has revealed that nonhuman animals and young children avoid unfair distributions of goods (Bloom, 2013; Brosnan & de Waal, 2003; Range, Horn, Viranyi, & Huber, 2009). However, this does not necessarily indicate that maintaining fairness or cooperation is the evolutionary root of morality.

In contrast, other researchers have taken the opposite approach, viewing morality as a multitude of distinct categories of adaptations or decision rules connected only loosely or terminologically (Graham et al., 2013; Haidt & Joseph, 2004; Kohlberg & Hersh, 1977; Shweder, Much, Mahapatra, & Park, 1997; Stich, 2006). Haidt and Joseph, for example, argue that the variety of moral rules can be traced to a few distinct moral foundations: care/harm, fairness/cheating, loyalty/betrayal, authority/subversion, and sanctity/degradation. For example, compassion and kindness evolved as a specific response to the suffering of one's offspring. In the modern environment, compassion is extended to many other contexts including strangers and baby seals.

Within the past decade, however, researchers in evolutionary psychology have made explicit an important implicit distinction between two major classes of morality: condemnation and conscience (DeScioli & Kurzban, 2009, 2013, Sperber & Baumard, 2012). Broadly, a distinction is made between moral decisions applied to others' behavior and those applied to one's own behavior. Furthermore, DeScioli and Kurzban were the first to highlight the importance of distinguishing between moral judgment and moralistic punishment.

In contrast to earlier approaches, we argue that the concept of morality comprises three distinct adaptations: moral judgment, moral influence, and moral conscience, each consisting of mechanisms designed to solve distinct adaptive problems. *Moral judgment*, we propose, is an adaptation subsuming a suite of evolved psychological mechanisms designed to determine whether a conspecific is *exploitative* or *prosocial*, that is, intentionally imposes either a net cost or a net benefit on one's inclusive fitness. The moral judgment mechanisms then store that information in the memory systems by identifying individuals as costly or beneficial relationship partners. This information is then used to avoid costly people, switch relationship partners, and track social relationships across time. For example, if a new acquaintance has a history of betraying friends, he probably will not make a good ally in times of crisis.

shackelf@oakland.edu

Moral influence, we propose, is a closely related, yet functionally distinct, adaptation. It consists of mechanisms that evolved to identify the most efficient and cost-effective way to alter another's future behavior to be less fitness cost-inflicting and more benefit-bestowing. Examples of moral influence include praise, bestowing rewards, rehabilitation, inflicting reputational damage, enlisting others for coordinated ostracism, and inflicting physical punishment.

The third moral component we propose, *moral conscience*, is a set of psychological mechanisms designed *to guide one's own behavior* toward others to avoid negative fitness consequences as a result of judgment and influence mechanisms in others. The emotion of anticipated guilt, for example, may function to deter the temptation to betray a friend to reap a short-term gain because of the long-term cost of a lost friendship. Taken together, these three psychological adaptations make up a tripartite theory of Machiavellian morality.

A key explanatory task is to identify classes of behaviors that humans moralize. The most obvious of these is intentionally exploitative behavior, such as stealing, which involves a gain for the perpetrator and a clear loss to the victim (Buss & Duntley, 2008). Behaviors intended to bestow a benefit on others, in contrast, would be judged as morally good. Examples include returning a lost wallet containing cash to its rightful owner or maintaining loyalty to a friend or mate when being disloyal might produce a temporary gain for the performer, but at a net long-term cost to the other.

The next explanatory step is to posit design features of the hypothesized adaptations, focusing on the inputs to the psychological mechanisms, the decision rules on which these mechanisms operate, and the behavioral or psychological output of the mechanisms. Certain inputs, decision rules, and outputs will be constant across the range of content domains and create the unity and consistency that researchers find in research on morality (DeScioli, Christner, & Kurzban, 2011). For example, the *intentions of the perpetrator* should influence moral judgments whether the violation was in the domain of physical harm, lying, or property damage (Keltikangas-Järvinen & Lindeman, 1997; Knobe, 2003; Nelson, 1980). However, other design features will be unique to solving a specific adaptive problem and add to the complexity and diversity of the moral content and behavior that we find across generations and cultures. Moral judgments about promiscuity should take certain inputs that are specific to sexuality, such as number of previous sexual partners, incidence of mate poaching, and incidence of infidelity. In contrast, moral judgments about disrespect would involve inputs such as existing hierarchical relationship and the size of the audience witnessing the disrespect.

The current chapter extends previous theorizing about morality by starting with an adaptive problem that the moral mechanisms evolved to solve and hypothesizing the design features necessary to solve it. The previous literature on morality lacks clear definitions of what constitutes morality due to the diversity of content areas (Krebs, 2011). Some have hypothesized distinct evolutionary origins of categories of moral content (e.g., purity, authority, harm), but doing so ignores the overwhelming similarities in decision rules across different content domains (DeScioli et al., 2011). Viewing the moral mechanisms as adaptations that function across content

shackelf@oakland.edu

areas by using content-specific mechanisms should help to identify the design features common across content areas and those that are unique to a specific adaptive problem. In turn, this form of analysis aids our understanding of both the universality and diversity of moral rules within and across cultures. Additionally, this approach may lead to more nuanced, novel, and testable predictions. By working through two examples, sexual infidelity and property theft, we will illustrate our tripartite framework of Machiavellian morality.

This approach refocuses research on morality to important content areas that have been neglected in the literature, such as moral rules governing the sexual domain. Additionally, this framework highlights the brighter side of moral influence in the form of gratitude, respect, praise, and rewards. Approaching the topic of morality from this evolutionary perspective may encourage new questions and new lines of research previously unexplored.

The Moral Mechanisms

The moral mechanisms comprise three distinct adaptations, each designed to solve broad classes of adaptive problems. Moral judgment, moral influence, and moral conscience are functionally distinct adaptations that operate across content areas. They encompass many subordinate mechanisms to solve the higher-order adaptive problems. These include categorizing a conspecific's behavior as intentionally cost-inflicting or benefit-bestowing, controlling or changing future behavior of conspecifics to minimize fitness costs and maximize benefits, and monitoring one's own behavior toward conspecifics to avoid condemnation, respectively.

Moral judgment comprises a set of evolved information-processing mechanisms that determines whether a moralized behavior has occurred, calculates a cost/benefit ratio of the behavior to self and to others, and stores that information for use in future social interactions. Specifically, moral judgment is designed for condemnation of cost-inflicting behavior and approbation of prosocial or benefit-bestowing behavior. Moral influence is activated after a moral judgment has been made, and avoidance of the cost-inflicting individual is unlikely. Moral influence mechanisms work to identify the most economical and efficient way of changing or controlling a conspecific's behavior for future interactions. That is, the mechanisms determine the best course of action given the specific circumstances of a moral violation to avoid future exploitative behavior from others and encourage prosocial behavior.

Moral conscience is a separate but related psychological adaptation designed to guide one's own benefit-bestowing or exploitative behavior. Importantly, moral judgment and moral influence are adaptations that focus on the decisions of conspecifics. Moral conscience, on the other hand, is concerned with behaviors of the self. Moral conscience therefore takes as input several internal regulatory variables (IRVs) and uses them to calculate the cost-benefit ratio of engaging in various forms of cost-inflicting or benefit-bestowing behavior (see Tooby, Cosmides, Sell, Lieberman, & Sznycer, 2008, for a fuller discussion of IRVs). Due to the crucial

shackelf@oakland.edu

distinctions between the moral mechanisms, the hypothesized inputs, decision rules, and outputs should diverge and converge in predictable ways. By analyzing each adaptation into its hypothesized design features, researchers can better predict and understand multidimensional phenomena such as morality.

Moral Judgment: Evolution and Design Features

In the ancestral environment of small-scale group living, individuals would have varied in the extent to which they engaged in exploitative strategies that intentionally imposed costs on others (Daly & Wilson, 1988). For example, some individuals would have been, on average, more selfish or more prone to exploitative aggression than others (Duntley & Buss, 2004). Risk of victimization by conspecifics was likely a frequent and recurrent problem in the small-scale societies (Chagnon, 1988). The ability to correctly identify and subsequently avoid these individuals as relationship partners would have afforded an evolutionary advantage. Importantly, moral judgment operates on a continuum ranging from morally evil to morally good. An underemphasized, but equally critical, function of moral judgment is to identify prosocial (i.e., intentionally benefit-bestowing) individuals and seek them out as future relationship partners. If certain individuals were consistently more likely to intentionally bestow benefits or suppress exploitative behaviors, the moral judgment mechanisms could motivate pursuit of these prosocial others as mates, friends, and cooperative partners.

If moral judgment is designed to evaluate a conspecific as either intentionally exploitative or prosocial, the inputs to the mechanisms should be factors of the individual and moralized behavior that shift the likelihood of costs and benefits to oneself, one's kin, and one's close social partners. The mechanism should first determine the likelihood that a given individual engaged in moralized behavior. The mechanisms should gather and review information through direct observation of the behavior, indirect information (e.g., gossip), and probabilistic cues to the moral violation. For example, in the absence of direct evidence, gossip about past moral behavior (e.g., moral reputation) can be used as an indicator of an exploitative or prosocial disposition (Dunbar, 1998).

Another input to the mechanism is whether the behaviors were intentional. If the morally good or bad behaviors were the result of incompetence, accident, or coercion, then these behaviors would not be indicative of the individual's intentions. Furthermore, if the individual's behavior was not intentional, then one could not reliably predict future exploitative or prosocial inclinations. Research has suggested that individuals who commit moral violations resulting from incompetence, chance, and strategy are categorized distinctly in the mind (Delton et al., 2012).

Finally, the key inputs include the overall probabilistic fitness costs and benefits to the self and to others with whom one's fitness is linked. The decision rules of the mechanisms should take the magnitude of the fitness costs (and benefits) imposed and the likelihood of receiving these fitness consequences by interacting with this individual.

shackelf@oakland.edu

The mechanisms then weigh these various inputs to determine how morally good or bad an individual is. The consequences of these weightings lead to a range of emotional, behavioral, and psychological outputs. Specifically, how morally good or evil a person is judged to be depends on the inputs: evidence of the behavior, intentions, and the likelihood of receiving fitness consequences as a result of the behavior. Direct evidence will be weighed more than indirect evidence; for example, catching someone inflicting harm will elicit a stronger reaction than hearing about it later from another person. The more likely a behavior was intentional, the more weight it will be given; for example, someone caught lying would be judged more harshly than someone omitting information because creating a lie is intentional, whereas omitting could be an artifact of forgetfulness. Finally, the magnitude of fitness consequences should shift moral judgment; for example, the act of murder imposes extreme fitness costs on an individual and should be more severely judged than less costly behaviors (Buss, 2006).

The outputs of the moral judgment mechanisms include activation of emotions (e.g., moral outrage, disgust, respect), thoughts (e.g., labeling the person who committed the behavior as morally good or evil), memory systems (e.g., storing the information about this person for future use), and behaviors (e.g., avoiding morally bad and pursuing morally good relationship partners). Although partner choice is a key output of moral judgment mechanisms, other behavioral and cognitive responses are notable. These responses include tracking social relationships based on history of exploitative or prosocial behaviors between third parties. Knowledge of who is an enemy or ally of whom would have been useful in close group living. Such knowledge would enable people to forge beneficial alliances, avoid costly ones, and predict side-taking in future conflicts. Thus, attention to the exploitative or prosocial interactions between third parties would have helped one navigate the complex and ever-changing network of alliances within one's social group. This could partially explain why people are fascinated by moral gossip surrounding unrelated third parties. Research has found that even infants as young as 3 months are able to use evidence of past moral behavior to avoid antisocial others (Hamlin, Wynn, & Bloom, 2010). Furthermore, infants are surprised when individuals choose to affiliate with a previously harmful character over a previously helpful one (Bloom, 2013).

The specific outputs of the moral judgment mechanisms depend on the circumstances surrounding the moralized behavior. A prosocial act, such as food sharing, will activate the emotions of respect and gratitude, will cause one to identify and remember this individual as a morally good person, and will motivate future interactions with this individual. These responses will interact in complex ways. The activation of emotions, in particular, may function to coordinate among the various cognitive, physiological, and behavioral responses of moral judgment mechanisms (Cosmides & Tooby, 2000). This could help explain why emotions are activated first, with other behavioral and psychological responses being activated later. Conscious moral deliberation is slow and unnecessary for the underlying than emotional reactions mechanisms to operate, giving the impression that rational rules matter much less (Haidt, 2001) (Fig. 1).

shackelf@oakland.edu

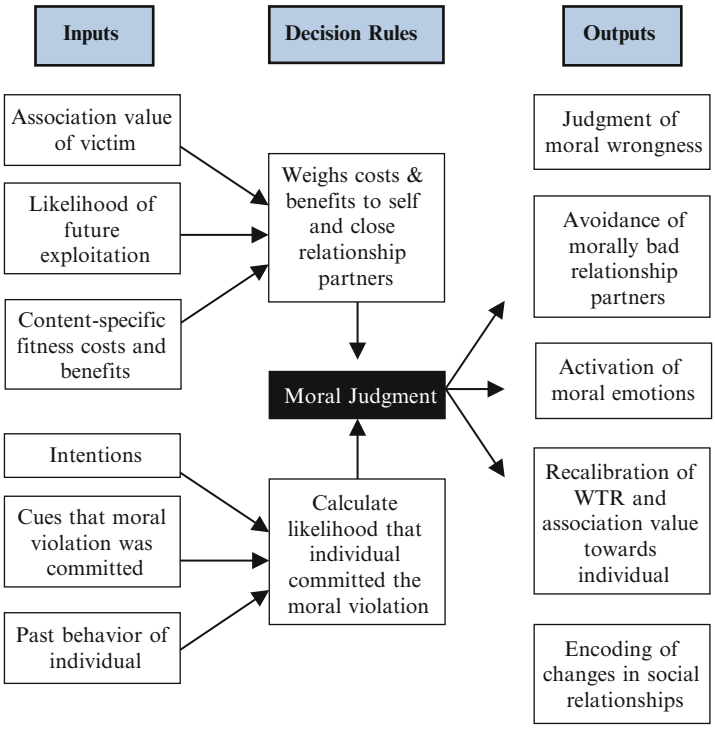

Fig. 1 Moral judgment

Moral Influence: Evolution and Design Features

Whereas moral judgment is concerned with tracking the likelihood of exploitation or prosociality by specific individuals and subsequent partner choice depending on those calculations, moral influence mechanisms function to *control and alter the behavior* of conspecifics to be less exploitative and more prosocial toward oneself and one's close relationship partners (Frank, 1988). Given the ability to freely choose with whom one interacts, avoidance is a relatively low cost means of solving the adaptive problem of exploitation by others (Baumard, André, & Sperber, 2013). However, in the ancestral environment of small group living, complete avoidance may have been unlikely or prohibitively costly. For example, the exploitative individual would have an extended kin and social alliance network with whom one would likely interact, occasionally putting one in social contact with the exploitative individual. Furthermore, the likelihood that a morally bad group member could impose costs on one's kin, mate, or friends would have increased the negative fitness consequences of allowing an exploitative individual's behavior to remain unchecked. When repeated interaction with a cost-inflicting individual was likely, then the

shackelf@oakland.edu

moral influence mechanisms for deciding how best to control and alter future behavior should become activated. Additionally, even if effective future avoidance of exploitative individuals were common in the ancestral environment, allowing oneself to be exploited without retribution could create a reputation as exploitable. This would encourage others within one's social group to inflict fitness costs without fear of retaliation.

Regarding third-party moral punishment, if an individual within one's social group were exploitative toward others within the group, the likelihood of that individual eventually inflicting costs on oneself or one's kin, mates, and friends would often be nonzero. Frequent exploitative behavior toward other group members could be indicative of a disposition that systematically undervalues other people's welfare relative to the self's welfare. This baseline low welfare trade-off ratio (WTR) could be correlated with likelihood of aggressive or exploitative acts that harm one's inclusive fitness. Thus, when the costs of exploitation by an individual outweigh the costs of punishment, then third-party punishment could theoretically evolve. Engaging in third-party punishment could act as an honest signal of one's ability and willingness to retaliate against moral violations without incurring the costs of being a direct victim of exploitation. Additionally, garnering a reputation as someone who punishes wrongdoers could lead to indirect fitness benefits including those associated with reciprocal altruism (Trivers, 1971; Wright, 1995) and strengthening of alliances with victims and their extended social networks. Thus, although punishment should be less costly and less frequent for the exploitation of unrelated others, people should still be willing to engage in third-party punishment under specific conditions: when the likelihood of future exploitation by a wrongdoer is high and reputational benefits are probable. Although evidence of third-party punishment is abundant, recent work attempting to remove audience effects and experimenter demand has shown little or no third-party punishment (Pedersen, Kurzban, & McCullough, 2013). The evolution of mechanisms for moral influence would have presented an effective solution to the adaptive problem of exploitation in situations of repeat interaction and reputation concerns, which likely characterized the close group living of the ancestral environment during which moral influence mechanisms evolved.

Moral influence decision rules should take as input factors specific to the individual. These person variables will moderate the likelihood and magnitude of punishment or rewards. Examples include how valuable a relationship partner the individual is (i.e., the association value of the individual; see Petersen, Sell, Tooby, & Cosmides, 2012) and how physically or socially formidable the individual is (Sell, Tooby, & Cosmides, 2009). Generally, the higher the association value, the less likely one is to seek punishment for a moral offense and the less severe that punishment will be. Family, friends, mates, and irreplaceable association partners all have high association value (Cosmides & Tooby, 1992). Heavy punishments inflicted upon these individuals will have negative fitness consequences for oneself. One will also be less likely to severely punish individuals who are physically or socially formidable for fear of retaliation.

Another key individual variable will be the individual welfare trade-off ratio (WTR), that is, the ratio between how much the individual values another person's welfare relative to his own welfare (Sell et al., 2009). Individuals with a low WTR

shackelf@oakland.edu

value another's welfare much less than their own, indicating the likelihood of their employing an exploitative strategy. The same individual could have a low WTR toward one person and a high WTR toward another. Perhaps even more important than a single measure of WTR would be a measure of an individual's WTR to oneself relative to other people. Since reproductive success is defined relative to conspecifics, if someone uniquely values one's welfare, especially if he does not particularly care about others in the social group, he could be a useful relationship partner. An individual who is not prosocial toward everyone in the group equally, but who instead is exclusively prosocial to oneself, would be preferred to an individual who is indiscriminately prosocial (Lukaszewski & Roney, 2010). Such a finding would provide evidence against group selection since people who have a low WTR toward other group members and only value oneself will be chosen as mates, friends, and allies over group-benefiting others. It also provides evidence against people's everyday application of utilitarian ethics since nearly everyone in the group would be better off if all individuals were indiscriminately prosocial. However, this is not the ethical code most people prefer in others.

Circumstantial factors should also be taken into account, such as the likelihood of future interaction with this person, likelihood of repeat moral violations, and the likelihood of deterring others from committing a similar violation against oneself in the future. If the likelihood of future interaction with the exploitative individual is very low, then there is no point in wasting time, energy, or resources to alter his future behavior (see Krasnow, Cosmides, Pedersen, and Tooby 2012 for a full discussion). However, if the likelihood of future interaction is high, then the magnitude of both rewards for good behavior and punishment for bad behavior should be large.

Other key factors include victim characteristics, such as association value, vulnerability, age, and sex. For example, harsher punishments will be imposed on those who harm or exploit more vulnerable victims since this behavior is indicative of a marked lack of empathy or a heightened tendency toward exploitative strategies. Content-specific variables relevant for solving a particular adaptive problem will also be taken as input.

In support of these hypothesized inputs to moral influence mechanisms, recent research suggests that people selectively impose costly punishment on those with whom they expect to cooperate in the future (Krasnow et al., 2012). Some research suggests that association value of the perpetrator influences whether punishment or rehabilitation is endorsed by third parties judging moral dilemmas (Lieberman & Linke, 2007; Petersen et al., 2012).

The output of moral influence decision rules should show efficiency in controlling moralized behavior with the least amount of effort and cost to self. Potential outputs of the moral influence mechanism are direct bestowal of rewards, public praise, rehabilitation, coordinated ostracism, direct punishment, indirect punishment (e.g., recruiting other individuals to inflict the punishment) or some combination of these strategies. Additional information-processing mechanisms could be responsible for determining the magnitude of consequences depending on the endogenous factors (e.g., WTR of individual) and the exogenous factors (e.g., relationship between self and other) mentioned above. The probability and magnitude of moral consequences will be proportional to the probability and magnitude of fitness costs

shackelf@oakland.edu

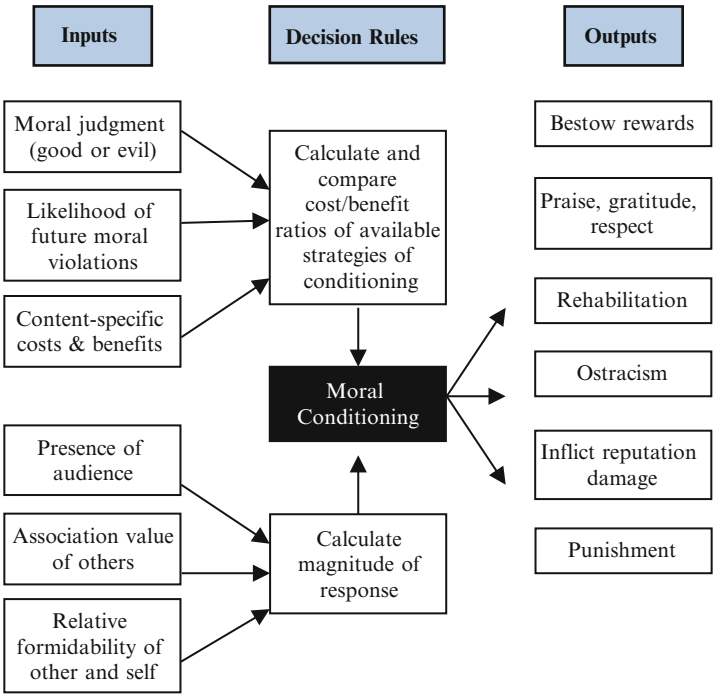

Fig. 2 Moral influence

and benefits to oneself or one's kin and close relationship partners. Behaviors that impose smaller fitness costs on oneself, such as intentional lying, will receive less punishment than relatively more costly violations, such as physical assault. Similarly, behaviors that provide larger benefits to oneself, such as saving one's life, will receive more reward than less beneficial behaviors, such as sharing food (Fig. 2).

Moral Conscience: Evolution and Design Features

Once the mechanisms for moral judgment and moral influence evolved, there would be selection pressure to regulate one's own behavior so as to strategically avoid reputation damage and other forms of punishment from others (DeScioli et al., 2011; Krebs, 2011) and to reap the rewards of building a good moral reputation (Sperber & Baumard, 2012). If moral conscience is a self-regulatory psychological adaptation, the features of the mechanism should be designed to economically and efficiently solve the adaptive problem of regulating one's behavior toward others to avoid negative moral judgment and influence. Specifically, moral conscience has two key functions: (1) to genuinely motivate one to resist costly exploitative

shackelf@oakland.edu

strategies in favor of benefit-bestowing behaviors that would have lead to greater long-term fitness benefits and (2) to strategically engage in self-beneficial behaviors while publicly garnering a good moral reputation. Having a reputation as a morally good individual would lead to a variety of long-term fitness benefits, including a large alliance network, increased mate value, direct rewards, and indirect reciprocity (Sperber & Baumard, 2012).

Proactively, moral conscience serves to anticipate the costs and benefits associated with exploitative actions and those associated with prosocial actions to guide behavior toward the least costly option. To do so, the information-processing mechanisms would take as inputs memory of past consequences of morally stigmatized or sanctioned behavior, the likelihood of punishment or reward (e.g., formidability of victim and kin), and the likelihood of reputation damage or enhancement (e.g., presence of an audience). Recent research examining the effects of reputation on morality has found evidence of audience effects (Haley & Fessler, 2005; Kurzban, Descioli, & O'Brien, 2007; Powell, Roberts, & Nettle, 2012; Rigdon, Ishii, Watabe, & Kitayama, 2009).

The inputs are then sent to the decision rules that weigh the costs and benefits of engaging in a given behavior compared to alternatives. The output of the system should take the form of a net cost-benefit ratio that should motivate subsequent behavior. Specifically, if the net cost-benefit ratio is higher than other available strategies, the individual will be motivated to avoid engaging in the behavior. If the net cost-benefit ratio is lower than alternative strategies, the individual will be motivated to engage in the behavior. Even a costly exploitative behavior (e.g., theft) could be the best available strategy if the alternatives are relatively more costly (e.g., starvation). Importantly, moral conscience is designed to forego short-term gains via exploitation in favor of the long-term benefits of prosociality. Therefore, if one does not expect to survive long enough to reap the long-term benefits, then the best strategy would be to engage in exploitation and receive immediate benefits. This could help explain the correlation between future discounting and various indices of criminal behavior. Researchers are increasingly recognizing the importance of considering criminal behavior as evolved strategies for acquiring resources, status, and mates under certain circumstances (Buss, 2012; Duntley & Buss, 2010; Durrant & Ward, 2012).

Retroactively, moral conscience performs two key functions. One mechanism is designed to do strategic damage control after a moral violation has been performed. A second is designed to encode the negative or positive consequences of moralized behavior to prevent oneself from engaging in costly moral violations in the future. The strategic damage control mechanisms should take as input cues to the likelihood of punishment or reputation damage (e.g., presence or absence of an audience, formidability), the magnitude of fitness costs inflicted on the victim, and characteristics of the victim that could influence others' moral judgments (e.g., vulnerability, age, and sex of victim).

These mechanisms should then weigh the different inputs to determine the best course of action to minimize the costs of having already engaged in a morally impermissible behavior. A key function of moral conscience is to engage in strategic

shackelf@oakland.edu

multi-person games, in which the best course of action is dependent on the behavior of others. For example, if many other people know about a moral transgression, the system should motivate expressions of guilt and remorse, apologizing to the victim and the victim's kin and making reparations depending on the self-assessed severity of the violation. If few people are aware of the transgression, the best strategy may be to deny the transgression altogether, make plausible excuses for the immoral behavior (e.g., blame the victim or circumstances surrounding the transgression), hide evidence of wrongdoing, and recruit kin, mates, and allies to defend oneself against retaliation. Individual difference variables such as formidability, and contexts such as need, should lead to different strategies of strategic damage control.

The encoding mechanisms are responsible for gathering information about the fitness consequences of engaging in a moralized behavior. Benefit-bestowing behavior could lead to a variety of positive consequences, including increased mate value, reputation enhancement, strengthening of existing alliances, and formation of new alliances. Conversely, morally impermissible or fitness cost-inflicting behavior could lead to numerous negative fitness consequences, including physical, economical, or emotional retaliation by the victim or the victim's kin, reputation damage, ostracism, loss of close relationships, and creation of enemies. The retroactive mechanisms responsible for encoding moral consequences should search for information about the positive and negative consequences of engaging in moralized behavior and store that information in memory systems.

Since moral conscience is a self-control mechanism, there is no reason why the mechanism should make the Machiavellian decision rules available to consciousness. All that is needed to motivate individuals to suppress their immediate immoral urges to avoid condemnation and seek rewards is positive feedback when committing morally good acts and negative feedback when committing morally bad acts. In much the same way that humans derive pleasure from food and sex without an explicit understanding of the link between food and differential survival or sex and differential reproductive success, moral behavior could lead to positive emotions and cognitions in the absence of conscious awareness of the decision rules underlying moral conscience. The proximate phenomenology of engaging in morally good behavior would include positive emotions and cognitions to motivate such prosocial acts, although the function is selfish in the ultimate sense (Krebs, 2005) (Fig. 3).

Sexual Infidelity

The sexual domain is one of the most highly moralized content domains (Buss & Asao, 2013). People make moral judgments about which sexual acts are permissible or impermissible, who can perform those acts with whom, when and where those acts can take place, and why others decide to engage in sex acts (Shweder et al., 1997). This is because the consequences of other people's sexual behavior historically had dramatic consequences for one's own reproductive success (Symons, 1979).

shackelf@oakland.edu

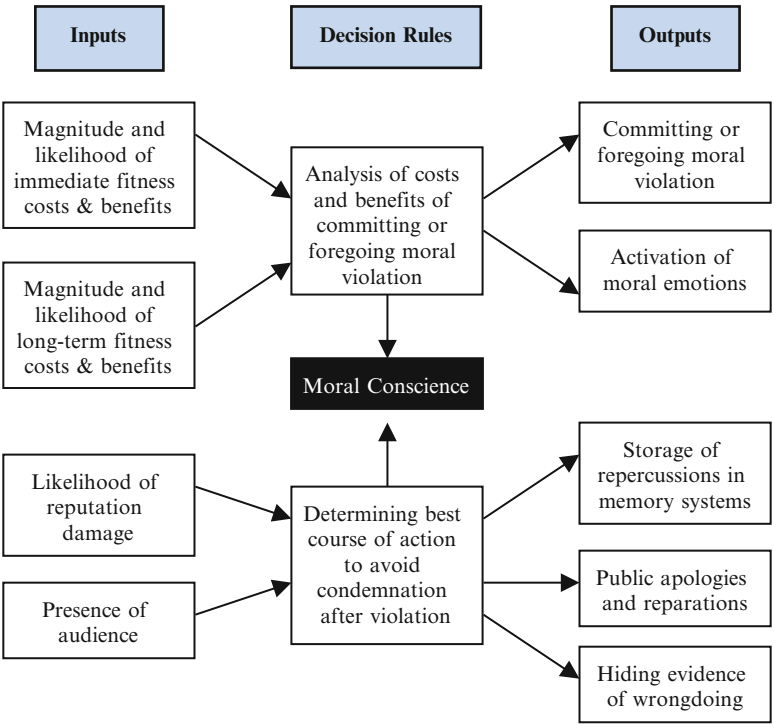

Fig. 3 Moral conscience

One commonly moralized sexual behavior is infidelity, which provides a useful illustration of the heuristic value of the proposed tripartite framework of Machiavellian morality. Sexual infidelity committed by one's partner is a particularly costly behavior for both men and women. Sexual infidelity could lead to a variety of costs to one's reproductive fitness, from termination of the existing relationship and all economic, social, and sexual benefits therein to sexually transmitted disease (Buss, 2000). Men face an additional cost of cuckoldry in the case of insemination by a rival male, whereas women face the diversion of partner's investment to a rival female (Symons, 1979). Sexual infidelity by one's partner is, therefore, an extremely cost-inflicting behavior, and changing the future likelihood of this behavior constitutes a recurrent adaptive problem that would have required a solution.

Moral judgments are often automatic, spontaneous, and associated with strong emotions (Haidt, 2001). However, the mechanisms responsible for the activation of these physiological, psychological, and behavioral responses need to be examined in greater detail. The inputs to the moral judgment mechanism would include (1) information that the infidelity occurred (e.g., proof of infidelity directly, indirect knowledge via gossip, or probabilistic cues to infidelity) (Shackelford & Buss, 1997), (2) potential costs associated with sexual infidelity (e.g., termination of a beneficial relationship, loss of sexual access to mate, loss of resources provided

shackelf@oakland.edu

by mate, reputation damage, sexually transmitted diseases, and cuckoldry), and (3) potential benefits associated with infidelity (e.g., termination of a detrimental relationship).

The decision rules would then weigh the evidence to determine the likelihood that the conspecific performed the moralized behavior and generate a net cost-benefit ratio of the behavior. For infidelity, the decision rules would calculate the likelihood that one's partner engaged in sexual infidelity and would calculate the large net cost associated with such infidelity. The output of the moral judgment mechanism could be emotional (e.g., moral outrage, disgust, or feeling betrayed), rational (e.g., identifying the cheater as a bad relationship partner), involve the memory system (e.g., encoding and storing the information about the infidelity, such as the identity of the mate poacher and the cues associated with discovery of the infidelity), or some combination of these outputs.

Moral influence mechanisms would then be activated to determine the best course of action for avoiding such large fitness costs in the future. One strategy would be to terminate the relationship, ensuring no possibility of sexual infidelity by this partner in the future. Another option would be to remain in the relationship but inflict some form of direct punishment on one's partner in the form of imposing physical or emotional costs or withdrawing benefits, such as withholding sexual access or resources (Buss & Duntley, 2011). This punishment would make sexual infidelity a prohibitively costly behavior, effectively preventing future instances of partner infidelity. These behavioral responses only represent two choices in a large array of options including indirect punishment (e.g., enlisting family members to inflict costs on an unfaithful partner), reputation damage (e.g., gossip to lower a partner's mate value and the likelihood of attracting future affair partners), and retaliatory affairs (Buss & Shackelford, 1997).

The influence strategy that an individual pursues depends on various inputs to the information-processing mechanisms. Some of these inputs will be unique to the adaptive problem of sexual infidelity, such as the relative mate value of self and the unfaithful partner and the likelihood of finding another sexual partner of equal or greater mate value in the future. Other inputs will be general to moral influence mechanisms across adaptive problem and content domain, such as the value and irreplaceability of the individual (i.e., association value), relative physical and social formidability of self and others, formidability of one's kin and alliance network, and likelihood of future moral violations by the individual. Empirically, one of the key motivators of intimate partner violence is suspicion or discovery of a sexual affair (Buss & Duntley, 2011; Daly & Wilson, 1988).

Unlike moral influence, moral conscience is independent of moral judgment. The distinction between moral judgment and moral conscience could help to explain moral hypocrisy, in general, and sexual double standards, in particular. Since moral conscience is designed to control and alter one's own behavior in response to past or probabilistic negative consequences of immoral actions, the inputs to the mechanism are orthogonal to the inputs to moral judgment mechanisms. In the case of one's own sexual infidelity, if either the benefits of sexual infidelity outweigh the costs or the chances of discovery are sufficiently low, then the mechanisms underlying moral

shackelf@oakland.edu

conscience could motivate an individual to engage in a clandestine affair. The conscience mechanism would take as input the likelihood that the sexual affair will be discovered by one's partner (e.g., suspiciousness of current partner, ease of coordinating or concealing the affair), the costs associated with discovery (e.g., loss of current relationship partner, reputation damage, risk of emotional or physical retaliation), and the benefits associated with engaging in sexual infidelity (e.g., mate switching to a higher-quality mate).

The decision rules would then weigh the positive and negative consequences of infidelity by their probabilities and generate a cost-benefit ratio. If the ratio is higher than alternative courses of action, then the mechanism could activate systems to avoid engaging in infidelity. If the ratio is lower than alternatives, then the mechanism could activate systems to motivate the affair, but only if one can conceal the affair with minimal costs to self, including avoiding the large cost of losing one's current partner.

Importantly, there are some circumstances in which the moral conscience mechanisms could motivate one's own sexual affair, while one's moral judgment mechanisms could judge another's sexual affair as morally wrong and worthy of punishment. This hypocrisy is possible due to the separate nature of the moral mechanisms. Specifically, the mechanisms that guide one's own behavior (i.e., moral conscience) are not necessarily those involved in judging other people's behavior (i.e., moral judgment). To the extent that the sexes differ in the costs and benefits of engaging in a given moralized behavior, the moral judgment and moral conscience mechanisms in men and women may provide diverging judgments of moral wrongness and deserved punishment of behaviors such as own and partner's infidelities (Buss, 2009).

In short, sexual infidelity illustrates the heuristic value of distinguishing among moral judgment, influence, and conscience. Additionally, highlighting the importance of sexual morality opens new lines of research. The next section applies our Machiavellian tripartite scheme to the domain of theft. Although theft is as ubiquitous as sexual infidelity, the two behaviors have little else in common. We chose such a divergent content area to examine which design features of the moral mechanisms operate across content areas and which are unique to the sexual domain.

Property Theft

Another problem that has been recurrently faced by humans over deep evolutionary time is the taking of one's personal belongings by conspecifics (Buss & Duntley, 2008; Duntley & Shackelford, 2008). Historically, personal property has taken numerous forms, from physical resources (e.g., animals, food, and modern currency) to intellectual property (e.g., ideas and written works). The negative ramifications of personal property theft are threefold: the actual loss of access to a valued personal item, the relative fitness costs associated with a rival gaining access to the valued item, and increased perceived exploitability due to the theft. Since these costs are

shackelf@oakland.edu

large, theft constitutes an adaptive problem of controlling or changing a conspecific's behavior. Theft, therefore, falls within the purview of the moral mechanisms.

To determine whether property theft actually occurred, moral judgment mechanisms will take as input certain information that is specific to the adaptive problem of theft prevention, such as local norms concerning ownership of property (e.g., documents or social rules that define what constitutes personal property and ownership in the culture) and information about previous and current possession of the property (e.g., who was first in possession, was the property lost or out of possession before it was taken by another). Other inputs will be the same across all contexts, such as whether the conspecific stole the property intentionally or as a result of accident, incompetence, or coercion. Similar to the inputs for sexual infidelity, moral judgment mechanisms will also take in information about the costs and benefits associated with having property stolen (e.g., how valuable and irreplaceable the property was and the probability of increased future exploitation by others) to calculate how morally wrong the violation was based on the costs and benefits to the burglarized individual. The higher the probability that a theft actually occurred and the more valuable or irreplaceable the item stolen is, the more morally wrong the thief will be judged. Other output of the moral judgment system includes activation of the moral emotions, and memory and avoidance of known thieves.

Once a moral judgment has been made, the moral influence mechanisms are activated such that the output of moral judgment is one input into the decision rules governing the type and magnitude of influence used to alter the behavior of others. In the case of an exploitative behavior, such as property theft, the influence would occur in the form of moralistic punishment. General inputs into the moral influence mechanism will include those mentioned previously for sexual infidelity (e.g., association value of the individual, relative formidability of self versus individual, and likelihood of future interactions with the individual). There will also be theft-specific inputs, such as whether the stolen property can be easily returned without damage or devaluation. The influence mechanism will then conduct a cost-benefit analysis for each potential punishment strategy to find the most efficient and economical means of deterring future theft either from the individual or from other conspecifics. If the stolen item is minimally valuable or one that can be easily returned without devaluation, the least costly strategy would be to simply ask the thief to return the property. However, if the piece of property were highly valuable or irreplaceable, if the item cannot be returned without devaluation, or if the theft increases one's reputation for exploitability (see Buss & Duntley, 2008), then the punishment should be more severe. Costly forms of punishment, such as direct physical, financial, or emotional punishment, could be worthwhile strategies if the long-term benefits of deterring future property theft by others outweigh the immediate costs of engaging in punishment.

The moral conscience mechanisms function to proactively deter theft, a costly strategy of resource acquisition, unless either the chances of detection are low or the benefits outweigh the costs weighed by their respective probabilities. For example, if an individual is on the brink of starvation and has no other feasible means of securing food resources, then food theft could be a useful strategy. In this case, the

shackelf@oakland.edu

benefits of immediate caloric intake necessary to sustain life outweigh the costs of potential detection and punishment. If the individual engages in property theft, the strategic damage control mechanisms should activate behaviors that decrease the likelihood of detection and punishment, such as leaving the site of the theft, avoiding the victim of theft, and denying accusations of theft unless the likelihood of detection is overwhelmingly high. If discovery of theft is inevitable, a different set of emotional, psychological, and behavioral mechanisms should be activated, including genuine feelings and confessions of guilt and remorse to appease the victim and victim's kin, verbal attempts to exculpate oneself from blame, and the return of or reparations for stolen property. The negative consequences associated with committing property theft should be encoded into memory systems to deter use of costly resource acquisition strategies in the future. The importance of circumstances in determining the output of the moral conscience mechanisms highlights the role of context in certain content domains. Some moralized behaviors are almost always condemnable when directed toward one's in-group members, for example, rape or torture. Other content areas, such as theft, may elicit moral decisions concerning wrongness, deserved punishment, and permissibility that are more context-dependent. Research currently being conducted suggests that violations involving theft may be more amenable to shifts in judgments of moral wrongness and deserved punishment than violations involving rape (Asao & Buss, 2014).

The examples of sexual infidelity and property theft illustrate the usefulness of the tripartite framework of Machiavellian morality. While some aspects of the moral mechanisms operate across content areas, other design features are content specific. By treating moral judgment, influence, and conscience as three related but separate suites of mechanisms, more nuanced predictions about the design features of those evolved psychological adaptations can be made. Specifically, the inputs, decision rules, and outputs can be examined in greater detail. This level of specificity is particularly helpful when dealing with a multidimensional and complex topic such as morality. This approach could help better define the borders between the different moral mechanisms and avoid the ambiguity, confusion, and polarization that have characterized research on morality in the past.

Ambiguity in Morality Research

Everyday discussions of morality are as ambiguous as they are ubiquitous. People show intense interest in the morality of other people. Gossip concerning the moral behaviors of others is commonplace. The media is dominated by information about the moral decisions of others. This is apparent from headline news stories such as the U.S. drone bombings in Yemen to small town reports of local heroes. Despite the widespread interest in morality, there are currently no agreed-upon definitions of what constitutes morality or the moral domain, nor is there even a consensus on whether such definitions would prove helpful when discussing morality.

shackelf@oakland.edu

Part of the conceptual confusion lies in the fact that the term "morality" is a broad term applied to at least three distinct adaptations designed to solve related but distinct adaptive problems. Some researchers have treated morality as synonymous with altruism (Wilson, 2012). However, altruistic behavior is only one subset of behaviors driven by moral conscience mechanisms. Although altruism is a theoretically important class of behaviors that can be partially explained by reference to moral conscience mechanisms, it is not the only behavioral output of those mechanisms. The function of moral conscience is to motivate people to strategically forego immediate fitness gains from moral wrongdoing in favor of delayed gains associated with maintaining a positive moral reputation and avoiding punishment. If correct, much of the function of moral conscience is to suppress exploitative tendencies out of concern for reputation and punishment, instead of promoting prosociality or group cohesion. A major component of moral conscience mechanisms will be to determine the circumstances in which one can effectively engage in exploitation without fear of negative fitness consequences. Thus, attempts to hide moral wrongdoing are important outputs of the moral conscience mechanisms. Shedding a light on the darker aspects of people's moral conscience mechanisms may change the way scientists think about moral conscience.

Another complication with studying morality is that "morality" is an umbrella term that covers a wide spectrum of content domains. Moralized content areas include sexual activity, food taboos, physical harm, theft, property rights, cultural norms, and fairness. This diversity is further confounded because certain behaviors are moralized to solve a specific evolutionarily relevant problem and are nearly universally condemned (e.g., cheating in social exchanges, theft, and murder); however, moralization of other behaviors lacks grounding in solving an adaptive problem and is highly variable across culture and time (e.g., homosexuality and cultural norms). The term "morality" is applied equally to these two quite different categories of behavior. The approach advocated here (i.e., starting with conspecifics' behaviors that would have constituted an adaptive problem for the individual that necessitated solving) leads to investigations of the former category of moralized behavior. However, once the moral mechanisms evolved to solve adaptive problems associated with avoiding and controlling cost-inflicting people, those mechanisms could have been used to moralize behaviors that were undesirable for reasons other than probabilistic fitness costs to promote one's religious, social, or political ideology (Rozin, Markwith, & Stoess, 1997). Additionally, public moralization of behaviors could advertise one's own moral goodness or to coordinate third-party side-taking during conflicts (DeScioli & Kurzban, 2013; Petersen, 2013).

Despite the diversity of moral content, there is evidence that people use similar underlying principles when making moral evaluations across contexts. The omission effect, in which actions are consistently judged more morally wrong than equivalent omissions (e.g., a lie is worse than an omission of the truth), is one such principle that operates across content areas (Cushman, Young & Hauser, 2006). Another is the doctrine of double effect which states that moral violations done to achieve another goal are less morally wrong than the same violation committed as a means to an end (Descioli, Asao, & Kurzban, 2012). Additionally, there is considerable agreement in

shackelf@oakland.edu

the relative ranking of moral violations across individuals and cultures, indicating a basic level of agreement in criteria used to assign severity of moral violations (e.g., Mikhail, 2007; Robinson, Kurzban, & Jones, 2007). Given the consistency in moral information processing, it is unlikely that moralization of different content domains evolved independently of one another. These surprising findings could be explained if each moral mechanism (moral judgment, punishment, and conscience) constitutes a unified set of adaptations that operate across distinct content areas by taking both content-general and content-specific factors as input, performing cost-benefit analyses, and coordinating behavioral, physiological, and psychological outputs to evaluate and control the behavior of a conspecific or oneself.

Discussion and Future Directions

The current framework builds on the extant body of research by integrating ideas from different models of morality and expanding these models to generate novel hypotheses about unexplored content areas. Baumard, André, and Sperber (2013) have emphasized the important role that partner choice plays in the evolution of fairness. We argue that evaluating potential relationship partners is the ultimate goal of moral judgment. We extend the idea of partner choice and switching into previously unexplored domains of morality, such as sexuality and property rights. Furthermore, we highlight the equal importance of moralistic rewards and punishment in shaping the evolution of moral norms. Partner choice was likely limited in the close group living that characterized ancestral hunter gatherers. Repeated exposure to most members of the group, lack of new available partners, and coercion could have made avoidance of exploitative individuals difficult. Therefore, the mechanisms for enacting moral rewards and punishments evolved to alter the behavior of others within one's social circle.

DeScioli and Kurzban (2009) raised crucial questions about the existence of third-party moral judgment and moralistic punishment. Additionally, they were the first to posit that moral conscience mechanisms are designed to avoid condemnation from others. Building on this work, the current framework offers a related, but different, account of the evolution of moral judgment and moralistic punishment. Furthermore, the model aspires to encompass the full spectrum of moral influence behaviors. In addition to moralistic punishment, the model sheds light on the equally puzzling phenomena of why humans praise, reward, and seek out as relationship partners those who help others. The media is filled with stories of everyday heroes, and people experience more positive emotions than negative emotions in everyday life (Algoe & Haidt, 2009). However, these positive aspects of morality and moral emotions are often overlooked in psychology. By emphasizing these positive aspects of moral influence, we hope to provide a more complete picture of morality.

The tripartite scheme of Machiavellian morality also recognizes the importance of perspectival shifts in moral decision-making (Duntley & Buss, 2004). There will be predictable differences in moral judgments depending on the

shackelf@oakland.edu

perspective of the evaluator. For instance, the same moral violation will elicit harsh negative judgments and punishment from the victim's close relationship partners, but may receive less harsh judgments from unrelated strangers or enemies. Furthermore, since moral judgment mechanisms are hypothesized to operate independently from moral conscience mechanisms, we expect that a moral violation committed by oneself will not be judged the same as a similar moral violation committed by another. Many moral double standards are accountable by the Machiavellian approach to morality. Thus, not all victims or perpetrators are created equal. A deeper understanding of the moral information-processing mechanisms will help to uncover potentially harmful biases in people's moral intuitions and behaviors. This information could then be used to guide public policy concerning morality to create more impartial and egalitarian policies.

Avoiding the Naturalistic Fallacy in Morality Research

The naturalistic fallacy is the logical error of assuming that because something is "natural" it must be morally right. We emphasize here that uncovering the evolutionary origins of people's moral intuitions and biases **does not** equate to condoning or endorsing those intuitions and biases. Importantly, this evolutionary psychological approach to morality, while theoretically useful, does not attempt to uncover what is objectively morally good or bad. The approach outlined here is concerned with understanding information-processing mechanisms housed in the brain and is agnostic about objective moral truth. Since our moral intuitions are shaped by evolution by natural selection, they were ultimately designed to increase reproductive success in ancestral environments, not to search for moral truths. The extent to which people's moral intuitions coincide with what any given philosophical perspective deems morally right is incidental.

Conclusion

Outside of the laboratory setting, questions of morality are rarely as simple and straightforward as the dilemmas commonly used to uncover people's moral reasoning. People infrequently conform neatly to a specific philosophical moral doctrine, such as utilitarianism, and instead use their biased, imperfect, and complex intuitions to guide their moral judgments and behaviors. Emotions are activated to coordinate among the various, competing subordinate mechanisms, while cost-benefit analyses are performed to determine the best courses of action. Prior relationships, reputation concerns, and formidability influence moral decision-making, despite people's desire for a justice system that is blind to those factors.

Further research using this tripartite evolutionary framework can help to recognize the biases in moral thinking. For example, people may be less impartial or egalitarian than previously believed. This approach aims to explore the role of social

shackelf@oakland.edu

relationships, formidability, physical attraction, and welfare trade-off ratios on moral decision-making. This knowledge can then be used to inform public policy, especially when our intuitions and behaviors do not coincide with what we deem objectively morally right.

Finally, the Machiavellian tripartite schema outlined above can help to organize the important advances that evolutionary psychologists have made in understanding the complexities of human morality. Moral judgment, influence, and conscience are adaptations designed to solve adaptive problems that span content areas. Starting from the adaptive problems that other people's behaviors pose, researchers can tease apart the content-specific and global design features of the moral mechanisms. This framework attempts to appreciate both the consistency and diversity within the moral domain.

References

Algoe, S. B., & Haidt, J. (2009). Witnessing excellence in action: The "other-praising" emotions of elevation, gratitude, and admiration. *The Journal of Positive Psychology, 4*(2), 105–127.

Asao, K., & Buss, D. M. (2014, June). *When justice isn't blind.* Poster presentation at CEU Summer University program on Morality: Evolutionary Origins and Cognitive Mechanisms, Budapest, Hungary.

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

Bloom, P. (2013). *Just babies: The origins of good and evil*. New York: Random House LLC.

Brosnan, S. F., & de Waal, F. D. M. (2003). Monkeys reject unequal pay. *Nature, 425*, 297–299.

Buss, D. M. (2009). *Sexual double standards*. Talk presented at annual meeting of the Human Behavior and Evolution Society.

Buss, D. M. (2000). *The dangerous passion: Why jealousy is as necessary as love and sex*. New York: Simon and Schuster.

Buss, D. M. (2006). *The murderer next door: Why the mind is designed to kill*. New York: Penguin Group.

Buss, D. M. (2012). The evolutionary psychology of crime. *Journal of Theoretical and Philosophical Criminology, 1*(1), 90–98.

Buss, D. M., & Asao, K. (2013, March). *The evolution of sexual morality*. Talk presented at Evolution of Morality conference at Oakland University, Rochester, MI.

Buss, D. M., & Duntley, J. D. (2008). Adaptations for exploitation. *Group Dynamics: Theory, Research, and Practice, 12*(1), 53.

Buss, D. M., & Duntley, J. D. (2011). The evolution of intimate partner violence. *Aggression and Violent Behavior, 16*(5), 411–419.

Buss, D. M., & Shackelford, T. K. (1997). From vigilance to violence: Mate retention tactics in married couples. *Journal of Personality and Social Psychology, 72*(2), 346.

Chagnon, N. (1988). Life histories, blood revenge, and warfare in a tribal population. *Science, 239*(4843), 985–992.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind*. New York: Oxford University Press.

Cosmides, L., & Tooby, J. (2000). Evolutionary psychology and the emotions. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 91–115). New York: Guilford.

Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment: Testing three principles of harm. *Psychological Science, 17*(12), 1082–1089.

Daly, M., & Wilson, M. (1988). *Homicide*. Hawthorne, New York: Aldine.

shackelf@oakland.edu

Delton, A., Cosmides, L., Guermo, M., Robertson, T. E., & Tooby, J. (2012). The psychosemantics of free riding: dissecting the Architecture of a Moral Concept. *Journal of Personality and Social Psychology, 102*(6), 1252–1270.

Descioli, P., Asao, K., & Kurzban, R. (2012). Omissions and byproducts across moral domains. *PLoS One, 7*(10), e46963.

DeScioli, P., Christner, J., & Kurzban, R. (2011). The omission strategy. *Psychological Science, 22*(4), 442–446.

DeScioli, P., & Kurzban, R. (2009). Mysteries of morality. *Cognition, 112*(2), 281–299.

DeScioli, P., & Kurzban, R. (2013). A solution to the mysteries of morality. *Psychological Bulletin, 139*(2), 477–496.

Dunbar, R. I. M. (1998). *Grooming, gossip, and the evolution of language*. Cambridge, MA: Harvard University Press.

Dunbar, R. I. M. (2004). Gossip in evolutionary perspective. *Review of General Psychology, 8*(2), 100–110.

Duntley, J., & Buss, D. M. (2004). The evolution of evil. In A. G. Miller (Eds.), *The social psychology of good and evil* (pp. 102–123). New York: Guilford.

Duntley, J. D., & Buss, D. M. (2010). The evolution of stalking. *Sex Roles*. doi:10.1007/s11199-010-9832-0.

Duntley, J. D., & Shackelford, T. K. (2008). Darwinian foundations of crime and law. *Aggression and Violent Behavior, 13*(5), 373–382.

Durrant, R., & Ward, T. (2012). The role of evolutionary explanations in criminology. *Journal of Theoretical and Philosophical Criminology, 4*(1), 1–37.

Frank, R. H. (1988). *Passions within reason: The strategic role of the emotions*. New York: WW Norton & Co.

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior, 24*(3), 153–172.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology, 47*, 55–130.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*(4), 814.

Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus, 133*(4), 55–66.

Haley, K. J., & Fessler, D. M. T. (2005). Nobody's watching? *Evolution and Human Behavior, 26*(3), 245–256.

Hamlin, J. K., Wynn, K., & Bloom, P. (2010). Three-month-olds show a negativity bias in their social evaluations. *Developmental Science, 13*(6), 923–929.

Keltikangas-Järvinen, L., & Lindeman, M. (1997). Evaluation of theft, lying, and fighting in adolescence. *Journal of Youth and Adolescence, 26*(4), 467–483.

Knobe, J. (2003). Intentional action in folk psychology: An experimental investigation. *Philosophical Psychology, 16*(2), 309–324.

Kohlberg, L., & Hersh, R. H. (1977). Moral development: A review of the theory. *Theory Into Practice, 16*(2), 53–59.

Krasnow, M., Cosmides, L., Pedersen, E. J., & Tooby, J. (2012). What are punishment and reputation for? *PLoS One, 7*(9), e45662.

Krebs, D. (2005). The evolution of morality. In D. M. Buss (Ed.), *The handbook of evolutionary psychology*. New Jersey: John Wiley & Sons.

Krebs, D. (2011). *The origins of morality: An evolutionary account*. Retrieved from http://www.eblib.com.

Kurzban, R., Descioli, P., & O'Brien, E. (2007). Audience effects on moralistic punishment. *Evolution and Human Behavior, 28*(2), 75–84.

Lieberman, D., & Linke, L. (2007). The effect of social category on third party punishment. *Evolutionary Psychology, 5*(2), 289–305.

shackelf@oakland.edu

Lukaszewski, A. W., & Roney, J. R. (2010). Kind toward whom? Mate preferences for personality traits are target specific. *Evolution and Human Behavior, 31*(1), 29–38.
Mikhail, J. (2007). Universal moral grammar: Theory, evidence and the future. *Trends in Cognitive Sciences, 11*(4), 143–152.
Nelson, S. A. (1980). Factors influencing young children's use of motives and outcomes as moral criteria. *Child Development, 51*(3), 823.
Pedersen, E. J., Kurzban, R., & McCullough, M. E. (2013). Do humans really punish altruistically? A closer look. *Proceedings of the Royal Society B: Biological Sciences, 280*(1758), 20122723.
Petersen, M. B. (2013). Moralization as protection against exploitation: Do individuals without allies moralize more? *Evolution and Human Behavior, 34*, 78–85.
Petersen, M. B., Sell, A., Tooby, J., & Cosmides, L. (2012). To punish or repair? Evolutionary psychology and lay intuitions about modern criminal justice. *Evolution and Human Behavior, 33*(6), 682–695.
Powell, K. L., Roberts, G., & Nettle, D. (2012). Eye images increase charitable donations: Evidence from an opportunistic field experiment in a supermarket. *Ethology, 118*(11), 1096–1101.
Range, F., Horn, L., Viranyi, Z., & Huber, L. (2009). The absence of reward induces inequity aversion in dogs. *Proceedings of the National Academy of Sciences of the United States of America, 106*(1), 340–345.
Rigdon, M., Ishii, K., Watabe, M., & Kitayama, S. (2009). Minimal social cues in the dictator game. *Journal of Economic Psychology, 30*(3), 358–367.
Robinson, P., Kurzban, R., & Jones, O. D. (2007). The origins of shared intuitions of justice. *Vanderbilt Law Review, 60*(6), 1633–1688.
Rozin, P., Markwith, M., & Stoess, C. (1997). Moralization and becoming a vegetarian: The transformation of preferences into values and the recruitment of disgust. *Psychological Science, 8*(2), 67–74.
Sell, A., Tooby, J., & Cosmides, L. (2009). Formidability and the logic of human anger. *Proceedings of the National Academy of Sciences of the United States of America, 106*(35), 15073–15078.
Shackelford, T. K., & Buss, D. M. (1997). Cues to infidelity. *Personality and Social Psychology Bulletin, 23*(10), 1034–1045.
Shweder, R. A., Much, N. C., Mahapatra, M., & Park, L. (1997). The "Big Three" of morality (autonomy, community, and divinity), and the "Big Three" explanations of suffering. In A. M. Brandt & P. Rozin (Eds.), *Morality and health* (pp. 119–169). New York, NY: Routledge.
Sperber, D., & Baumard, N. (2012). Moral reputation: An evolutionary and cognitive perspective. *Mind & Language, 27*(5), 495–518.
Stich, S. (2006). Is morality an elegant machine or a kludge? *Journal of Cognition and Culture, 6*(1), 181–189.
Symons, D. (1979). *The evolution of human sexuality*. New York: Oxford University Press.
Tomasello, M., & Vaish, A. (2013). Origins of human cooperation and morality. *Annual Review of Psychology, 64*, 231–255.
Tooby, J., Cosmides, L., Sell, A., Lieberman, D., & Sznycer, D. (2008). Internal regulatory variables and the design of human motivation: A computational and evolutionary approach. In *Handbook of approach and avoidance motivation* (Vol. 15, p. 251).
Trivers, R. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46*(1), 35–57.
Wilson, E. O. (2012). *The social conquest of earth*. New York: WW Norton & Company.
Wright, R. (1995). *The moral animal: The new science of evolutionary psychology*. New York: Vintage Books.

shackelf@oakland.edu

Morality as Cooperation: A Problem-Centred Approach

Oliver Scott Curry

Introduction

Your country is under attack and you are preparing to join the fight to defend it. Just then, your mother calls and tells you she is seriously ill and needs your help. Do you take care of your mother, or do you abandon her to fight for your country? You are a member of a sports team that always loses to a rival team. You have an opportunity to join that rival team. Do you take it? You borrow £10 from a wealthy friend. The friend forgets all about it. Do you give him the £10 back? You and another friend are walking along the street when you spot a £20 note on the ground. You bend down and pick it up. Do you offer to share it with your friend?

In most people, these scenarios evoke a range of thoughts, feelings, emotions, and intuitions about what to do, what is the right thing to do, what one ought to do—what is the *moral* thing to do. What are these moral thoughts and feelings, where do they come from, how do they work, and what are they for? Scholars have struggled with these questions for millennia, and for many people the nature of morality is so baffling that they assume it must have a supernatural origin (Pew, 2014).

The good news is that we now have a scientific answer to these questions. Previous approaches have noticed that morality has *something* to do with cooperation (see Table 1). But now it is possible to use the mathematical theory of cooperation—the theory of nonzero-sum games—to transform this commonplace into a precise and comprehensive theory, capable of making specific testable predictions about the nature of morality.

In this chapter, I use game theory to identify the fundamental problems of human social life, and show how—in principle and in practice—they are solved. I argue

O.S. Curry (✉)
Institute of Cognitive & Evolutionary Anthropology, University of Oxford,
64 Banbury Road, Oxford OX2 6PN, UK
e-mail: oliver.curry@anthro.ox.ac.uk

© Springer International Publishing Switzerland 2016

T.K. Shackelford, R.D. Hansen (eds.), *The Evolution of Morality*,
Evolutionary Psychology, DOI 10.1007/978-3-319-19671-8_2

shackelf@oakland.edu

Table 1 Some previous views of morality and cooperation

Aristotle	Justice is 'what is for the benefit of the whole community' or 'to the common advantage' (Aristotle, 1992, p. 207, 1160a10-14)
St. Augustine	Human law consists of 'an ordered concord of civic obedience and rule in order to secure a kind of co-operation of men's wills for the sake of attaining the things which belong to this mortal life' (Augustine, 1998, p. 945)
Thomas Aquinas	'If then a group of free men is directed by a rule to the common good of the group, his government will be right and just …' (Aquinas, 1988, pp. 15–16)
David Hume	Moral passions promote the 'public interest', the 'public good', a 'common end', 'the general interests of society', and 'the good of mankind' (Hume, 1739/1985, p. 532, p. 580, p. 590, p. 620, p. 628)
Bishop Joseph Butler	'That mankind is a community, that we all stand in a relation to each other, that there is a public end and interest of society which each particular is obliged to promote, is the sum of morals' (Butler, 1856, IX)
Bertrand Russell	'[M]en's desires conflict, and 'good' is, to my mind, mainly a social concept, designed to find issue from this conflict' (Russell, 1927, p. 230)
Henry Hazlitt	'Social cooperation is the foremost means by which the majority of us attain most of our ends. It is on the implicit if not the explicit recognition of this that our codes of morals, our rules of conduct, are ultimately based. 'Justice' itself … consists in observance of the rules or principles that do most, in the long run, to preserve and promote social cooperation' (Hazlitt, 1964)
John Rawls	'The circumstances of justice may be described as the normal conditions under which human cooperation is both possible and necessary' (Rawls, 1971, p. 126)
John Mackie	'Protagoras, Hobbes, Hume and Warnock are all at least broadly in agreement about the problem that morality is needed to solve: limited resources and limited sympathies together generate both competition leading to conflict and an absence of what would be mutually beneficial cooperation' (Mackie, 1977, p. 111)
David Wong	'Human beings have needs to resolve internal conflicts between requirements and to resolve interpersonal conflicts of interest. Morality is a social creation that evolved in response to these needs' (Wong, 1984, p. 175)
Daniel Hausman and Michael McPherson	'[T]he normative principles governing individual interactions are human contrivances to adjudicate conflicts of interest and to secure the benefits of cooperation' (Hausman & McPherson, 1996, p. 186)
Jonathan Haidt	'Moral systems are interlocking sets of values, virtues, norms, practices, identities, institutions, technologies, and evolved psychological mechanisms that work together to suppress or regulate selfishness and make cooperative social life possible' (Haidt & Kesebir, 2010)
Alan Fiske	'Morality functions to facilitate the generation and maintenance of long-term social-cooperative relationships with others' (Rai & Fiske, 2011)
Michael Tomasello	'Human morality arose evolutionarily as a set of skills and motives for cooperating with others' (Tomasello & Vaish, 2013)
Joshua Greene	'[T]he core function of morality is to promote and sustain cooperation' (Greene, 2015)

shackelf@oakland.edu

that it is the solutions to these problems that philosophers and others have called 'morality'. Thus, morality turns out to be a collection of biological and cultural solutions to the problems of cooperation and conflict recurrent in human social life. I show how this theory of 'morality as cooperation' incorporates the best elements of previous theories, and moves beyond them to create a principled taxonomy of moral values of unprecedented depth and breadth. I derive from this theory testable predictions about the structure and content of moral thought and outline how they differ from those of rival theories. And I conclude that, because the debate between these theories can be resolved using standard scientific method, the study of morality has at last become a branch of science. Let's get started.

A Natural History of Morality

Life begins when molecules start making copies of themselves. These 'replicators' are 'selfish' in the technical sense that they promote their own replication (Dawkins, 1976/2006). But they can promote their own replication at the expense of other replicators, or in concert with them (Dawkins, 1998). Game theory analyses these competitive and cooperative interactions as zero-sum and nonzero-sum, respectively (Maynard Smith, 1982; Von Neumann & Morgenstern, 1944). Competitive zero-sum interactions have a winner and a loser; one's gain is another's loss. But cooperative nonzero-sum interactions can have two winners; they can be win–win situations.

Natural selection for genes that employ cooperative strategies has driven several 'major transitions' in the evolution of life on Earth, including the formation of cells, chromosomes and multicellular organisms (Maynard Smith & Szathmáry, 1995). Natural selection has also favoured genes for cooperation between individuals, in a wide variety of species (Dugatkin, 1997), including humans. Humans descend from a long line of social primates; they have spent 50 million years living in social groups (Shultz, Opie, & Atkinson, 2011) and two million years making a living as intensely collaborative hunter–gatherers (Tooby & DeVore, 1987). This has equipped humans with a range of biological—including psychological—adaptations for cooperation. These adaptations can be seen as natural selection's attempts to solve the problems of cooperation. And ever since entering the 'cognitive niche' (Boyd, Richerson, & Henrich, 2011; Pinker, 2010), humans have attempted to improve upon natural selection's solutions by inventing evolutionarily novel cultural solutions—'tools and rules'—for further bolstering cooperation (Binmore, 1994a, 1994b; Nagel, 1991; Popper, 1945).

Together, these biological and cultural mechanisms provide both the motivation for social, cooperative and altruistic behaviour—leading individuals to value and pursue specific mutually beneficial outcomes—and the standards by which individuals evaluate the social behaviour of others. And it is precisely these mechanisms—these solutions to problems of cooperation this collection of instincts, intuitions, ideas, and institutions that constitute human morality (Curry, 2005).

shackelf@oakland.edu

This theory of morality as cooperation predicts that there will be not one but many domains of morality. This is because game theory tells us that there is not one problem of cooperation, but many, and many solutions. And the theory predicts what these problem-centred domains will be: (1) the allocation of resources to kin; (2) coordination to mutual advantage; (3) exchange; and (4) conflict resolution by means of (a) contests featuring displays of hawkish and dove-ish traits, (b) division, and (c) possession. Let's look at each of these problems, how natural selection and human ingenuity have attempted to solve them, and what predictions this problem-centred approach makes about human morality.

(1) Kinship

A gene has the potential to influence not only its own replication but also the replication of replicas of itself. In some situations, a gene in one individual can best promote its replication by diverting resources to copies of itself that reside in other individuals—that is, in genetic relatives or family members. Genes that benefit replicas will be favoured by natural selection if the cost of helping is outweighed by the benefit to the recipient gene(s) (Dawkins, 1979; Hamilton, 1964). So, evolutionary theory leads us to expect that organisms will possess adaptations for detecting and delivering benefits (or avoiding harm) to kin.

And, as expected, numerous species do indeed have adaptations for identifying (Hepper, 1991) and being altruistic to genetic relatives—with parental care and eusociality among insects being the most widespread and conspicuous examples (Clutton-Brock, 1991; Royle, Smiseth, & Kölliker, 2012).

Humans and their recent primate ancestors have always lived in groups composed mostly of genetic relatives, and so they have always faced the problem of allocating resources to kin (Chapais, 2014). Research into adaptations for kin altruism in humans has focussed on kin detection and incest aversion (Lieberman, Tooby, & Cosmides, 2003, 2007), paternal investment (Geary, 2000) and its absence (Daly & Wilson, 1996), and the effects of uncertainty of paternity on paternal and grandparental investment (Euler & Weitzel, 1996; Gaulin & Schlegel, 1980; Platek et al., 2003). Culturally, humans have invented institutions—such as naming conventions (Oates & Wilson, 2002) and inheritance rules (Smith, Kish, & Crawford, 1987)—to extend the reach of kin altruism. Behaviourally, kin altruism in humans is evident in the universality of family structure in human societies, patterns of alliance (Chagnon & Bugos, 1979), and homicide (Daly & Wilson, 1988). Humans have also invented a variety of rules for regulating inbreeding and avoiding incest (Thornhill, 1991).

Morality as cooperation predicts that solutions to the problem of efficiently allocating resources to kin—such as caring for offspring, helping family members, and avoiding inbreeding—are component parts of human morality and will be considered morally good. And there is evidence to suggest that they are.

For example, Edvard Westermarck's classic cross-cultural survey of ethics concluded: 'There is one duty so universal and obvious that it is seldom mentioned:

shackelf@oakland.edu

the mother's duty to rear her children…Another duty…is incumbent on the married man: the protection and support of his family' (Westermarck, 1906). The anthropologist May Edel and her philosopher husband Abraham Edel concurred: 'the moral obligation for a mother to take care of her children…is a universal imperative' (Edel & Edel, 1959/1968). And in Confucian ethics, 'Duty to the family trumped all other duties' (Fukuyama, 1996). Obligations to family—an ethic of care, an obligation to distribute goods on the basis of need and relationship, not abstract rules—also figure prominently in some feminist moral philosophy (Noddings, 1978; Ruddick, 1980). And 'the horror of incest is well nigh universal in the human race' (Westermarck, 1906).

(2) Mutualism

Situations in which individuals benefit more by working together than they do by working alone are referred to as mutualisms (Connor, 1995). Such mutualisms can provide economies of scale, efficient divisions of labour, and strength (or safety) in numbers. Darwin provides a typically charming example of the benefits of teamwork: 'Hamadryas baboons turn over stones to find insects, &c.; and when they come to a large one, as many as can stand round, turn it over together and share the booty' (Darwin, 1871). Because individuals must coordinate their behaviour in order to realise these benefits, these situations are modelled as coordination problems (Lewis, 1969; Schelling, 1960)—including 'stag hunts' (Skyrms, 2004) and soldier's dilemmas' (Clutton-Brock, 2009)—and the ensuing relationships are referred to as friendships, alliances, and coalitions (Tooby & Cosmides, 1996).

In principle, coordination problems can be solved by focal points and precedence ('return to the same breeding grounds each year'), simple decision rules ('follow the leader'; Van Vugt, Hogan, & Kaiser, 2008), signalling and communication ('I'm over here!'), as well as more sophisticated abilities to anticipate and predict others' behaviour (proto-theory of mind; Whiten, 1996). There has been relatively little empirical work on adaptations for coordination per se (but see Boos, Kolbe, Kappeler, & Ellwart, 2011). However, there is little doubt that many species are able to solve coordination problems, as evident in the ubiquity of herds, shoals, flocks, and collaborative hunting (Boinski & Garber, 2000; Clutton-Brock, 2009), as well as the formation of alliances and coalitions (Bissonnette et al., 2015; Harcourt & de Waal, 1992).

The problem of coordinating to mutual advantage has been a recurrent feature of the social lives of humans and their recent ancestors, especially with regard to collaborative hunting (Alvard, 2001; Alvard & Nolin, 2002) and forming coalitions to compete with rival coalitions (Wrangham, 1999). Research on adaptations for mutualism and coordination in humans has focussed on coalitionary psychology (Kurzban, Tooby, & Cosmides, 2001; Tooby & Cosmides, 2010), adaptations for representing common knowledge (Thomas, DeScioli, Haque, & Pinker, 2014), and 'theory of mind' (Curry & Jones Chesters, 2012; Tomasello, Carpenter, Call, Behne,

shackelf@oakland.edu

& Moll, 2005; Young, Camprodon, Hauser, Pascual-Leone, & Saxe, 2010). Theory of mind, in particular, seems to have taken human cooperation to new heights. This ability allows us to think about what others are thinking; to infer their desires, beliefs, and intentions; and to factor these into our judgments of their conduct—distinguishing, for example, between intentional and accidental harms. Theory of mind also seems to play a central role in the formation of conventions and other 'social constructions' that can be used to solve an indefinite array of novel coordination problems (Berger & Luckmann, 1966). Culturally, humans have enhanced their ability to coordinate their behaviour by means of maps, clocks, calendars and communication technology, and badges of membership. Behaviourally, mutualism is apparent in the widespread and spontaneous tendency of humans to form groups and to benefit those groups at the expense of others (Balliet, Wu, & De Dreu, 2014; Sherif, Harvey, White, Hood, & Sherif, 1954/1961; Tajfel, 1970).

Morality as cooperation predicts that solutions to the problems of mutualism—such as forming friendships, participating in collaborative endeavours, favouring your own group, and adopting local conventions—are component parts of human morality and will be considered morally good. There is evidence to suggest that they are.

Aristotle devoted two books of his *Nichomachean Ethics* to friendship (Aristotle, 1962); for Cicero, friendship was 'the noblest and most delightful of all the gifts the gods have given mankind' (Cicero, 1971); and G. E. Moore ranked friendship as one of 'the most valuable things that we can know or imagine' and the one that provides the only justification for 'performing any public or private duty' (Moore, 1903). Plato argued that life was a one big coordination problem, and that justice consisted of an efficient division of labour where everyone played their part (Plato, 1974). Loyalty—commitment to a common cause, such as the 'devotion of a patriot to his country'—has been described as 'the heart of all the virtues, the central duty amongst all duties' (Royce, 1908). More recently, many theorists have agreed that loyalty—'giving special consideration to a person or group of persons' (Gert, 2013, p. 18)—is a moral issue, even if they have not agreed on the reasons why (Levinson, Parker, & Woodruff, 2013). And the moral philosopher Allan Gibbard has argued that people possess 'biological adaptations for coordination' that enable them to identify and adopt norms and conventions and thereby coordinate individuals to mutual advantage: 'The key to human moral nature lies in coordination broadly construed' (Gibbard, 1990a, 1990b).

(3) Exchange

In some situations, the benefits of mutualism are uncertain, perhaps because the benefits are transferred at different times; here, individuals might be exploited by 'free riders', who accept a benefit, but neglect to return it. These situations are modelled as prisoner's dilemmas (social dilemmas, public goods games, and so on) (Ostrom & Walker, 2002)—games in which non-cooperation is the only viable

shackelf@oakland.edu

strategy. However, if individuals meet repeatedly, then the situation becomes an 'assurance game', and cooperation can be maintained by a strategy of conditional cooperation—such as 'tit for tat'—that begins by cooperating and then reciprocates the other individual's behaviour (returning a benefit or avenging an injury) (Axelrod, 1984; Trivers, 1971).

Surprisingly, few if any examples of full-blown 'reciprocal altruism' have been found in non-human species (Amici et al., 2014; Clutton-Brock, 2009), although some aspects of reciprocity have been identified in cleaner fish (Bshary & Grutter, 2006), vampire bats (Carter & Wilkinson, 2013), and primates (Mitani, 2009).

Social exchange may have been a recurrent feature of the social lives of humans since our last common ancestors with chimpanzees six million years ago (Jaeggi & Gurven, 2013); and there is some suggestive evidence for trade between groups from 82,000 years ago (Bouzouggar et al., 2007). Research on adaptations for exchange in humans has focussed on trust (Kosfeld, Heinrichs, Zak, Fischbacher, & Fehr, 2005), gratitude (McCullough, Kimeldorf, & Cohen, 2008), cheater detection (Cosmides & Tooby, 2005), punishment (Price, Cosmides, & Tooby, 2002), revenge, and forgiveness (McCullough, Kurzban, & Tabak, 2013). Culturally, humans have extended the scope of exchange and reciprocity through such 'technologies of trust' as money, written contracts, 'mechanical cheater detectors' such as '[c]ash register tapes, punch clocks, train tickets, receipts, accounting ledgers', handcuffs, prisons, electric chairs, CCTV, branding of criminals, and criminal records (Pinker, 1997). Behaviourally, reciprocity emerges early in children's behaviour (Harbaugh, Krause, Liday, & Vesterlund, 2002) and is used as a strategy for social exchange cross-culturally (Henrich et al., 2005; Kocher, Cherry, Kroll, Netzer, & Sutter, 2008).

Morality as cooperation predicts that solutions to the problems of exchange—especially the mechanisms that implement reciprocity—are component parts of human morality and will be considered morally good. There is evidence to suggest that they are.

Reciprocity in general is the guiding principle of many moral philosophies. When asked for a single word that could sum up morality, Confucius answered: 'Reciprocity perhaps? Do not inflict on others what you yourself would not wish done to you' (Confucius, 1994). 'Social contract' theorists—from 'Glaucon' (Plato, 1974) to Hobbes (1651/1958) to Rawls (1971)—have viewed all of morality through the lens of reciprocity. The golden rule of 'do as you would be done by' is present in all major world religions (Chilton & Neusner, 2009). And in its negative form, reciprocity provides the guiding principle of theories of punishment and retribution—from the Code of Hammurabi's 'eye for an eye' onwards (Daly & Wilson, 1988). The specific subcomponents of reciprocity—trust (Baier, 1995), patience (Curry, Price, & Price, 2008), gratitude (Emmons, 2004), guilt (Gibbard, 1990b), apology (Ohtsubo & Watanabe, 2009), and forgiveness (Downie, 1965; Godfray, 1992; Richards, 1988)—have also been regarded as important facets of morality.

shackelf@oakland.edu

(4) Conflict Resolution

Organisms often come into conflict over resources such as food, territory, and mates (Huntingdon & Turner, 1987). Although such conflicts appear zero-sum, in fact there are costs involved in conflict—time, energy, and injury—that individuals have a common interest in avoiding. For this reason, animal conflicts are modelled not as zero-sum games, but as nonzero-sum hawk–dove games, in which the worst outcome occurs only if both players adopt a 'hawkish' strategy of all-out aggression (Maynard Smith & Price, 1973). Thus, conflict presents combatants with an opportunity to cooperate, by competing in less mutually destructive ways. There are three ways of achieving this: contests (featuring the display of hawkish and dove-ish traits), division, and possession.

(a) Contests

Instead of fighting, one option is for contestants to display reliable indicators of 'fighting ability' (or 'resource holding power' or 'formidability') and for the weaker party to cede the resource to the stronger. In this way, the stronger party still wins, but both avoid the costs of a real fight (Gintis, Smith, & Bowles, 2001; Maynard Smith & Price, 1973).

Animal contests in which contestants follow such 'display and defer' strategies are widespread in nature. Depending on the species, 'hawkish' displays of size, weight, age, or experience may carry the day (Hardy & Briffa, 2013; Riechert, 1998). Such displays may also involve costly acts that benefit others (Zahavi & Zahavi, 1997). Conversely, 'dove-ish' cues of submission involve exaggerated concealment of these same attributes, or conspicuous displays of their absence (Darwin, 1872/1998; Preuschoft & van Schaik, 2000). In stable social groups, in which relative 'power' is already known by reputation (through direct experience or third-party observation), individuals can dispense with the contest, and allocate disputed resources by 'rank'. Such 'dominance hierarchies' represent a further de-escalation of conflict, and are also widespread in nature (Preuschoft & van Schaik, 2000).

Humans and their recent ancestors have always faced the problem of conflict resolution, because such problems are inherent in group living (Shultz & Dunbar, 2007). Research into human adaptations for resolving conflicts via contests has focussed on cues of dominance and deference, including facial expressions, voice pitch, and height (Sell et al., 2010; Sell et al., 2009; Watkins et al., 2010), and testosterone—the hormonal system responsible for prompting competitive displays, elating winners, and deflating losers (Mazur, 2005). And experiments suggest that a tendency for the strong to display status by helping the weak—*noblesse oblige*—is present cross-culturally (Fiddick, Cummins, Janicki, Lee, & Erlich, 2013). Culturally, humans have invented numerous means of minimising the costs of

shackelf@oakland.edu

conflict through stylised contests—including single ('champion') combat (Cowan, 2007), duels, tournaments, rules of combat (Queensberry rules, Geneva Conventions), and competitive games and sports (Deaner & Smith, 2012). There has been relatively little research on human adaptations for navigating hierarchies, apart from the finding that human hierarchies are less pronounced than those of our nearest primate relatives (Boesch, 1999; Gavrilets, Duenez-Guzman, & Vose, 2008). But culturally, humans have invented countless ways of displaying status and regulating relationships accordingly, such as honorifics, etiquette, dress codes, medals, decorations and honours, and caste systems. Behaviourally, humans—especially males—commonly engage in costly and conspicuous displays of prowess, resources, and even altruism, especially in the context of mate competition (Hardy & Van Vugt, 2006; Hawkes, 1991; Hawkes, O'Connell, & Blurton Jones, 2001; Miller, 2000). Children spontaneously form dominance hierarchies relatively early in their development (Edelman & Omark, 1973), and status hierarchies are a ubiquitous feature of human societies (Boone, 1992; Rubin, 2000).

Morality as cooperation predicts that resolving conflicts by means of contests will give rise to two apparently opposing sets of moral values, reflecting the two branches of the 'display–defer' strategy—the virtues of the hawk and the virtues of the dove. The theory predicts that hawkish signals of prowess (strength, fortitude, bravery, heroism generosity, largesse) and also dove-ish displays of submission (humility, deference, respect, obedience) are component parts of human morality and will be considered morally good. There is evidence to suggest that they are.

Traits that establish status and forestall disputes have been celebrated as 'excellences' or 'virtues' throughout history (MacIntyre, 1981a, 1981b). The philosopher David Hume gives a particularly cogent account (Hume, 1739/1985). He recognised that many animals take pride in their 'beauty, strength, swiftness'; in addition, humans take pride in their 'imagination, judgment, memory or disposition; wit, good-sense, learning, courage, justice, [and] integrity', and differences in the ability give rise to hierarchies in which 'certain deferences and mutual submissions' are required 'of the different ranks of men towards each other'. High status then motivates altruistic acts by fostering the 'heroic virtues': '[c]ourage, intrepidity, ambition, love of glory, magnanimity, and all the other shining virtues'. Hume contrasted these 'heroic' virtues with the 'monkish' virtues of '[c]elibacy, fasting, penance, mortification, self-denial, humility, silence, solitude', and so on (Hume, 1757/1889). A monkish virtue such as humility—'a just sense of our weakness'—'is esteem'd virtuous, and procures the good-will of everyone' (Hume, 1757/1889). Aristotle, Machiavelli, Nietzsche, and Mill have celebrated similar virtues, for similar reasons (Curry, 2007). And, in keeping with the theory, the original meaning of 'respect' evoked 'an element of fear' directed towards 'dangerous things'. 'In olden days… the scale of respect was one with the scales of power and status'. Later, the term came to be applied not just to physical power, but to the power of ideas, 'not the ability to make demands backed up by force, but the ability to make claims backed up by reasons', and in this way, 'moral terms which in their original senses had to

shackelf@oakland.edu

do with power, pressure, force, coercion…come to be applied to 'moral' force, or power' (Feinberg, 1973).

Consistent with the theory, both hawkish and dove-ish traits tend to be seen as moral when there is an obvious power differential—as in Plato's Republic (workers ought to obey their 'virtuous' philosophical superiors), Aristotle's polis (slaves ought to obey their 'rational' masters), and feudal monarchies (subjects ought to obey their 'divine' sovereigns). Similarly, respect and obedience seem appropriate when arguing that children ought to obey their parents or soldiers ought to obey their superior officers. But, as the theory also predicts, in societies that are, or profess to be, more equal—such as Western, Educated, Industrialised, Rich Democracies (WEIRD) (Henrich, Heine, & Norenzayan, 2010)—deference and respect for power appear 'obsolete' (Berger, 1970).

(b) Division

If the contested resource is divisible (such as spoils from a hunt, or a disputed border between territories), then game theory models the situation as a 'bargaining problem' (Nash, 1950). Here, one solution is to divide the resource in proportion to the relative (bargaining) power of the protagonists (Skyrms, 1996). In the case of equally powerful individuals, this results in equal shares (Maynard Smith, 1982).

Among animals, indirect evidence for a 'sense of fairness' in non-human primates comes from reactions to unequal treatment in economic games (Brosnan, 2013).

There has been relatively little research on human adaptations for resolving conflicts using division. It has been found that males with elevated levels of testosterone make (Zak et al., 2009) and reject (Burnham, 2007) lower offers in ultimatum bargaining games. And there is also some evidence that individuals will exhibit deference to the preferences of more powerful individuals (de Kwaadsteneit & van Dijk, 2010). Nevertheless, rules such as 'I cut, you choose', 'meet in the middle', 'split the difference', and 'take turns' are ancient and widespread means of resolving disputes (Brams & Taylor, 1996). And behaviourally, it has been found that 'equal shares' is a spontaneous and cross-culturally prevalent decision rule in economic games (Güth, Schmittberger, & Schwarze, 1982; Henrich et al., 2005) and other situations (Messick, 1993).

Morality as cooperation predicts that resolving conflicts by means of division—negotiation, compromise, fairness—is a component part of human morality and will be considered morally good. There is evidence to suggest that it is.

Negotiating a compromise—whether directly between two individuals, or by means of a third party (arbitration, mediation)—has been described as a 'fair and rational way of reaching a reasonable agreement' (Pennock & Chapman, 1979). And fairness itself has been viewed as synonymous with morality, as in John Rawls' (1958) influential work 'Justice as Fairness'.

shackelf@oakland.edu

(c) Possession

Finally, game theory shows that conflicts over resources can also be resolved by deference to prior ownership (Gintis, 2007; Maynard Smith, 1982). The recognition of prior ownership is widespread in nature: 'in almost all territorial species, intruders respect territory ownership'—'The space that a territory owner defends is functionally equivalent to his property, and an intruder's respect reveals his acknowledgment of ownership and property rights' (Hauser, 2001, p. 303; see also Strassmann & Queller, 2014).

There has been relatively little research on human adaptations for ownership—although some have interpreted the 'endowment effect' (Gintis, 2007; Kahneman & Tversky, 1979) and international disputes over territory (Johnson & Toft, 2014) in this light. Culturally, humans have invented a range of institutions—title and land registries—to keep track of who owns what (No Title, 2001), and 'first possession' is the basis of much property law (Rose, 1985). Behaviourally, the notion that objects can be 'owned' emerges early in child development (Friedman & Neary, 2008; Ross & Friedman, 2011) and (in various forms) is cross-culturally universal: 'in all groups personal ownership of some goods and rights exists…private property, in this sense, is known everywhere' (Herskovits, 1952, p. 372); 'the phenomenon is a universal one, since there is no group who live so precariously that there is not some tool, some weapon, some bit of ornament or clothing that is not regarded as indisputably the possession of its maker, its user, its wearer' (Herskovits, 1952, p. 327).

Morality as cooperation predicts that resolving conflicts by deferring to prior ownership—respecting others' property and territory and not stealing—is a component part of human morality and will be considered morally good. There is evidence to suggest that it is.

In another astute analysis, David Hume noted that property rights are acquired primarily through 'first possession' or 'occupation', and he argued that such rights serve 'to cut off all occasions of discord and contention' (Hume, 1739/1985). Many others have agreed that there can be a moral right to own property, even while disagreeing as to the reasons why (Becker, 1977; Locke, 2000; Pennock & Chapman, 1980). And Westermarck reports that 'When we examine the moral rules of uncivilised races…[i]n every savage community homicide is prohibited by custom, and so is theft' (Westermarck, 1906).

A Periodic Table of Ethics

Thus, morality as cooperation predicts that there will be multiple moral domains, and it predicts what these domains will be. It uses the game theory of cooperation to create a novel taxonomy of moral values—a 'Periodic Table of Ethics'—that incorporates a wide variety of moral phenomena: obligations to family, group loyalty, reciprocity, bravery, respect for hierarchy, fairness, and property rights (see Table 2).

shackelf@oakland.edu

Table 2 A periodic table of ethics: an overview of morality as cooperation

Problem	Theory	Animal examples	Human examples	Morals
Kinship	Kin selection (Dawkins, 1979; Hamilton, 1964)	Kin recognition (Hepper, 1991), parental care (Clutton-Brock, 1991; Royle et al., 2012)	Kin detection and incest aversion (Lieberman et al., 2003, 2007), paternal investment (Geary, 2000), patterns of homicide (Daly & Wilson, 1996). Rules against incest (Thornhill, 1991)	Obligations to kin (Fukuyama, 1996), duty of parental care (Edel & Edel, 1959/1968; Westermarck, 1906), prohibition of incest (Westermarck, 1906)
Mutualism	Mutualism (Connor, 1995), coordination (Lewis, 1969; Schelling, 1960), coalition formation (Tooby & Cosmides, 1996; Von Neumann & Morgenstern, 1944)	Mutualism (Clutton-Brock, 2009), coordination (Boinski & Garber, 2000; Boos et al., 2011), coalitions (Bissonnette et al., 2015; Harcourt & de Waal, 1992)	Coalitionary psychology (Kurzban et al., 2001), common knowledge (Thomas et al., 2014), 'theory of mind' (Tomasello et al., 2005). Ingroup favouritism (Balliet et al., 2014; Sherif et al., 1954/1961; Tajfel, 1970). Social construction (Berger & Luckmann, 1966)	Friendship (Aristotle, 1962), loyalty (Royce, 1908), conformity (Gibbard, 1990a, 1990b)
Exchange	Reciprocal altruism (Axelrod, 1984; Trivers, 1971)	Vampire bats? (Carter & Wilkinson, 2013)	Trust (Kosfeld et al., 2005), gratitude (McCullough et al., 2008), cheater detection (Cosmides & Tooby, 2005), punishment (Price et al., 2002), revenge and forgiveness (McCullough et al., 2013). Technologies of trust (Pinker, 1997). Ubiquity of reciprocity (Henrich et al., 2005; Kocher et al., 2008)	Reciprocity (Rawls, 1971), punishment (Daly & Wilson, 1988), trust (Baier, 1995), gratitude (Emmons, 2004), guilt (Gibbard, 1990b), apology (Ohtsubo & Watanabe, 2009), forgiveness (Downie, 1965; Godfray, 1992; Richards, 1988)

(continued)

shackelf@oakland.edu

Table 2 (continued)

Problem	Theory	Animal examples	Human examples	Morals
Conflict resolution: contests (hawk–dove)	Animal conflict and costly signals (Gintis et al., 2001; Maynard Smith & Price, 1973), dominance and deference (Mazur, 2005)	Animal contests (Hardy & Briffa, 2013; Riechert, 1998), dominance hierarchies (Preuschoft & van Schaik, 2000)	Formidability (Sell et al., 2010), costly signalling (Hawkes, 1991; Hawkes et al., 2001; Miller, 2000), *noblesse oblige* (Fiddick et al., 2013), dominance and deference (Mazur, 2005). Games and sports (Deaner & Smith, 2012). Ubiquity of status hierarchies (Boone, 1992; Rubin, 2000)	Virtues and excellences (Curry, 2007; MacIntyre, 1981b). Hawkish virtues (fortitude, bravery, skill, generosity, beauty) (Hume, 1739/1985), dove-ish virtues (humility, respect, deference, obedience) (Feinberg, 1973; Hume, 1757/1889)
Conflict resolution: division	Bargaining and fairness (Maynard Smith, 1982; Nash, 1950; Skyrms, 1996)	Primates? (Brosnan, 2013)	Ultimatum games (Güth et al., 1982; Henrich et al., 2005), equality (Messick, 1993). 'Cut the cake' (Brams & Taylor, 1996)	Fairness (Rawls, 1958), negotiation, and compromise (Pennock & Chapman, 1979)
Conflict resolution: possession	Prior ownership (Gintis, 2007; Maynard Smith, 1982)	Ownership and territoriality (Strassmann & Queller, 2014)	Endowment effect (Gintis, 2007; Kahneman & Tversky, 1979), territoriality (Johnson & Toft, 2014). Property law (Rose, 1985). Ubiquity of property (Herskovits, 1952)	Property rights (Becker, 1977; Hume, 1739/1985; Locke, 2000; Pennock & Chapman, 1980). Theft (Westermarck, 1906)

And, as we have just seen, this approach receives some support from the existing literature on morality. But morality as cooperation is also brimming with further novel testable predictions about the structure and content of moral thought. Developing this promising, principled, problem-centred approach will involve making these predictions explicit and putting them to the test.

First, the good, the bad, and the neutral. As we have seen, morality as cooperation predicts that people will regard specific types of cooperative behaviour—behaviour that solves some problem of cooperation—as morally good. Thus, people will regard helping your family, being loyal to your group, reciprocating favours, being brave, deferring to authority, dividing disputed resources, and respecting property, as morally good. And they will regard failing to cooperate—by neglecting your family, betraying your group, cheating, being cowardly, rebelling against

shackelf@oakland.edu

authority, being unfair, and stealing—as morally bad. The theory also predicts that behaviour that has nothing to do with cooperation—nonsocial behaviour or competition in zero-sum games ('all's fair in love and war')—will be regarded as morally neutral.

Second, universality and diversity. Morality as cooperation also predicts that—because these problems are universal features of human social life—these cooperative behaviours will be considered morally good in every human culture, at all times and in all places. There will be no cultures where morality is about something other than cooperation—say, aesthetics or nutrition. And there will be no cultures where helping your family, being loyal to your group, reciprocating favours, being brave, deferring to authority, dividing disputed resources, respecting property, and so on are considered morally bad. However, the theory does not predict that moral systems will everywhere be identical. On the contrary, the prediction is that, to the extent that different people and different societies face different portfolios of problems, different domains of morality will loom larger—different cultures will prioritise different moral values. For example, differences in family size, frequency of warfare, or degree of inequality may lead to differences in the importance attached to family values, bravery, and respect.

Third, uncharted territory. Morality as cooperation predicts that as yet poorly understood aspects of morality will also turn out to be about cooperation. For example, sexual morality will consist of a collection of solutions to the specific problems of cooperation and conflict that arise within and between the sexes. Political morality will regard leaders as morally good if they promote cooperation among their followers—by solving coordination problems (especially in the context of group defence), enforcing contracts, punishing cheats, resolving (violent) conflicts, displaying prestigious virtues (especially bravery and wisdom), maintaining hierarchies, impartially arbitrating disputes, redistributing the rewards of collective action equitably, and respecting their subjects' property. Conversely, morally bad leaders will be those who do none of the above and instead parasitise their followers' cooperation. Ethics in international relations—grand alliances, trade agreements, diplomacy, rules of war, and so on—will consist of solutions to the problems of cooperation that arise between groups, as opposed to individuals. Religious morality—ancestor worship, food taboos, karma, reverence, and so on—will turn out to be the product of mechanisms designed for mundane cooperation (McKay & Whitehouse, 2014).

Finally, extending the foundations. Morality as cooperation predicts that developments in game theory will expand the theory's explanatory power. Already, by drawing on all nonzero-sum games, the theory goes beyond most existing reviews of cooperation, which tend to focus on kin and reciprocal altruism, and overlook mutualism and conflict resolution (see Table 3). The discovery of new game-theoretical problems and solutions will open up new horizons for the explanation of further aspects of morality.

shackelf@oakland.edu

Table 3 Previous reviews of cooperation are incomplete

			Kin	Mutualism	Exchange	Contest (hawk)	Contest (dove)	Division	Possession	Other
Cooperation among animals	Dugatkin (1997)	Kinship	1							
		Reciprocity			1					
		Byproduct mutualism		1						
		Group selection								1
Cooperation and collective action in animal behaviour	Nunn and Lewis (2001)	Kinship	1							
		Prisoner's dilemma			1					
		Coordination/mutualism		1						
		Chicken/hawk–dove				1	1			
The evolution of cooperation	Sachs, Mueller, Wilcox, and Bull (2004)	Directed reciprocation			1					
		Shared genes	1							
		By-product benefits		1						
Five rules for the evolution of cooperation	Nowak (2006)	Kin selection	1							
		Direct reciprocity			1					
		Indirect reciprocity			1					
		Network reciprocity			1					
		Group selection								1
The evolution of cooperation and altruism	Lehmann and Keller (2006)	Direct benefits		1						
		Reciprocation			1					
		Kin selection	1							
		Greenbeard	1							
Evolutionary explanations for cooperation	West, Griffin, and Gardner (2007)	Direct benefits			1					
		Indirect benefits	1							
		By-product benefits		1						

shackelf@oakland.edu

Alternative Alchemies

Morality as cooperation is a naturalistic theory grounded in our understanding of the material world; it draws on the latest insights from empirical sciences such as ethology, psychology, and anthropology; it offers a unified, universal view of morality; and it uses the principles of game theory to identify specific problems of cooperation and their corresponding solutions and to make predictions about moral phenomena. As such, morality as cooperation differs from existing theories in a number of ways.

It differs from those theories that invoke the supernatural (it has no need of that hypothesis). It differs from those that attempt to explain morality using only pre-scientific folk ontologies—such as belief, desire, passion, reason, and the will (Jackson, Pettit, & Smith, 2004).

It differs from theories that maintain that there is nothing that unifies the diverse array of moral phenomena (Sinnott-Armstrong & Wheatley, 2013) and that we must therefore settle for a plethora of low-level generalisations about morality (Bartels, Bauman, Cushman, Pizarro, & McGraw, 2015).

It differs from theories that argue that the very definition of morality varies from culture to culture, that there are no universal moral values, and that morality varies radically or arbitrarily across cultures (Ladd, 1985).

It differs from theories that hold that morality is not about cooperation, but about fulfilling natural human functions or fully expressing human capacities (Arnhart, 1998; Casebeer, 2003). And it differs from theories that hold that morality is about maximising welfare, well-being or utility by any means, not necessarily cooperation (Mill & Bentham, 1987).

It differs from approaches that do not use game theory (or indeed any theory at all) to derive their taxonomies of morality and that consequently conflate, omit, and misconstrue different types of cooperation (see Table 4). For example, morality as cooperation suggests that Fiske's Relational Models (based on ethnographic field work and, oddly, the theory of measurement; Stevens, 1946), Shweder's CAD Triad (based on a small study in one culture), and Haidt's Moral Foundations (based on a literature review of five sources, including Fiske and Shweder) err in conflating kinship and mutualism, and exchange and division, and in omitting hawkish traits and possession. Further, morality as cooperation suggests that the Moral Foundations approach also errs by interpreting mutualism as group selection (Haidt, 2012) and including a category—purity, avoiding 'people with diseases, parasites [and] waste products'—that has no apparent connection to cooperation.

And, it differs from theories that, because they lack any underlying theory, cannot make principled predictions about the nature of morality (Haidt & Joseph, 2011).

shackelf@oakland.edu

Table 4 Previous moral taxonomies are incomplete

			Kin	Mutualism	Exchange	Contest (hawk)	Contest (dove)	Division	Possession	Other
Relational Models	Fiske (1992) and Rai and Fiske (2011)	Unity	1	1						
		Respect					1			
		Equality			1			1		
		Proportionality			1					
CAD Triad	Shweder, Much, Park, and Mahapatra (1997/2003)	Community	1	1			1			
		Autonomy			1			1		1
		Divinity								1
Moral Foundations	Graham et al. (2011) and Haidt and Joseph (2004)	Care	~1							1
		Fairness			1			1		
		Ingroup	1	1						1
		Authority					1			
		Purity								1

shackelf@oakland.edu

Conclusion

Morality is no mystery. We have a theory. Morality is a collection of biological and cultural solutions to the problems of cooperation and conflict recurrent in human social life; and game theory reveals what those problems and solutions are. Morality as cooperation explains what morality is, where it comes from, how it works, and what it is for.

Crucially, because this theory makes predictions about morality—predictions that can be tested against those of rival theories using standard scientific method—it makes clear that the study of morality, theory driven and empirically tested, is simply another branch of science. And it is this realisation, more than any particular theory, that will set the study of morality on the firm scientific foundation that will finally allow it to flourish.

Acknowledgements I am very grateful to Helena Cronin and Daniel Mullins for their enormously helpful comments on numerous drafts of this chapter.

References

Alvard, M. (2001). Mutualistic hunting. In C. Stanford & H. Bunn (Eds.), *Meat-eating and human evolution* (pp. 261–278). New York: Oxford University Press.

Alvard, M., & Nolin, D. (2002). Rousseau's whale hunt? *Current Anthropology, 43*(4), 533–559.

Amici, F., Aureli, F., Mundry, R., Amaro, A., Barroso, A., Ferretti, J., et al. (2014). Calculated reciprocity? A comparative test with six primate species. *Primates, 55*(3), 447–457. doi:10.1007/s10329-014-0424-4.

Aquinas, T. (1988). *On politics and ethics*. New York, London: WW Norton & Company.

Aristotle. (1962). *Nichomachean ethics* (M. Oswald, Trans.). Englewood Cliffs, NJ: Prentice-Hall.

Aristotle. (1992). *The politics*. London: Penguin Books.

Arnhart, L. (1998). *Darwinian natural right: The biological ethics of human nature*. Albany, NY: SUNY Press.

Augustine. (1998). *The city of god against the pagans*. Cambridge, UK: Cambridge University Press.

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

Baier, A. (1995). *Moral prejudices: Essays on ethics*. Cambridge, MA: Harvard University Press.

Balliet, D., Wu, J., & De Dreu, C. K. W. (2014). Ingroup favoritism in cooperation: A meta-analysis. *Psychological Bulletin*, No Pagination Specified. doi:10.1037/a0037737.

Bartels, D. M., Bauman, C. W., Cushman, F. A., Pizarro, D. A., & McGraw, A. P. (2015). Moral judgment and decision making. In G. Keren & G. Wu (Eds.), *Blackwell reader of judgment and decision making*. Malden, MA: Blackwell.

Becker, L. C. (1977). *Property rights: Philosophic foundations*. London, Henley and Boston: Routledge & Kegan Paul.

Berger, P. L. (1970). On the obsolescence of the concept of honor. *European Journal of Sociology, 11*, 338–347.

Berger, P. L., & Luckmann, T. (1966). *The social construction of reality*. London: Allen Lane The Penguin Press.

Binmore, K. (1994a). *Game theory and the social contract, Vol 1: Playing fair?* (Vol. I). Cambridge, MA: MIT Press.

shackelf@oakland.edu

Binmore, K. (1994b). *Game theory and the social contract, Vol 2: Just playing* (Vol. II). Cambridge, MA: MIT Press.

Bissonnette, A., Perry, S., Barrett, L., Mitani, J. C., Flinn, M., Gavrilets, S., et al. (2015). Coalitions in theory and reality: A review of pertinent variables and processes. *Behaviour, 152*(1), 1–56. doi:10.1163/1568539X-00003241.

Boesch, C. (1999). *Hierarchy in the forest*. Cambridge, MA: Harvard University Press.

Boinski, S., & Garber, P. A. (Eds.). (2000). *On the move: How and why animals travel in groups*. Chicago: Chicago University Press.

Boone, J. (1992). Competition, cooperation and the development of social hierarchies. In E. A. Smith & B. Winterhalder (Eds.), *Ecology, evolution and social behavior* (pp. 301–337). New York: Aldine de Gruyter.

Boos, M., Kolbe, M., Kappeler, P. M., & Ellwart, T. (Eds.). (2011). *Coordination in human and primate groups*. Heidelberg, Germany: Springer.

Bouzouggar, A., Barton, N., Vanhaeren, M., d'Errico, F., Collcutt, S., Higham, T., et al. (2007). 82,000-year-old shell beads from North Africa and implications for the origins of modern human behavior. *Proceedings of the National Academy of Sciences of the United States of America, 104*(24), 9964–9969. doi:10.1073/pnas.0703877104.

Boyd, R., Richerson, P. J., & Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 10918–10925. doi:10.1073/pnas.1100290108.

Brams, S. J., & Taylor, A. D. (1996). *Fair division: From cake-cutting to dispute resolution*. Cambridge, UK: Cambridge University Press.

Brosnan, S. F. (2013). Justice- and fairness-related behaviors in nonhuman primates. *Proceedings of the National Academy of Sciences of the United States of America, 110*(Suppl. 2), 10416–10423. doi:10.1073/pnas.1301194110.

Bshary, R., & Grutter, A. S. (2006). Image scoring and cooperation in a cleaner fish mutualism. *Nature, 441*(7096), 975–978. http://www.nature.com/nature/journal/v441/n7096/suppinfo/nature04755_S1.html.

Burnham, T. C. (2007). High-testosterone men reject low ultimatum game offers. *Proceedings of the Royal Society B: Biological Sciences, 274*(1623), 2327–2330.

Butler, J. (1856). *Fifteen sermons preached at the Rolls Chapel*. London: Longman.

Carter, G. G., & Wilkinson, G. S. (2013). Food sharing in vampire bats: reciprocal help predicts donations more than relatedness or harassment. *Proceedings of the Royal Society B: Biological Sciences, 280*(1753). doi:10.1098/rspb.2012.2573.

Casebeer, W. (2003). *Natural ethical facts: Evolution, connectionism, and moral cognition*. Cambridge, MA: MIT Press.

Centre, P. R. (2014). *Worldwide, many see belief in god as essential to morality*.

Chagnon, N. A., & Bugos, P. E. (1979). Kin selection and conflict: An analysis of a Yanomamö ax fight. In N. A. Chagnon & W. Irons (Eds.), *Evolutionary biology and human social behaviour: An anthropological perspective*. North Scituate, MA: Duxbury Press.

Chapais, B. (2014). Complex kinship patterns as evolutionary constructions, and the origins of sociocultural universals. *Current Anthropology, 55*(6), 751–783. doi:10.1086/678972.

Chilton, B. D., & Neusner, J. (Eds.). (2009). *The golden rule: The ethics of reciprocity in world religions*. Bloomsbury Academic.

Cicero, M. T. (1971). *On the good life*. London: Penguin Classics.

Clutton-Brock, T. H. (1991). *The evolution of parental care*. Princeton, NJ: Princeton University Press.

Clutton-Brock, T. H. (2009). Cooperation between non-kin in animal societies. *Nature, 462*, 51–57.

Confucius. (1994). A single word. In P. Singer (Ed.), *Ethics* (p. 76). Oxford, UK: Oxford University Press.

Connor, R. C. (1995). The benefits of mutualism: A conceptual framework. *Biological Reviews, 70*(3), 427–457.

shackelf@oakland.edu

Cosmides, L., & Tooby, J. (2005). Neurocognitive adaptations designed for social exchange. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 584–627). New York: Wiley.
Cowan, R. (2007). *For the glory of Rome: A history of warriors and warfare*. London: Greenhill Books.
Curry, O. S. (2005). *Morality as natural history: An adaptationist account of ethics*. PhD, London School of Economics, London. Retrieved from http://etheses.lse.ac.uk/2/
Curry, O. S. (2007). The conflict-resolution theory of virtue. In W. P. Sinnott-Armstrong (Ed.), *Moral psychology* (Vol. I, pp. 251–261). Cambridge, MA: MIT Press.
Curry, O. S., & Jones Chesters, M. (2012). 'Put yourself in the other fellow's shoes': The role of 'theory of mind' in solving coordination problems. *Journal of Cognition and Culture, 12*, 147–159.
Curry, O. S., Price, M. E., & Price, J. G. (2008). Patience is a virtue: Cooperative people have lower discount rates. *Personality and Individual Differences, 44*, 778–783.
Daly, M., & Wilson, M. (1988). *Homicide*. New York: Aldine de Gruyter.
Daly, M., & Wilson, M. (1996). Violence against stepchildren. *Current Directions in Psychological Science, 5*, 77–81.
Darwin, C. (1871). *The Descent of Man and Selection in Relation to Sex*. London, John Murray.
Darwin, C. (1872/1998). *The expression of the emotions in man and animals* (3rd ed.). London: John Murray/HarperCollins.
Dawkins, R. (1976/2006). *The selfish gene* (3rd ed.). Oxford, UK: Oxford University Press.
Dawkins, R. (1979). Twelve misunderstandings of kin selection. *Zeitschrift für Tierpsychologie, 51*(2), 184–200. doi:10.1111/j.1439-0310.1979.tb00682.x.
Dawkins, R. (1998). *Unweaving the rainbow: Science, delusion and the appetite for wonder*. London: Penguin Books.
de Kwaadsteneit, E. W., & van Dijk, E. (2010). Social status as a cue for tacit coordination. *Journal of Experimental Social Psychology, 46*, 515–524.
Deaner, R. O., & Smith, B. A. (2012). Sex differences in sports across 50 societies. *Cross-Cultural Research*. doi:10.1177/1069397112463687.
Downie, R. S. (1965). Forgiveness. *Philosophical Quarterly, 15*(59), 128–134.
Dugatkin, L. A. (1997). *Cooperation among animals: An evolutionary perspective*. New York: Oxford University Press.
Edel, M., & Edel, A. (1959/1968). *Anthropology and ethics: The quest for moral understanding*. Cleveland, OH: Case Western Reserve University Press.
Edelman, M. S., & Omark, D. R. (1973). Dominance hierarchies in young children. *Social Science Information, 12*(1), 103–110. doi:10.1177/053901847301200105.
Emmons, R. A. (Ed.). (2004). *The psychology of gratitude*. Oxford, UK: OUP.
Euler, H. A., & Weitzel, B. (1996). Discriminative grandparental solicitude as reproductive strategy. *Human Nature, 7*, 39–59.
Feinberg, J. (1973). Some conjectures about the concept of respect. *Journal of Social Philosophy, 4*(2), 1–3. doi:10.1111/j.1467-9833.1973.tb00163.x.
Fiddick, L., Cummins, D., Janicki, M., Lee, S., & Erlich, N. (2013). A cross-cultural study of noblesse oblige in economic decision-making. *Human Nature, 24*(3), 318–335. doi:10.1007/s12110-013-9169-9.
Fiske, A. P. (1992). The four elementary forms of sociality – Framework for a unified theory of social-relations. *Psychological Review, 99*(4), 689–723. doi:10.1037//0033-295X.99.4.689.
Friedman, O., & Neary, K. R. (2008). Determining who owns what: Do children infer ownership from first possession? *Cognition, 107*(3), 829–849. http://dx.doi.org/10.1016/j.cognition.2007.12.002.
Fukuyama, F. (1996). *Trust: The social virtues and the creation of prosperity*. London: Penguin Books.
Gaulin, S. J. C., & Schlegel, A. (1980). Paternal confidence and paternal investment: A cross cultural test of a sociobiological hypothesis. *Ethology and Sociobiology, 1*(4), 301–309. doi:10.1016/0162-3095(80)90015-1.

shackelf@oakland.edu

Gavrilets, S., Duenez-Guzman, E. A., & Vose, M. D. (2008). Dynamics of alliance formation and the egalitarian revolution. *PLoS One, 3*(10), e3293.
Geary, D. C. (2000). Evolution and proximate expression of human paternal investment. *Psychological Bulletin, 126*, 55–77.
Gert, B. (2013). Loyalty and morality. In S. Levinson, J. Parker, & P. Woodruff (Eds.), *Nomos* (Vol. LIV, pp. 3–21). New York & London: New York University Press.
Gibbard, A. (1990a). Norms, discussion, and ritual: Evolutionary puzzles. *Ethics, 100*(July), 787–802.
Gibbard, A. (1990b). *Wise choices, apt feelings*. Oxford, UK: Clarendon Press.
Gintis, H. (2007). The evolution of private property. *Journal of Economic Behavior & Organization, 64*(1), 1–16.
Gintis, H., Smith, E. A., & Bowles, S. (2001). Costly signaling and cooperation. *Journal of Theoretical Biology, 213*, 103–119. doi:10.1006/jtbi.2001.2406.
Godfray, H. C. J. (1992). The evolution of forgiveness. *Nature, 355*, 206–207.
Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology, 101*(2), 366–385. doi:10.1037/A0021847.
Greene, J. D. (2015). The rise of moral cognition. *Cognition, 135*, 39–42. http://dx.doi.org/10.1016/j.cognition.2014.11.018.
Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization, 3*(4), 367–388.
Haidt, J. (2012, June 18). To see group-selected traits, look at groupishness during intergroup competition. *Edge*.
Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus, 133*(4), 55–66.
Haidt, J., & Joseph, C. (2011). How moral foundations theory succeeded in building on sand: A response to Suhler and Churchland. *Journal of Cognitive Neuroscience, 23*(9), 2117–2122.
Haidt, J., & Kesebir, S. (2010). Morality. In S. Fiske, G. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (5th ed., pp. 797–832). Hoboken, NJ: Wiley.
Hamilton, W. D. (1964). The genetical evolution of social behaviour. *Journal of Theoretical Biology, 7*, 1–16, 17–52. doi:10.1016/0022-5193(64)90039-6.
Harbaugh, W. T., Krause, K., Liday, S. G., & Vesterlund, L. (2002). Trust in children. In E. Ostrom & J. Walker (Eds.), *Trust, reciprocity and gains from association: Interdisciplinary lessons from experimental research* (pp. 302–322). New York: Russell Sage Foundation.
Harcourt, A., & de Waal, F. B. M. (Eds.). (1992). *Coalitions and alliances in humans and other animals*. Oxford: Oxford University Press.
Hardy, C. L., & Van Vugt, M. (2006). Nice guys finish first: The competitive altruism hypothesis. *Personality and Social Psychology Bulletin, 32*(10), 1402–1413. doi:10.1177/0146167206291006.
Hardy, C. W., & Briffa, M. (Eds.). (2013). *Animal contests*. Cambridge, UK: Cambridge University Press.
Hauser, M. (2001). *Wild minds: What animals really think*. London: Penguin.
Hausman, D. M., & McPherson, M. S. (1996). *Economic analysis and moral philosophy*. Cambridge, UK: Cambridge University Press.
Hawkes, K. (1991). Showing off: Tests of another hypothesis about men's foraging goals. *Ethology and Sociobiology, 12*(1), 29–54.
Hawkes, K., O'Connell, J. F., & Blurton Jones, N. G. (2001). Hadza meat sharing. *Evolution and Human Behavior, 22*(2), 113–142.
Hazlitt, H. (1964). *The foundations of morality*. Princeton, NJ: D. Van Nostrand Company Inc.
Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., et al. (2005). 'Economic Man' in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences, 28*(6), 795–855.
Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Beyond WEIRD: Towards a broad-based behavioral science. *Behavioral and Brain Sciences, 33*(2–3), 111–135. doi:10.1017/S0140525x10000725.

shackelf@oakland.edu

Hepper, P. G. (Ed.). (1991). *Kin recognition*. Cambridge, UK: Cambridge University Press.
Herskovits, M. J. (1952). *Economic anthropology: A study in comparative economics*. New York: Afred A. Knopf.
Hobbes, T. (1651/1958). *Leviathan*. New York: Macmillan.
Hume, D. (1739/1985). *A treatise of human nature*. London: Penguin Classics.
Hume, D. (1757/1889). *The natural history of religion*. London: Freethought.
Huntingdon, F. A., & Turner, A. K. (1987). *Animal conflict*. London & New York: Chapman and Hall.
Jackson, F., Pettit, P., & Smith, M. (2004). *Mind, morality, and explanation: Selected collaborations*. Oxford, UK: OUP.
Jaeggi, A. V., & Gurven, M. (2013). Reciprocity explains food sharing in humans and other primates independent of kin selection and tolerated scrounging: A phylogenetic meta-analysis. *Proceedings of the Royal Society of London B: Biological Sciences, 280*(1768).
Johnson, D. D. P., & Toft, M. D. (2014). Grounds for war: The evolution of territorial conflict. *International Security, 38*(3), 7–38. doi:10.1162/ISEC_a_00149.
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263–291.
Kocher, M. G., Cherry, T., Kroll, S., Netzer, R. J., & Sutter, M. (2008). Conditional cooperation on three continents. *Economics Letters, 101*, 175–178. doi:10.1016/j.econlet.2008.07.015.
Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature, 435*, 673–676.
Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences of the United States of America, 98*(26), 15387–15392.
Ladd, J. (Ed.). (1985). *Ethical relativism*. Lanham, MD: University Press of America.
Lehmann, L., & Keller, L. (2006). The evolution of cooperation and altruism – A general framework and a classification of models. *Journal of Evolutionary Biology, 19*(5), 1365–1376. doi:10.1111/j.1420-9101.2006.01119.x.
Levinson, S., Parker, J., & Woodruff, P. (Eds.). (2013). *Loyalty* (Vol. LIV). New York & London: New York University Press.
Lewis, D. K. (1969). *Convention: A philosophical study*. Cambridge, MA: Harvard University Press.
Lieberman, D., Tooby, J., & Cosmides, L. (2003). Does morality have a biological basis? An empirical test of the factors governing moral sentiments relating to incest. *Proceedings of the Royal Society of London Series B-Biological Sciences, 270*(1517), 819–826. doi:10.1098/Rspb.2002.2290.
Lieberman, D., Tooby, J., & Cosmides, L. (2007). The architecture of human kin detection. *Nature, 445*(7129), 727–731. doi:10.1038/Nature05510.
Locke, J. (2000). Labour as the basis of property. In M. Rosen, W. Wolff, & C. McKinnon (Eds.), *Political thought* (p. 73). Oxford, UK: Oxford University Press.
MacIntyre, A. C. (1981a). *After virtue*. London: Gerald Duckworth & Co. Ltd.
MacIntyre, A. C. (1981b). The nature of the virtues. *Hastings Center Report, 11*(2), 27–34.
Mackie, J. L. (1977). *Ethics: Inventing right and wrong*. London: Penguin.
Maynard Smith, J. (1982). *Evolution and the theory of games*. Cambridge, UK: Cambridge University Press.
Maynard Smith, J., & Price, G. R. (1973). The logic of animal conflict. *Nature, 246*, 15–18.
Maynard Smith, J., & Szathmáry, E. (1995). *The major transitions in evolution*. Oxford, UK: Oxford University Press.
Mazur, A. (2005). *Biosociology of dominance and deference*. Lanham, MD: Rowan & Littlefield.
McCullough, M. E., Kimeldorf, M. B., & Cohen, A. D. (2008). An adaptation for altruism? The social causes, social effects, and social evolution of gratitude. *Current Directions in Psychological Science, 17*(4), 281–285. doi:10.1111/j.1467-8721.2008.00590.x.

shackelf@oakland.edu

McCullough, M. E., Kurzban, R., & Tabak, B. A. (2013). Cognitive systems for revenge and forgiveness. *Behavioral and Brain Sciences, 36*(01), 1–15. doi:10.1017/S0140525X11002160.
McKay, R., & Whitehouse, H. (2014). Religion and morality. *Psychological Bulletin*.
Messick, D. M. (1993). Equality as a decision rule. In B. Mellers & J. Baron (Eds.), *Psychological perspectives on justice* (pp. 11–31). Cambridge, UK: Cambridge University Press.
Mill, J. S., & Bentham, J. (1987). *Utilitarianism and other essays*. London: Penguin.
Miller, G. F. (2000). *The mating mind*. London: William Heinemann.
Mitani, J. C. (2009). Cooperation and competition in chimpanzees: Current understanding and future challenges. *Evolutionary Anthropology, 18*(5), 215–227. doi:10.1002/Evan.20229.
Moore, G. E. (1903). *Principia ethica*. Cambridge, UK: Cambridge University Press.
Nagel, T. (1991). *Mortal questions*. London: Canto.
Nash, J. (1950). The bargaining problem. *Econometrica, 18*, 155–162.
No Title. (2001, March 29). *The Economist*.
Noddings, N. (1978). *Caring: A feminine approach to ethics and education*. Berkeley, CA: University of California Press.
Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science, 314*(5805), 1560–1563. doi:10.1126/science.1133755.
Nunn, C. L., & Lewis, R. J. (2001). Cooperation and collective action in animal behaviour. *Economics in nature: Social dilemmas, mate choice and biological markets* (pp. 42–66). Cambridge, UK: CUP.
Oates, K., & Wilson, M. (2002). Nominal kinship cues facilitate altruism. *Proceedings: Biological Sciences, 269*(1487), 105–109.
Ohtsubo, Y., & Watanabe, E. (2009). Do sincere apologies need to be costly? Test of a costly signaling model of apology. *Evolution and Human Behavior, 30*(2), 114–123.
Ostrom, L., & Walker, J. (Eds.). (2002). *Trust and reciprocity: Interdisciplinary lessons from experimental research*. New York: Russell Sage Foundation.
Pennock, J. R., & Chapman, J. W. (Eds.). (1979). *Compromise in ethics, law and politics* (Vol. XXI). New York: New York University Press.
Pennock, J. R., & Chapman, J. W. (Eds.). (1980). *Property* (Vol. XXII). New York: New York University Press.
Pinker, S. (2010). The cognitive niche: Coevolution of intelligence, sociality, and language. *Proceedings of the National Academy of Sciences of the United States of America, 107*, 8993–8999. doi:10.1073/pnas.0914630107.
Pinker, S. (1997). *How the mind works*. New York: W W Norton.
Platek, S. M., Critton, S. R., Burch, R. L., Frederick, D. A., Myers, T. E., & Gallup, G. G., Jr. (2003). How much resemblance is enough? Sex difference in reactions to resemblance, but not the ability to detect resemblance. *Evolution and Human Behavior, 24*(3), 81–87.
Plato. (1974). *The republic*. London: Penguin Books.
Popper, K. R. (1945). *The open society and its enemies*. London: Routledge.
Preuschoft, S., & van Schaik, C. P. (2000). Dominance and communication: Conflict management in various social settings. In F. Aureli & F. B. M. de Waal (Eds.), *Natural conflict resolution* (pp. 77–105). Berkeley, CA: University of California Press.
Price, M. E., Cosmides, L., & Tooby, J. (2002). Punitive sentiment as an anti-free rider psychological device. *Evolution and Human Behavior, 23*(3), 203–231. doi:10.1016/j.evolhumbehav.2004.08.009.
Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review, 118*(1), 57–75. doi:10.1037/a0021867.
Rawls, J. (1958). Justice as fairness. *The Philosophical Review, 67*(2), 164–194.
Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.
Richards, N. (1988). Forgiveness. *Ethics, 99*(1), 77–97.
Riechert, S. E. (1998). Game theory and animal contests. In L. A. Dugatkin & H. K. Reeve (Eds.), *Game theory and animal behavior* (pp. 64–93). Oxford, UK: Oxford University Press.

shackelf@oakland.edu

Rose, C. M. (1985). Possession as the origin of property. *University of Chicago Law Review, 52*, 73–88.
Ross, H., & Friedman, O. (Eds.). (2011). *Origins of ownership of property: New directions for child and adolescent development*. Hoboken, NJ: John Wiley & Sons.
Royce, J. (1908). *The philosophy of loyalty*. New York: Macmillan.
Royle, N. J., Smiseth, P. T., & Kölliker, M. (Eds.). (2012). *The evolution of parental care*. Oxford, UK: OUP.
Rubin, P. H. (2000). Hierarchy. *Human Nature, 11*(3), 259–279.
Ruddick, S. (1980). Maternal thinking. *Feminist Studies, 6*, 342–367.
Russell, B. (1927). *Philosophy*. New York: Norton.
Sachs, J. L., Mueller, U. G., Wilcox, T. P., & Bull, J. J. (2004). The evolution of cooperation. *The Quarterly Review of Biology, 79*(2), 135–160. doi:10.1086/383541.
Schelling, T. C. (1960). *The strategy of conflict*. Cambridge, MA: Harvard University Press.
Sell, A., Bryant, G. A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., et al. (2010). Adaptations in humans for assessing physical strength from the voice. *Proceedings of the Royal Society B-Biological Sciences, 277*(1699), 3509–3518. doi:10.1098/Rspb.2010.0769.
Sell, A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., & Gurven, M. (2009). Human adaptations for the visual assessment of strength and fighting ability from the body and face. *Proceedings of the Royal Society B: Biological Sciences, 276*(1656), 575–584. doi:10.1098/Rspb.2008.1177.
Sherif, M., Harvey, O. J., White, B. J., Hood, W. R., & Sherif, C. W. (1954/1961). *Intergroup conflict and cooperation: The Robbers cave experiment*. Norman, OH: University of Oklahoma Book Exchange.
Shultz, S., & Dunbar, R. I. M. (2007). The evolution of the social brain: Anthropoid primates contrast with other vertebrates. *Proceedings of the Royal Society of London B: Biological Sciences, 274*(1624), 2429–2436.
Shultz, S., Opie, C., & Atkinson, Q. D. (2011). Stepwise evolution of stable sociality in primates. *Nature, 479*(7372), 219–222. http://www.nature.com/nature/journal/v479/n7372/abs/nature10601.html#supplementary-information.
Shweder, R. A., Much, N., Park, L., & Mahapatra, M. M. (1997/2003). The 'Big Three' of morality (autonomy, community, divinity) and the 'Big Three' explanations of suffering. In A. Brandt & P. Rozin (Eds.), *Morality and health*. New York: Routledge.
Sinnott-Armstrong, W., & Wheatley, T. (2013). Are moral judgments unified? *Philosophical Psychology, 27*(4), 451–474. doi:10.1080/09515089.2012.736075.
Skyrms, B. (1996). *Evolution of the social contract*. Cambridge, UK: Cambridge University Press.
Skyrms, B. (2004). *The stag hunt and the evolution of social structure*. Cambridge, UK: Cambridge University Press.
Smith, M. S., Kish, B. J., & Crawford, C. B. (1987). Inheritance of wealth as human kin investment. *Ethology and Sociobiology, 8*(3), 171–182. http://dx.doi.org/10.1016/0162-3095(87)90042-2.
Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103*(2684), 677–680.
Strassmann, J. E., & Queller, D. C. (2014). Privatization and property in biology. *Animal Behaviour*. http://dx.doi.org/10.1016/j.anbehav.2014.02.011.
Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American, 223*(5), 96–102.
Thomas, K. A., DeScioli, P., Haque, O. S., & Pinker, S. (2014). The psychology of coordination and common knowledge. *Journal of Personality and Social Psychology, 107*(4), 657–676. doi:10.1037/a0037037.
Thornhill, N. W. (1991). An evolutionary analysis of rules regulating human inbreeding and marriage. *Behavioral and Brain Sciences, 14*(02), 247–261. doi:10.1017/S0140525X00066449.
Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28*(5), 675.
Tomasello, M., & Vaish, A. (2013). Origins of human cooperation and morality. *Annual Review of Psychology, 64*(1), 231–255. doi:10.1146/annurev-psych-113011-143812.
Tooby, J., & Cosmides, L. (1996). Friendship and the banker's paradox: Other pathways to the evolution of adaptations for altruism. In W. G. Runciman, J. Maynard Smith, & R. I. M. Dunbar

shackelf@oakland.edu

(Eds.), *Evolution of social behaviour patterns in primates and man* (pp. 119–143). Oxford, UK: British Academy/Oxford University Press.
Tooby, J., & DeVore, I. (1987). The reconstruction of hominid behavioral evolution through strategic modeling. In W. G. Kinzey (Ed.), *The evolution of human behavior: Primate models* (pp. 183–237). Albany, NY: SUNY Press.
Tooby, J., & Cosmides, L. (2010). *Groups in mind: The coalitional roots of war and morality* (pp. 191–234).
Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46*(1), 35–57. doi:10.1086/406755.
Van Vugt, M., Hogan, R., & Kaiser, R. B. (2008). Leadership, followership, and evolution: Some lessons from the past. *American Psychologist, 63*(3), 182–196.
Von Neumann, J., & Morgenstern, O. (1944). *The theory of games and economic behavior*. Princeton, NJ: Princeton University Press.
Watkins, C. D., Fraccaro, P. J., Smith, F. G., Vukovic, J., Feinberg, D. R., DeBruine, L. M., et al. (2010). Taller men are less sensitive to cues of dominance in other men. *Behavioral Ecology*. doi:10.1093/beheco/arq091.
West, S. A., Griffin, A. S., & Gardner, A. (2007). Evolutionary explanations for cooperation. *Current Biology, 17*(16), R661–R672. doi:10.1016/j.cub.2007.06.004.
Westermarck, E. A. (1906). *The origin and development of the moral ideas*. London: Macmillan.
Whiten, A. (1996). When does smart behaviour-reading become mind-reading? In P. Carruthers & P. K. Smith (Eds.), *Theories of theories of mind*. Cambridge, UK: CUP.
Wong, D. (1984). *Moral relativity*. Berkeley, CA: UC California Press.
Wrangham, R. (1999). Evolution of coalitionary killing. *Yearbook of Physical Anthropology, 42*, 1–30.
Young, L., Camprodon, J. A., Hauser, M., Pascual-Leone, A., & Saxe, R. (2010). Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proceedings of the National Academy of Sciences of the United States of America, 107*(15), 6753–6758.
Zahavi, A., & Zahavi, A. (1997). *The handicap principle: A missing piece of Darwin's puzzle*. Oxford, UK: Oxford University Press.
Zak, P. J., Kurzban, R., Ahmadi, S., Swerdloff, R. S., Park, J., Efremidze, L., et al. (2009). Testosterone administration decreases generosity in the ultimatum game. *PLoS One, 4*(12), e8330.

shackelf@oakland.edu

An Evolving and Developing Field of Study: Prosocial Morality from a Biological, Cultural, and Developmental Perspective

Gustavo Carlo, Christa Christ, Deborah Laible, and Zehra Gulseven

Scholars have long debated the aspects of our species that define us as distinct from other species. Early theorists suggested that our ability to engage in self-reflection and abstract thinking is uniquely human, others assert that our capacity for empathy is characteristic of humans, and still others propose that our ability to manipulate and master our environment is uniquely human (see Bandura, 1986; Chudek & Henrich, 2011). In recent years, however, researchers have shown the capacity for these skills in nonhuman primates, although humans may exhibit more mature abilities (Subiaul, Vonk, Okamoto-Barth, & Barth, 2008; Warneken & Tomasello, 2009). Although debates regarding these issues will continue, some scientists have suggested that aspects of morality are uniquely human (Chudek & Henrich, 2011; Jensen & Silk, 2014; see Brosnan, 2010). In part, this belief is grounded in the notion that morality is a co-constructive notion that can only be understood in social and cultural contexts and involves high-level sociocognitive skills (Chudek & Henrich, 2011; Kurtines, Mayock, Pollard, Lanza, & Carlo, 1991). Moreover, morality involves the integration of sociocognitive skills and emotion-related processes that can manifest as value-based social actions and behaviors. These social cognitions, social emotions, and behaviors change across time and differ

G. Carlo (✉)
Department of Human Development and Family Science,
University of Missouri, 410 Gentry Hall, Columbia, MO, 65211, USA
e-mail: carlog@missouri.edu

C. Christ
University of Nebraska-Lincoln, Lincoln, USA

D. Laible
Lehigh University, Bethlehem, USA

Z. Gulseven
University of Missouri, Columbia, MO, USA

© Springer International Publishing Switzerland 2016

T.K. Shackelford, R.D. Hansen (eds.), *The Evolution of Morality*,
Evolutionary Psychology, DOI 10.1007/978-3-319-19671-8_3

shackelf@oakland.edu

across individuals and social and cultural contexts. However, most traditional theories of human moral development do not integrate these biological, cultural, and developmental elements. The present chapter presents a brief review of our current understanding of human moral development with a particular focus on childhood and adolescence. Moreover, our chapter focuses on our understanding of biological and sociocultural proximal causes and correlates of morality.

There are several theories as to why morality evolved in humans. Developing conceptions of morality provide a framework for common and differing ideas, concepts, and issues surrounding what is deemed acceptable or unacceptable. These evaluative notions then support and organize interactions among individuals and groups of people. As such, morality is a concept that has undergone changes over time and across contexts. As Campbell (1975) noted, evolutionary processes occur both at the biological level and societal level. Although morality has both biological and environmental foundations, human understanding of morality occurs in social context. Thus, human moral concepts have undergone centuries of social evolutionary changes.

One common evolutionary framework for understanding morality is kin selection (Sober & Wilson, 1998; Wilson, 1978). Kin selection theory posits that humans are motivated to act more generously toward and to benefit others who are genetically related (de Waal, 2008). In other words, individuals will be more likely to act in ways that benefit others when those others are genetically related to those individuals. These scholars further espouse that the benefits can be measured by the concept of "inclusive fitness," defined as the shared genetic material with kin. However, this theory was wrought with concerns regarding its ability to explain cooperation among individuals who are not closely genetically related.

Trivers (1971) sought to address the limitations of kin selection theory by proposing the concept of reciprocal altruism. This asserts that individuals will cooperate with nonrelated others if they anticipate benefits from doing so. This notion implies that all behaviors are driven by self-interest at the genetic level. There are other variations of reciprocal altruism including indirect reciprocity (i.e., which works at the level of reputation at the social or community level) and a strong reciprocity version (i.e., posits cooperation that does not emphasize self-interests but rather behavior manifests out of intrinsic "rightness") (Alexander, 1987). In essence, most evolutionary theories share common assumptions including the notion that moral behaviors have adaptive value (with respect to proliferation of genes), are based on self-interests (but see Sober & Wilson, 1998; Warneken & Tomasello, 2009), and have proximal instrumental value (Hawley, 2014). Moreover, many evolutionary theorists focus on explaining moral behaviors at a distal level of analysis and infer (rather than directly assess) underlying explanatory processes from observed behaviors, and few directly examine microlevel, proximal processes (e.g., psychological, neural, physiological).

In the present chapter, we present a conception of morality that highlights relatively microlevel, proximal psychological processes associated with human morality. We assume a social developmental perspective that incorporates biological, psychological, and cultural processes (see Fig. 1; based on Carlo & de Guzman,

shackelf@oakland.edu

2009). Briefly, the model depicts a number of antecedent variables that include characteristics of the child (e.g., genetic predispositions, sociocognitive abilities, self-regulation, temperament), peers (e.g., characteristics of the affiliative group), and family (e.g., parenting practices, household structure), school context (e.g., racial/ethnic composition, school engagement), life events (e.g., exposure to major life stressors), and receiving community characteristics (e.g., socioeconomic status, employment opportunities, violence rates). However, the effects of these antecedent variables depend on the child's perceptions via a variety of sociocognitive and socioemotive tendencies, including appraisals, moral reasoning, values, empathy, and racial/ethnic identity. Furthermore, these sociocognitive and socioemotive traits may influence the ethnic minority child's perceptions of culture-related stressors, which in turn predict prosocial and antisocial outcomes. Interrelations within each set of variables, interaction effects, and bidirectional relations are not depicted in the model. The model also emphasizes moral development in youth from different ethnic groups. Although morality cuts across both prosocial and antisocial dimensions, our work has mostly focused on prosocial moral development and we highlight this work. We begin with some definitional issues.

Theoretical and Definitional Issues in Morality

Morality refers to a code of conduct that individuals and groups adopt as normative to govern themselves. Morality can refer to normative standards that define goodness and rightness, whereas immorality often refers to behaviors deemed wrong or unacceptable. There are various other definitions of morality—some that specify principles and others that specify behaviors (Kohlberg, 1969; Mischel & Mischel, 1976). Our perspective posits that the human co-constructivist nature of morality necessitates an understanding of morality at the behavioral, cognitive, and emotive levels. Furthermore, the foundational processes of moral behaviors, cognitions, and emotions are biological (including genetics, neurotransmitters, neural, physiological), as well as interpersonal and social-communal. Although the cognitive and emotive aspects of morality are subject to societal evaluations, the aspect often most heavily evaluative is overt behavior. Therefore, we begin with a discussion of moral behaviors.

Behavioral Expressions

Moral behaviors can be classified into two broad categories: prosocial and antisocial. Prosocial behaviors are actions that benefit others and include behaviors that are selfishly or selflessly motivated (Carlo, 2014). Examples of prosocial behaviors include donating money or goods, sharing resources, volunteerism, comforting others, and helping when asked to do so. The conditions for prosocial behaviors

shackelf@oakland.edu

also vary considerably, including helping under personal duress, helping in front of an audience, spontaneously helping others, and helping in emergency situations. One interesting subset of prosocial behaviors is altruistic behaviors, often considered high-cost or high-risk actions enacted with little or no concern for self-rewards (Batson, 1998; Eisenberg, Fabes, & Spinrad, 2006; Staub, 1978). There is extensive scholarly debate and research about these forms of prosocial behaviors (Sober & Wilson, 1998; Warneken & Tomasello, 2009). Although space limits an extensive review of theory and research on altruistic behaviors, such actions are the subject of much interest and debate across many scholarly disciplines. Our own research focuses on prosocial and altruistic behaviors; thus, in this chapter we focus on prosocial actions.

Despite our focus on prosocial behaviors in our subsequent discussion, it is useful to contrast these actions with antisocial behaviors. Antisocial behaviors are defined as actions that contradict social norms, laws, and rules. These behaviors include aggression (i.e., actions that harm others), delinquency (e.g., actions that violate social laws and rules), and criminality (e.g., vandalism, burglary). Much of the early research in morality and moral development emphasized the study of antisocial moral actions (e.g., cheating, lying, stealing, aggression), and therefore, many early theories were developed primarily on the basis of our understanding of antisocial morality (Eisenberg, 1986; Staub, 1978). The more recent focus on prosocial morality provides an important theoretical and methodological contrast to that earlier work. However, the studies of both aspects of morality will likely produce richer conceptions of morality.

Although much of interest is in prosocial and antisocial actions, our developmental focus necessitates that we identify candidate processes that can account for changes in prosocial and antisocial behaviors across time. The interest in developmental trends across the life span led early theorists to examine sociocognitive processes that demonstrate clear changes across childhood and adolescence (Kohlberg, 1969; Piaget, 1965). Among the sociocognitive processes most often associated with prosocial and antisocial behaviors were moral reasoning and perspective taking.

Cognitive-Based Moral Processes

Moral Reasoning

Early work in moral development followed in the tradition of Greek philosophers (e.g., Plato), Immanuel Kant, and Lawrence Kohlberg, which emphasized the central role of reasoning, rationality, and logic as the essential elements of morality (Kohlberg, 1969). For these scholars, morality is determined via rational thinking and reasoning regarding issues of justice, welfare, and fairness. Moreover, reasoning plays a powerful role in defining morality through the notions of "categorical

shackelf@oakland.edu

imperatives" and "deontic responsibility." Kant introduced the concept of the categorical imperative as a moral obligation or duty that stems from reason and is independent of circumstances or desires (Kant, 1755). Similarly, Kohlberg asserted that individuals who engage in high-level moral reasoning are motivated to act upon their reasoning as a result of a felt sense of duty and obligation (Kohlberg & Candee, 1984). Kohlberg delineated six stages of moral judgment that he posited to develop across childhood, adolescence, and adulthood. According to cognitive-developmental theorists, these transformations across time are primarily due to changes in perspective taking (i.e., understanding others' thoughts and emotions or social situations). Furthermore, exposure to different social perspectives induces cognitive disequilibrium, which results in accommodation and assimilation processes and new moral conceptions (Kohlberg, 1969).

For decades, the research on moral development was dominated by Kohlberg's strong stage reasoning perspective. This pioneering work resulted in the demonstrated age-related changes in moral understanding and explanations for the role of moral reasoning in moral decision-making and behaviors (Colby & Kohlberg, 1987). However, several scholars noted mixed findings, including the lack of a strong, reliable link between moral reasoning and behavior and the relative neglect of moral emotions and attention to culture-related processes (Rest, 1983; Wong, 2009). Despite these limitations, research on moral reasoning continues (e.g., Ongley et al., 2014) and newer conceptions and revisions to the early ideas have been proposed. For example, some work has focused on distinguishing among personal, social, and moral domains of reasoning (Turiel, 2006). This line of work presents evidence that some of the mixed previous findings likely resulted from confusion surrounding the definitions of morality. Another line of work that extends prior research on moral reasoning reflects the study of prosocial moral reasoning (Eisenberg, 1986). In this research, scholars focus on individuals' thinking about dilemmas when faced with opportunities to help others in the relative absence of laws, social rules, or guidelines. This latter body of work demonstrates distinct developmental trends in moral reasoning, evidence of compassion-related forms of reasoning, gender and culture-related differences in moral reasoning, and links between prosocial moral reasoning and prosocial behaviors (e.g., Carlo, Knight, McGinley, Goodvin, & Roesch, 2010; Carlo, Knight, McGinley, & Hayes, 2011; see Carlo, 2014; Eisenberg et al., 2006).

Values

Although the study of values has a long history in psychology (Allport, 1937), much of the early work related to morality has been relatively neglected. However, in recent years, there is a resurging interest in the role of values in understanding moral development (McAdams, 2009; Schwartz, 2010). Critics of the moral judgment approach noted inadequate accounts of moral motivation and moral content (Blasi, 1983). Furthermore, there was newfound interest in attention to culture-related

shackelf@oakland.edu

processes, moral socialization, and moral identity and motivation (Grusec & Goodnow, 1994; Hardy & Carlo, 2005; Knight, Bernal, & Carlo, 1995)—aspects that had been relatively neglected in much of the moral development research. Many theories of cultural socialization posit shared values and beliefs as core elements of culture. According to cultural scholars, parents transmit their culture-related values and beliefs to their offspring through culture-related practices and rituals, as well as through direct tuition (see Carlo & de Guzman, 2009, for a discussion of relevant theories). Children thus acquire cultural values and beliefs, which can provide standards and guidance for navigating moral situations. According to this perspective, age-related changes in value acquisition and endorsement can occur, at least partly, as a result of socialization experiences. Therefore, developmental trends in moral behaviors may be linked to developmental trends in values.

With regard to research on culture-related values and moral development, much of the sparse work has focused on US Latino/Latina samples. Knight, Carlo, and colleagues have conducted a series of studies aimed at examining the role of culture-related values on Latino/Latina youth prosocial development (see Knight & Carlo, 2012). Based on cultural socialization theories (Carlo & de Guzman, 2009; Knight et al., 1995), these researchers hypothesized that specific forms of prosocial behaviors among Latino/Latinas stem from ethnic socialization practices that promote specific traditional cultural values. Familism (i.e., support, obligation, and identity with family) is a traditional cultural value that is strongly endorsed by many Latino/Latinas. This value is espoused and nurtured in Latino/Latina children via socialization practices and experiences. Because the value entails an orientation toward the needs of others, respect for others, and practice in helping others, familism is deemed to foster prosocial behaviors (Knight & Carlo, 2012). In a series of longitudinal and cross-sectional studies, researchers have demonstrated significant relationships between familism and prosocial behaviors (Armenta, Knight, Carlo, & Jacobson, 2011; Calderon-Tena et al., 2011; Knight, Carlo, Basilio, & Jacobson, 2014; Streit et al., 2015). Recent studies show that other traditional Latino/Latina values (e.g., respect, traditional gender roles) also show significant relationships with prosocial behaviors (Davis, Carlo, & Knight, in press). Furthermore, mainstream US values (e.g., personal wealth, competition) are related to distinct forms of prosocial behaviors (Armenta et al., 2011). Finally, in a few studies of European American youth, investigators showed support for the socialization of moral values and reported significant relationships between moral values (e.g., kindness, honesty, fairness) and prosocial behaviors (Caprara & Steca, 2007; Hardy & Carlo, 2005; Hardy, Carlo, & Roesch, 2010; Padilla-Walker, 2007; Padilla-Walker & Carlo, 2007).

Overall, the growing evidence supports the notion that moral cognitions play an important role in predicting (prosocial) moral behaviors. One line of previous research findings suggests that these relations are task or situation specific (see Carlo, Mestre, Samper, Tur, & Armenta, 2010). That is, some moral situations pull for a relatively more prominent role of specific forms of moral cognitions, whereas other moral situations pull for a relatively more prominent role of moral emotions. For example, observing a person who is hurt and bleeding from a fall may not require much cognitive processing to respond prosocially. However, volunteering to

shackelf@oakland.edu

donate blood in response to a natural disaster or deciding to donate money for charity may require more cognitive processing. In both cases, moral cognitions and emotions are jointly contributing, but one set of processes may play a more prominent role in predicting behaviors than the other. Thus, careful consideration of the sociocognitive and socioemotive traits relevant to specific moral behaviors is important.

Emotion-Based Moral Processes

Empathy and Sympathy

The most prominent moral emotion that is conceptually and empirically linked to moral behaviors is empathy and its associated vicarious responses. Scholars have noted important distinctions among empathy, sympathy, and personal distress. Empathy refers to feeling the same as another as a result of another person's plight (Eisenberg et al., 2006; Hoffman, 2000). Empathy, thus, can be either positively valenced (e.g., joy) or negatively valenced (e.g., sadness). According to these scholars, empathy can sometimes lead to sympathy or personal distress responding. Sympathy is defined as feelings of sorrow or concern for another person. This emotion results from another person's plight but is always negatively valenced. Both empathy and sympathy have been most strongly linked to moral behaviors and both are also linked closely to perspective taking (i.e., cognitively understanding another person's thoughts, emotions, or social situation). Finally, personal distress is an emotion that is also closely linked with empathy but results in an aversive, physiologically arousing response that stems from another's plight. In contrast to empathy and sympathy, personal distress reflects an orientation to the self and one's own state and is not conceptually linked to helping others (unless helping results in reducing their personal distress). Moreover, the elements of empathic and sympathetic responding include the consideration of the needs of others, which can lead to altruistic (selfless-motivated) behaviors and mitigate harm and injury on others (e.g., aggressive and antisocial behaviors). Although most often empathy and sympathy are not adequately distinguished in research, evidence on the central role of empathy and sympathy in predicting prosocial and antisocial behaviors is generally corroborative (Carlo, 2006; Eisenberg et al., 2006).

Guilt and Shame

Other scholars have conceptually linked guilt and shame to moral behaviors (Hoffman, 2000; Kochanska, 1994; see Tangney & Dearing, 2002). Individuals who perceive themselves to have failed to meet their own moral standards are prone to

shackelf@oakland.edu

guilt. In contrast, shame stems from a perceived failure to meet the normative standards of a valued other. For example, a child who fails to help a distressed sibling may feel shame for failing to meet their parent's expectations or that child may feel guilty knowing that they failed to act on their own moral standards. Tangney (1996) and others (Hoffman, 2000; Kochanska, 1994) posit that guilt most often elicits higher-level moral responding, whereas shame most often mitigates moral responding. Sparse research generally supports a positive link between guilt and prosocial and moral outcomes (Ferguson, Stegge, Miller, & Olsen, 1999; Kochanska, Gross, Lin, & Nichols, 2002; Roberts, Strayer, & Denham, 2014). However, cross-cultural research and other recent research suggest that both guilt and shame can play a prominent role in predicting higher-level moral responding (Carlo, McGinley, Davis, & Streit, 2012; Fung, 1999; Laible, Eye, & Carlo, 2008; Menesini & Camodeca, 2008). Thus, more research is needed to better understand the roles of guilt and shame in moral development.

Individual Differences in Justice-Based and Compassion-Based Considerations

In a further effort to clarify the relationships among moral cognitions, moral emotions, and moral behaviors, Carlo and Davis (in press) elaborated on the distinction between justice-based and compassion-based moral domains (cf. Wong, 1984). The former entails consideration of what is just, fair, or right. These scholars further suggest that there may be individual differences in the extent to which some persons are inclined to invoke justice-based principles, values, or norms whereas other persons may be more likely to invoke compassion-based considerations of humaneness and suffering. Furthermore, some moral situations may be more pertinent to justice-based considerations (especially if there are explicit societal laws or rules), and other moral situations may be more pertinent to compassion-based considerations. In addition, some moral situations pull for a tension between justice-based and compassion-based considerations—such situations can partially explain sociomoral dilemmas that are challenging to resolve or reconcile. Although there may be exceptions, individuals who invoke justice-based considerations usually rely on moral cognitive processes (such as moral reasoning), whereas individuals who invoke compassion-based considerations usually rely on moral emotive processes (such as empathy or sympathy). Therefore, both moral cognitions and emotions play a role in predicting moral behaviors but the relative importance of each process may differ depending upon the situation (i.e., moral behavior opportunity) and individual differences in moral orientations. This perspective is consistent with the core assumption that morality is subject to change across time, individuals, and situations. Now that we have briefly outlined proximal mechanisms associated with moral development, we present a brief review of the biological underpinnings of these mechanisms.

shackelf@oakland.edu

Biological Basis of Proximal Mechanisms

Genes and Neurotransmitters

A number of researchers have examined the genetic and biological bases of moral cognitions, emotions, and behaviors in humans. Much of the early work in this area was based on twin longitudinal studies designed to ascertain the heritability of various traits and behaviors. Several studies have demonstrated that empathy and prosocial behaviors have a considerable heritable component (e.g., Emde et al., 1992; Knafo, Israel, & Ebstein, 2011; Knafo & Plomin, 2006; Matthews, Batson, Horn, & Rosenman, 1981). Other studies show empathic and prosocial tendencies relatively early in human life (Sommerville, Schmidt, Yun, & Burns, 2013; Zahn-Waxler, Radke-Yarrow, & King, 1979). Developmental researchers have also produced substantial evidence on the stability of prosocial tendencies across the life span (Eisenberg et al., 1999). Presumably some of the stability in prosocial traits is partly due to the heritable and temperamental basis of these traits (Eisenberg et al., 1999, 2006).

Perhaps the strongest evidence on the genetic and biological basis of morality stems from research focusing on neurotransmitters and hormones associated with moral traits. Although much of the research has focused on nonhuman species (mostly voles), recent years have yielded a rich number of studies that focus on humans. Several candidate neurotransmitters have been identified as particularly relevant to prosocial traits, including oxytocin, serotonin, and, to a somewhat lesser extent, dopamine and vasopressin.

Oxytocin, sometimes referred to as the "love hormone," has been conceptually and empirically linked to attachment, bonding, approach behaviors, sociability, and prosocial traits (Decety & Svetlova, 2012). However, there are several methodological approaches to the study of oxytocin and prosocial behaviors, which introduce some difficulties in reaching confident conclusions and making direct comparisons across study findings. Genetic candidate studies identify a specific oxytocin gene allele and examine whether the presence of that allele is associated with exhibited prosocial behaviors. Other researchers experimentally administer oxytocin to subjects (usually via intranasal administration), and across these studies, the measures of prosocial behaviors vary greatly (e.g., questionnaires, behavioral tasks, tasks that specify helping targets). Genetic candidate studies generally find support for the notion that oxytocin gene alleles are associated with prosocial behaviors (Christ, Carlo, & Stoltenberg, 2015; Walter et al., 2012), though some studies find such relationships are moderated by other variables (Ci, Wu, & Su, 2014). Studies that experimentally administered oxytocin to subjects demonstrate mixed findings (Barraza, McCullough, Ahmadi, & Zak, 2011; Declerck, Boone, & Kiyonari, 2013; De Dreu et al., 2010; Huffmeijer, Alink, Tops, Bakermans-Kranenburg, & van IJzendoorn, 2012; Israel, Weisel, Ebstein, & Bornstein, 2012; Scheele et al., 2014; Shalvi & De Dreu, 2014; Zak, Stanton, & Ahmadi, 2007; see Pfeiffer, 2013).

shackelf@oakland.edu

Another line of research investigates possible direct and indirect (mediation) relationships between an oxytocin gene allele and prosocial behaviors. For example, in one recent study, Christ et al. (in press) showed evidence that an oxytocin gene allele was associated with sympathy and subsequent altruistic but not public prosocial behaviors. The findings are consistent with the notion that sympathy is an important predictor of selfless actions. On the other hand, because helping in front of others (i.e., public helping) is considered a selfish form of helping (to gain approval from others), the nonsignificant relationship between oxytocin and sympathy in public helping was expected. Studies such as these help bridge the gap across the biological, psychological, and behavioral divides.

Similarly, serotonin and dopamine have also been associated with prosocial traits and behaviors. The findings regarding these neurotransmitters are somewhat mixed (Crockett, Clark, Hauser, & Robbins, 2010; Jiang, Chew, & Ebstein, 2013; Knafo et al., 2011; Marsh et al., 2011; Stoltenberg, Christ, & Highland, 2012). However, serotonin has been linked to impulsivity, anxiety, substance use and risky behaviors, harm aversion, and a variety of other traits (Crockett et al., 2010; Stoltenberg et al., 2012). If one considers some forms of prosocial behaviors—forms that incur a sacrifice or high cost to the self as risky—then serotonin may be especially linked to such prosocial behaviors. Moreover, the relation of serotonin to prosocial behavior may be accounted for by links to traits related to prosociality (see Stoltenberg et al., 2012, for supportive evidence). In general, then, both of these neurotransmitters may facilitate prosociality in a complex or indirect manner.

Temperament

Children's temperamental characteristics, i.e., their biologically influenced "individual differences in reactivity and regulation" (Rothbart, Ahadi, & Evans, 2000, p. 122), have been implicated in individual differences in children's moral conduct and behavior. Temperament characteristics are present early in life, stable across the life span, and present in many species. Some theorists assert that temperament has a strong heritability component (Buss & Plomin, 1984), and other theorists suggest that such traits are modifiable by environmental forces (Derryberry & Rothbart, 1988; Thomas & Chess, 1977). Research generally indicates that both genetics and environment (especially non-shared environment) contribute to temperamental characteristics (see Saudino, 2005, for a review).

Although a number of temperamental dimensions have been implicated directly in children's moral conduct, the two that have been most explored by researchers are aspects of children's fear-proneness and children's effortful control (Hastings, Rubin, & DeRose, 2005; Kochanska & Aksan, 2006; Kochanska et al., 2002). Theoretically, fear-proneness should lead to children feeling discomfort and anxiety after wrongdoing or when anticipating wrongdoing (akin to guilt), and as a result,

shackelf@oakland.edu

children high in fear should be more likely to engage in moral conduct. Empirical research indicates that fear-proneness is inversely associated with children's rule violation behavior (e.g., cheating), and this link appears to be mediated by guilt (Kochanska et al., 2002). Research has also supported inverse links between children's fear-proneness and aggressive and delinquent behavior (Ladd & Profilet, 1996; Shaw, Gilliom, Ingoldsby, & Nagin, 2003). Fear-proneness, however, may have some costs with regard to prosocial behavior; children who are higher in inhibition (i.e., fear-proneness) may be less likely to empathize with or help unfamiliar others (Young, Fox, & Zahn-Waxler, 1999), although the research is equivocal on this issue (see Hastings et al., 2005).

Negative emotionality (or the intensity of experiencing negative affect, more generally) may be inversely related to moral affect and conduct. A number of researchers have argued for that negative emotionality inhibits children's prosocial behavior and increases their antisocial behavior (see Eisenberg et al., 2000, 2006). It seems likely that children who are high in negative emotionality might become easily over-aroused when their needs conflict with others or in situations where others are distressed, and this in turn limits their prosocial behavior. Similarly, children high in negative emotionality might have difficulty regulating or controlling their negative emotion when faced with conflict, which may increase their risk for responding with aggression. The research has generally supported the links between negative emotionality and children's aggressive behavior, but has been more mixed with regard to the links that negative emotionality shares with empathic and prosocial behavior (e.g., Eisenberg et al., 1996, 2006; Laible, Carlo, Murphy, Augustine, & Roesch, 2014; Laible, Carlo, Panfile, Eye, & Parker, 2010; see Eisenberg, 2000).

Effortful control, in contrast, is the self-regulatory capacity for temperament and involves several mechanisms for the regulation of both internal states and impinging stimuli (Derryberry & Rothbart, 1988; Eisenberg et al., 1996). Effortful control is composed of attentional, working memory and behavioral regulatory processes (Eisenberg, Hofer, & Vaughan, 2007; Rothbart, Ellis, Rosario Rueda, & Posner, 2003). According to scholars, effortful control seems particularly related to moral behaviors because the demands of attending to the needs of others, responding emotionally, and acting in a manner that helps others necessitate a well-regulated cognitive, affective, and behavioral set of responses (Eisenberg et al., 2007). The links between effortful control and children's moral empathic, prosocial, and moral responding have been well established (Kochanska, Murray, Jacques, Koenig, & Vandegeest, 1996; Kochanska, Murray, & Coy, 1997; Kochanska, Murray, & Harlan, 2000; Laible et al., 2014; Rothbart, Ahadi, & Hershey, 1994). In addition, effortful control may moderate the links between negative emotionality and children's prosocial and empathic responding (Eisenberg & Fabes, 1992, 2006; Laible et al., 2014). Thus, children moderate or high in negative emotionality may engage in moral conduct if they are also well regulated and able to successfully control their negative emotional reactions.

shackelf@oakland.edu

Puberty

Much attention has been devoted to the possible role of pubertal timing and status on aggressive and antisocial behaviors. In this research, though sometimes with mixed results, there is evidence that early pubertal timing is associated with higher prevalence of problem behaviors (although sometimes for girls and not boys; e.g., Graber, Seeley, Brooks-Gunn, & Lewinsohn, 2004; Ge, Conger, & Elder, 2001). Furthermore, the relationships seem to differ as a function of measurement issues (e.g., use of clinical measures versus self-reports). In general, theorists speculate that early-maturing youth may be more prone to problem behaviors as a result of identity issues or peer group processes (Coie & Dodge, 1998). Perhaps future researchers will more directly examine hormonal levels associated with pubertal status, and such research may help disentangle the mixed findings.

Somewhat surprisingly, very little research has been devoted to investigating the relationships between pubertal timing and status and prosociality. Fabes, Carlo, Kupanoff, and Laible (1999) speculated that associated increases of pubertal hormones could result in either increased emotional sensitivity or decreased impulsivity. In the former, one might expect positive relationships between pubertal timing and prosocial behaviors, whereas in the latter scenario, one might expect negative relationships between such processes. In the one most direct study, investigators found a positive relationship between early pubertal timing and prosocial behaviors, but only for boys and not for girls (Carlo, Crockett, Wolff, & Beal, 2012). Clearly, definitive conclusions are not appropriate given the scarcity of research on this topic, but these findings are provocative and suggest that early pubertal timing may enhance prosociality (at least for boys).

Biology × Environment Approaches

Scholarly discussion regarding the importance of research on Biology × Environment interactions abounds (see e.g., Warneken & Tomasello, 2009). No doubt, a comprehensive understanding of morality will necessitate models that account for both biological and environment processes. However, such research presents many challenges both conceptually and methodologically. Among the many challenges, one major challenge is the difficulty of disentangling biological-based and environment-based processes because such processes are inherently intertwined (Emde et al., 1992; Plomin et al., 1993). In light of such challenges, inferences and conclusions of findings from such study designs need to be cautious. Despite these challenges, developmental researchers have conducted studies designed to examine Parenting × Temperament and Culture × Biology interaction effects on children's moral tendencies.

shackelf@oakland.edu

Temperament × Parenting Studies

Recent theorizing has proposed that some children are more influenced by their rearing environment than others, and as a result of this differential susceptibility, these children can both benefit more from positive environmental experiences (such as maternal sensitivity) and be more at risk from negative environmental experiences (such as abuse) (Belsky & Pluess, 2009). In addition to exploring how specific genes operate in making children more vulnerable to their rearing environment (e.g., Caspi et al., 2002), aspects of temperament have also been explored as markers of sensitivity to the rearing environment. Most of the work on differential susceptibility with regard to temperament has focused on the role of difficult temperament. This research supports the idea that children with a difficult temperament (i.e., those high in negative emotionality and low in regulation) have a higher proportion of variance accounted for in outcomes by rearing environments than those without this temperamental profile (Belsky, 2005). Other work has confirmed that children with difficult temperaments seem to be more influenced for better or worse by their home environments (see Belsky & Pluess, 2009, for a review). For example, Van Aken, Junger, Verhoeven, Van Aken, and Deković (2007) found toddlers with difficult temperaments (followed across a 6-month period) displayed the smallest increases in externalizing behavior when they had sensitive mothers who used reasonable control, but had the highest increases when reared by mothers who used harsh control.

The work on differential susceptibility and on other models that support interactions between temperament and parenting, such as goodness of fit models (Thomas, Chess, & Birch, 1968), supports the idea that there is a complex interplay between environmental influences and biology in predicting children's developmental outcomes (and that this interplay might not always be captured by simple main effects or variable-centered models) (see, e.g., Laible et al., 2014). Thus, researchers need to explore the possibility of interactions between environmental and biological factors and consider using alternative statistical models (such as person-centered approaches using latent class analysis) to understand outcomes related to children's moral development. It is also worth pointing out, however, that researchers have not explored whether other temperamental characteristics beyond difficultness are indicative of susceptibility to environmental influences. Moreover, in this line of research, there are no controls for variance due to shared genes between caregiver and child. In addition, although there is much work investigating interactions between temperament and parenting for predicting children's aggressive and externalizing behavior (see Belsky & Pluess, 2009, for a review), there is very little work on how these two sets of variables might interact to produce prosocial or empathic behavior (for an exception, see Carlo, Roesch, & Melby, 1998). There has been some limited work investigating how particular genes (e.g., 7-repeat DRD4 allele) interact with sensitive parenting to predict prosocial behavior, and these works support the idea that genotype does interact with sensitive parenting to predict children's prosocial behavior. For example, a recent study found that children with the

shackelf@oakland.edu

DRD4 7-repeat allele are more prosocial as a result of maternal positive parenting than children without this gene who also experienced positive parenting (Knafo et al., 2011).

Culture × Biology Studies

There are recent attempts to integrate culture and evolution theories that show much promise for understanding morality. Chudek and Henrich (2011), for example, propose a culture-gene, coevolution theory that conceptualizes moral behaviors as norms and proposes culture- and evolution-based explanations for apparently unique aspects of human morality. At the more proximate level, a somewhat compatible approach to examining Biology × Environment interactions is based on identifying culture-related processes that may interact with individual difference traits. As noted previously, culture can be considered reflective of a child's environment, but rather than simply examine context variables, one can examine psychological variables that reflect the child's culture (similar to the notion of the developmental niche, see Harkness & Super, 2002). This approach is based on the assumption that there may be individual differences in children's perceptions of their cultural environment.

Research designed to directly examine interactive effects of Biology × Culture on prosocial and moral tendencies is rare. One recent study showed that variation in a serotonin receptor gene allele across different cultures that differed in ecological threat is associated with moral judgment (Mrazek, Chiao, Blizinsky, Lun, & Gelfand, 2013). Although these findings are suggestive of possible Biology × Culture interactions, the absence of direct measures of culture-related processes limits our ability to confidently infer culture effects. Therefore, although there is promise for new research that incorporates the moderating effects of culture-related processes, such research on moral development is still to be conducted.

Summary and Future Directions

The present chapter presents a brief summary of moral cognitions, moral emotions, and moral behaviors (with a strong focus on prosociality) and their biological and cultural underpinnings from a human developmental and cultural perspective. As can be surmised, the resurgence in the biological bases of prosocial and moral behaviors has resulted in a flurry of provocative findings that is pushing the boundaries of moral development theories. Much of the findings on the biological basis of morality from human research parallels closely findings from research on other animal species (Carlo, 2014). Recent attention to culture-related processes is also resulting in important discoveries regarding the transmission of morality from parents to offspring and is furthering our understanding of culture group differences

shackelf@oakland.edu

and changes in conceptions of morality across time. However, much more research is needed on the biological foundations (especially focusing on moral cognitions), on culture-related processes, and on the interaction of biology and environment. Furthermore, research in many of these areas remains conceptually and methodologically limited, and there are major gaps in our understanding. Many of these limitations stem from the relatively disparate lines of scholarship and the overly simple conceptions of morality and moral behaviors.

Recent conceptualizations of morality provide new windows of opportunity to significantly advance the field. For example, distinctions between moral domains such as the distinction between justice-based and compassion-based morality may help explain apparently mixed findings and further our understanding of sociomoral development. These advances will also make scholarly work much more relevant to addressing major sociomoral issues. Moreover, although scholars have studied prosocial and antisocial behaviors for several decades, much of the existing research operationalized such behaviors as relatively global constructs. Fortunately, scholars have shifted more attention to refining the study of moral behaviors such that prosocial and antisocial behaviors are conceptualized as relatively complex constructs. Researchers studying aggression, for example, have documented important and unique characteristics of distinct forms of aggression. These newer conceptualizations demonstrate that the developmental trajectories and the correlates of direct, indirect, verbal, physical, instrumental, and relational aggression are distinct (Dodge, Coie, & Lynam, 2006). This research led to significant advances in our understanding of the biological underpinnings of aggression, of the development of aggressive tendencies, and of individual and group differences in aggression (Dodge et al., 2006). This advancement also has facilitated scholarly exchanges across disparate disciplines, including across human and nonhuman animal research.

In contrast, prosocial development scholars have only much more recently emphasized distinct forms of prosocial behaviors (Padilla-Walker & Carlo, 2014). Interestingly, early work on prosocial behaviors highlighted the task specificity of such behaviors (Hartshorne et al., 1930), but such work had been virtually ignored for decades. In recent years, however, several lines of research point out the importance of distinguishing between prosocial behaviors across different contexts (e.g., home, school) and toward different targets (e.g., ingroup/outgroup, family, friends, strangers). This work has also resulted in further attention to the measurement of prosocial behaviors. Based on prior research and focus group data, our research team identified six different types of prosocial behaviors that differ across context and motives (Carlo & Randall, 2001). *Altruistic* actions refer to behaviors conducted with little or no regard for self reward. *Compliant* is defined as helping others when asked. Assisting others in front of an audience is referred to as *Public*. *Anonymous* helping is exhibited when persons assist others without revealing their identity. Helping others under emotionally evocative situations is labeled *Emotional* actions, whereas helping others in emergency situations is termed *Dire*. This typology and the resulting validated instruments (the Prosocial Tendencies Measure and Prosocial Tendencies Measure-Revised) have been used in several studies across several cultural and ethnic groups. Findings suggest that, indeed, these distinct

shackelf@oakland.edu

forms of prosocial behaviors have unique correlates (see McGinley, Opal, Richaud, & Mesurado, 2014). However, to date, research is limited examining the biological bases of specific forms of prosocial behaviors though, as noted previously, findings regarding neurotransmitters and prosociality suggest different findings across distinct forms of prosocial behaviors. In one study, researchers demonstrated a complex pattern of relationships between the presence of specific oxytocin genotypes (OXTR) and lower levels of emotional and anonymous prosocial behaviors and higher levels of public prosocial behaviors (also see Christ et al., in press). Furthermore, some research suggests bidirectional relationships between neurotransmitters and prosocial tendencies—for example, one study found that empathy-inducing situations may also trigger oxytocin release (Barraza & Zak, 2009).

Another future avenue concerns our conceptions of morality and their implications for understanding other human behaviors. Morality is a broad and encompassing area of study. Scholars have already presented evidence that the study of morality has important implications beyond the moral sphere, including extending our understanding of psychological and physical health, academic achievement, and close interpersonal relationships. For example, a recent study demonstrated that oxytocin administration can inhibit atherosclerotic lesion development and inflammation (Szeto et al., 2013). These findings are in line with other research indicating that prosociality may attenuate disease and illness (Schreier, Schonert-Reichl, & Chen, 2013; see Carlo, 2014). Perhaps the links between prosocial behaviors and health can be partly explained via changes in oxytocin (or other neurotransmitters). Stoltenberg et al. (2012) showed that the relationship between the presence of a serotonin receptor gene allele and prosocial behaviors is partly explained via social anxiety. Other research suggests negative relations between prosocial behaviors and substance use (Barber, Eccles, & Stone, 2001; Carlo, Crockett, Wilkinson, & Beal, 2011). Moreover, longitudinal research suggests that exhibiting relatively high levels of prosociality early in life predicted later mortality and medical problems across a 27-year span (Danner, Snowdon, & Friesen, 2001).

With regard to the links between prosocial behaviors and academic achievement, there is robust evidence that frequent early prosocial behaviors reliably predicts better academic outcomes in later adolescence and young adulthood (Caprara, Barbaranelli, Pastorelli, Bandura, & Zimbardo, 2000; Carlo, Mestre, Samper, Tur, & Armenta, 2011; Wentzel, 1993). For example, in one notable study, the investigators showed that frequent expression of prosocial behaviors in elementary school predicted better academic outcomes 5 years later (Caprara et al., 2000). Furthermore, the expression of early aggression did not predict such outcomes. Several mechanisms have been proposed to explain such findings, including the possibility that children with prosocial tendencies may be better liked by teachers, may have better relationships with teachers and fellow students, or may be engaged in school activities. However, another possibility is that children with prosocial tendencies may be less prone to impulsivity and deficits in self-regulation (associated with serotonin), which may facilitate academic performance. Indeed, as noted previously, researchers have documented relationships between prosocial behaviors and self-regulation, including effortful control and emotion regulation (Eisenberg et al., 2006).

shackelf@oakland.edu

Prosocial behaviors have also been linked to the development and maintenance of close interpersonal relationships. A common approach to developing a new relationship is to engage in prosocial actions toward others, and a good strategy for maintaining and building close interpersonal relationships is to express prosociality toward those others. As noted previously, prosocial behaviors are linked with trust and trustworthiness and with better quality parent–child, sibling, peer, and marital relationships (Rotenberg et al., 2005). Investigators have also demonstrated links between oxytocin and attachment relationships, bonding, approach actions, and sociability (Krueger et al., 2012). Thus, oxytocin may play an important role in facilitating the development of close interpersonal relationships and prosocial behaviors. Some studies suggest that the link between oxytocin and prosocial behaviors may not be direct but rather indirect via relationship variables. For example, Krueger et al.'s (2012) findings suggest that oxytocin may be linked to altruistic behaviors via its influence on trustworthiness. Alternatively, oxytocin might promote prosocial behaviors and subsequently foster close interpersonal relationships. Thus, it appears that neurotransmitters play a central role in explaining the relationships between prosocial behaviors and quality of interpersonal relationships and that future research is needed to examine possible bidirectional relations.

These future directions clearly do not do justice to the many possibilities in the study of the biological foundations of morality. We expect new and valuable advances will occur as the field continues to evolve. Thus, as the study of morality gradually becomes more interdisciplinary, we expect greater attention to more refined conceptualizations of morality and greater methodological advances in this field. Our hope is that these new conceptions will better integrate the biological, cultural, and developmental aspects of morality.

References

Alexander, R. D. (1987). *The biology of moral systems*. Piscataway, NJ: Transaction Publishers.

Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Holt.

Armenta, B. E., Knight, G. P., Carlo, G., & Jacobson, R. P. (2011). The relation between ethnic group attachment and prosocial tendencies: The mediating role of ethnically related cultural values. *European Journal of Social Psychology, 41*, 107–115.

Bandura, A. (1986). *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice Hall.

Barber, B. L., Eccles, J. S., & Stone, M. R. (2001). Whatever happened to the jock, the brain, and the princess? Young adult pathways linked to adolescent activity involvement and social identity. *Journal of Adolescent Research, 16*(5), 429–455.

Barraza, J. A., McCullough, M. E., Ahmadi, S., & Zak, P. J. (2011). Oxytocin infusion increases charitable donations regardless of monetary resources. *Hormones and Behavior, 60*(2), 148–151.

Barraza, J. A., & Zak, P. J. (2009). Empathy toward strangers triggers oxytocin release and subsequent generosity. *Annals of the New York Academy of Sciences, 1167*(1), 182–189.

Batson, C. D. (1998). Altruism and prosocial behavior. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 2, pp. 282–316). Boston: McGraw-Hill.

shackelf@oakland.edu

Belsky, J. (2005). Differential susceptibility to rearing influence: An evolutionary hypothesis and some evidence. In B. J. Ellis & D. F. Bjorklund (Eds.), *Origins of the social mind: Evolutionary psychology and child development* (pp. 139–163). New York: Guilford Press.

Belsky, J., & Pluess, M. (2009). Beyond diathesis stress: Differential susceptibility to environmental influences. *Psychological Bulletin, 135*(6), 885.

Blasi, A. (1983). Moral cognition and moral action: A theoretical perspective. *Developmental Review, 3*(2), 178–210.

Brosnan S. F. (2010). What do capuchin monkeys tell us about cooperation? In D. R. Forsyth & C. L. Hoyt (Eds.), *For the greater good of all: Perspectives on individualism, society, and leadership* (Vol. Jepson Studies in Leadership Series, pp. 11–28), Hampshire, UK: Palgrave Macmillan Publishers.

Buss, A. H., & Plomin, R. (1984). *Temperament: Early developing personality traits*. New York: Erlbaum.

Campbell, D. T. (1975). On the conflicts between biological and social evolution and between psychology and moral tradition. *American Psychologist, 30*(12), 1103.

Calderón-Tena, C. O., Knight, G. P., & Carlo, G. (2011). The socialization of prosocial behavioral tendencies among Mexican American adolescents: The role of familism values. *Cultural Diversity and Ethnic Minority Psychology, 17*(1), 98.

Caprara, G. V., Barbaranelli, C., Pastorelli, C., Bandura, A., & Zimbardo, P. G. (2000). Prosocial foundations of children's academic achievement. *Psychological Science, 11*(4), 302–306.

Caprara, G. V., & Steca, P. (2007). Prosocial agency: The contribution of values and self-efficacy beliefs to prosocial behavior across ages. *Journal of Social and Clinical Psychology, 26*(2), 218–239. doi:10.1521/jscp.2007.26.2.218.

Carlo, G. (2014). The development and correlates of prosocial moral behaviors. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (2nd ed., pp. 208–234). New York: Psychology Press.

Carlo, G. (2006). Care-based and altruistically-based morality. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 551–579). Mahwah, NJ: Lawrence Erlbaum Associates.

Carlo, G., Crockett, L. J., Wilkinson, J. L., & Beal, S. J. (2011). The longitudinal relationships between rural adolescents' prosocial behaviors and young adult substance use. *Journal of Youth and Adolescence, 40*(9), 1192–1202.

Carlo, G., Crockett, L. J., Wolff, J. M., & Beal, S. J. (2012). The role of emotional reactivity, self-regulation, and puberty in adolescents' prosocial behaviors. *Social Development, 21*(4), 667–685.

Carlo, G., Knight, G. P., McGinley, M., Goodvin, R., & Roesch, S. C. (2010). Understanding the developmental relations between perspective taking and prosocial behaviors: A meta-analytic review. In J. Carpendale, G. Iarocci, U. Muller, B. Sokol, & A. Young (Eds.), *Self- and social-regulation: Exploring the relations between social interaction, social cognition, and the development of executive functions* (pp. 234–269). Oxford, UK: Oxford University Press.

Carlo, G., Knight, G. P., McGinley, M., & Hayes, R. (2011). The roles of parental inductions, moral emotions, and moral cognitions in prosocial tendencies among Mexican American and European American early adolescents. *The Journal of Early Adolescence, 31*(6), 757–781.

Carlo, G., McGinley, M., Davis, A., & Streit, C. (2012). Behaving badly or goodly: Is it because I feel guilty, shameful, or sympathetic? Or is it a matter of what I think? *New Directions for Youth Development, 2012*(136), 75–93.

Carlo, G., Mestre, M. V., Samper, P., Tur, A., & Armenta, B. E. (2010). Feelings or cognitions? Moral cognitions and emotions as longitudinal predictors of prosocial and aggressive behaviors. *Personality and Individual Differences, 48*, 872–877.

Carlo, G., Mestre, M. V., Samper, P., Tur, A., & Armenta, B. E. (2011). The longitudinal relations among dimensions of parenting styles, sympathy, prosocial moral reasoning, and prosocial behaviors. *International Journal of Behavioral Development, 35*(2), 116–124.

Carlo, G., & Randall, B. A. (2001). Are all prosocial behaviors equal? A socioecological developmental conception of prosocial behavior. *Advances in Psychology Research, 2*, 151–170.

shackelf@oakland.edu

Carlo, G., Roesch, S. C., & Melby, J. (1998). The multiplicative relations of parenting and temperament to prosocial and antisocial behaviors in adolescence. *The Journal of Early Adolescence, 18*(3), 266–290.

Carlo, G. & Davis, A. (in press). Benevolence in a justice-based world: The power of sentiments in predicting prosocial behaviors. In J. Annas, D. Narvaez, & N. E. Snow (Eds.), *Developing the virtues: Integrating perspectives*. New York: Oxford University Press.

Carlo, G., & de Guzman, M. R. T. (2009). Theories and research on prosocial competencies among US Latinos/as. *Handbook of US Latino psychology: Developmental and community-based perspectives* (pp. 191–211).

Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W., et al. (2002). Role of genotype in the cycle of violence in maltreated children. *Science, 297*(5582), 851–854.

Christ, C. C., Carlo, G., & Stoltenberg, S. F. (2015). Oxytocin receptor (OXTR) single nucleotide polymorphisms indirectly predict prosocial behavior through perspective taking and empathic concern. *Journal of Personality.* doi:10.1111/jopy.12152.

Chudek, M., & Henrich, J. (2011). Culture–gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in Cognitive Sciences, 15*(5), 218–226.

Ci, H., Wu, N., & Su, Y. (2014). Clock gene modulates roles of OXTR and AVPR1b genes in prosociality. *PLoS One, 9*(10), e109086.

Coie, J. D., & Dodge, K. A. (1998). Aggression and antisocial behavior. In: W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology, Vol. 3: Social, emotional, and personality development* (5th ed., pp. 779–862). New York: John Wiley & Sons.

Colby, A., & Kohlberg, L. (1987). *The measurement of moral judgment: Theoretical foundations and research validation* (Vol. 1). New York: Cambridge.

Crockett, M. J., Clark, L., Hauser, M. D., & Robbins, T. W. (2010). Serotonin selectively influences moral judgment and behavior through effects on harm aversion. *Proceedings of the National Academy of Sciences, 107*(40), 17433–17438.

Crockett, M. J., Clark, L., Lieberman, M. D., Tabibnia, G., & Robbins, T. W. (2010). Impulsive choice and altruistic punishment are correlated and increase in tandem with serotonin depletion. *Emotion, 10*(6), 855.

Danner, D. D., Snowdon, D. A., & Friesen, W. V. (2001). Positive emotions in early life and longevity: Findings from the nun study. *Journal of Personality and Social Psychology, 80*(5), 804.

Davis, A. N., Carlo, G., & Knight, G. P. (in press). Perceived maternal parenting styles, cultural values, and prosocial tendencies among Mexican American youth. *Journal of Genetic Psychology*.

De Dreu, C. K., Greer, L. L., Handgraaf, M. J., Shalvi, S., Van Kleef, G. A., Baas, M., et al. (2010). The neuropeptide oxytocin regulates parochial altruism in intergroup conflict among humans. *Science, 328*(5984), 1408–1411.

De Waal, F. B. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology, 59*, 279–300.

Decety, J., & Svetlova, M. (2012). Putting together phylogenetic and ontogenetic perspectives on empathy. *Developmental Cognitive Neuroscience, 2*(1), 1–24.

Declerck, C., Boone, C., & Kiyonari, T. (2013). The effect of oxytocin on cooperation in a prisoner's dilemma depends on the social context and a person's social value orientation. *Social Cognitive and Affective Neuroscience*, nst040.

Derryberry, D., & Rothbart, M. K. (1988). Arousal, affect, and attention as components of temperament. *Journal of Personality and Social Psychology, 55*(6), 958.

Dodge, K. A., Coie, J. D., & Lynam, D. (2006). Aggression and antisocial behavior in youth. In W. Damon & R. M. Lerner (Series Ed.) and N. Eisenberg (Vol. Ed.), *Handbook of child psychology. Vol. 3. Social, emotional, and personality development* (6th ed.). New York: Wiley.

Eisenberg, N. (1986). *Altruistic emotion, cognition, and behavior*. Mahwah, NJ: Lawrence Erlbaum Associates.

Eisenberg, N. (2000). Emotion, regulation, and moral development. *Annual review of psychology, 51*(1), 665–97.

Eisenberg, N., & Fabes, R. A. (Eds.). (1992). *Emotion and its regulation in early development*. San Francisco: Jossey-Bass.

shackelf@oakland.edu

Eisenberg, N., & Fabes, R. A. (2006). Emotion-related regulation and children's socioemotional competence. In L. Balter & C. S. TamisLeMonda (Eds.), *Child psychology: A handbook of contemporary issues* (2nd ed., pp. 357–381). New York: Psychology Press.

Eisenberg, N., Fabes, R. A., Karbon, M., Murphy, B. C., Wosinski, M., Polazzi, L., et al. (1996). The relations of children's dispositional prosocial behavior to emotionality, regulation, and social functioning. *Child development, 67*(3), 974–992.

Eisenberg, N., Guthrie, I. K., Fabes, R. A., Shepard, S., Losoya, S., Murphy, B., et al. (2000). Prediction of elementary school children's externalizing problem behaviors from attentional and behavioral regulation and negative emotionality. *Child development, 71*(5), 1367–1382.

Eisenberg, N., Guthrie, I. K., Murphy, B. C., Shepard, S. A., Cumberland, A., & Carlo, G. (1999). Consistency and development of prosocial dispositions: A longitudinal study. *Child Development, 70*(6), 1360–1372.

Eisenberg, N., Hofer, C., & Vaughan, J. (2007). Effortful control and its socioemotional consequences. In J. J. Gross (Ed.), *Handbook of emotion regulation* (pp. 287–306). New York: Guilford.

Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial development. In W. Damon & R. M. Lerner (Series Eds.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology, Vol. 3: Social, emotional, and personality development* (6th ed., pp. 646–718). New York: John Wiley.

Emde, R. N., Plomin, R., Robinson, J., Corley, R., DeFries, J., Fulker, D. W., et al. (1992). Temperament, emotion, and cognition at fourteen months: The MacArthur Longitudinal Twin Study. *Child Development, 63*(6), 1437–1455.

Fabes, R. A., Carlo, G., Kupanoff, K., & Laible, D. (1999). Early adolescence and prosocial/moral behavior I: The role of individual processes. *The Journal of Early Adolescence, 19*(1), 5–16.

Ferguson, T. J., Stegge, H., Miller, E. R., & Olsen, M. E. (1999). Guilt, shame, and symptoms in children. *Developmental Psychology, 35*(2), 347.

Fung, H. (1999). Becoming a moral child: The socialization of shame among young Chinese children. *Ethos, 27*(2), 180–209.

Ge, X., Conger, R. D., & Elder, G. H., Jr. (2001). Pubertal transition, stressful life events, and the emergence of gender differences in adolescent depressive symptoms. *Developmental Psychology, 37*(3), 404.

Graber, J. A., Seeley, J. R., Brooks-Gunn, J., & Lewinsohn, P. M. (2004). Is pubertal timing associated with psychopathology in young adulthood? *Journal of the American Academy of Child & Adolescent Psychiatry, 43*(6), 718–726.

Grusec, J. E., & Goodnow, J. J. (1994). Impact of parental discipline methods on the child's internalization of values: A reconceptualization of current points of view. *Developmental Psychology, 30*(1), 4.

Hardy, S. A., & Carlo, G. (2005). Religiosity and prosocial behaviours in adolescence: The mediating role of prosocial values. *Journal of Moral Education, 34*(2), 231–249.

Hardy, S. A., Carlo, G., & Roesch, S. C. (2010). Links between adolescents' expected parental reactions and prosocial behavioral tendencies: The mediating role of prosocial values. *Journal of Youth and Adolescence, 39*(1), 84–95.

Harkness, S., & Super, C. M. (2002). Culture and parenting. In M. H. Bornstein (Ed.), *Handbook of parenting, Vol. 2: Biology and ecology of parenting* (2nd ed., pp. 253–280). Mahwah, NJ: Lawrence Erlbaum Associates.

Hartshorne, H., May, M. A., & Shuttleworth, F. K. (1930). Studies in the nature of character. Studies in the organization of character, vol. 3. New York: Macmillan.

Hastings, P. D., Rubin, K. H., & DeRose, L. (2005). Links among gender, inhibition, and parental socialization in the development of prosocial behavior. *Merrill-Palmer Quarterly (1982)*, 467–493.

Hawley, P. H. (2014). Evolution, prosocial behavior, and altruism: A roadmap for understanding where the proximate meets the ultimate. In L. M. Padilla-Walker & G. Carlo (Eds.), *Prosocial development: A multidimensional approach* (pp. 43–69). New York: Oxford University Press.

Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice*. Cambridge: University Press.

shackelf@oakland.edu

Huffmeijer, R., Alink, L. R., Tops, M., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2012). Asymmetric frontal brain activity and parental rejection predict altruistic behavior: Moderation of oxytocin effects. *Cognitive, Affective, & Behavioral Neuroscience, 12*(2), 382–392.

Israel, S., Weisel, O., Ebstein, R. P., & Bornstein, G. (2012). Oxytocin, but not vasopressin, increases both parochial and universal altruism. *Psychoneuroendocrinology, 37*(8), 1341–1344.

Jensen, K., & Silk, J. B. (2014). Searching for the evolutionary roots of human morality. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (2nd ed., pp. 475–494). New York: Psychology Press.

Jiang, Y., Chew, S. H., & Ebstein, R. P. (2013). The role of D4 receptor gene exon III polymorphisms in shaping human altruism and prosocial behavior. *Frontiers in Human Neuroscience, 7*.

Kant, I. (1755). *Inquiry concerning the distinctness of the principles of natural theology and morality*. KANT, I. Theoretical Philosophy, 1770, 248.

Knafo, A., Israel, S., & Ebstein, R. P. (2011). Heritability of children's prosocial behavior and differential susceptibility to parenting by variation in the dopamine receptor D4 gene. *Development and Psychopathology, 23*(01), 53–67.

Knafo, A., & Plomin, R. (2006). Prosocial behavior from early to middle childhood: Genetic and environmental influences on stability and change. *Developmental Psychology, 42*(5), 771.

Knight, G. P., & Carlo, G. (2012). Prosocial development among Mexican American youth. *Child Development Perspectives, 6*(3), 258–263.

Knight, G. P., Bernal, M. E., & Carlo, G. (1995). Socialization and the development of cooperative, competitive, and individualistic behaviors among Mexican American children. In *Meeting the challenge of linguistic and cultural diversity in early childhood education* (Vol. 6, pp. 85–102).

Knight, G. P., Carlo, G., Basilio, C. D., & Jacobson, R. P. (2014). Familism values, perspective taking, and prosocial moral reasoning: Predicting prosocial tendencies among mexican American Adolescents. *Journal of Research on Adolescence*. doi: 10.1111/jora.12164.

Kochanska, G. (1994). Beyond cognition: Expanding the search for the early roots of internalization and conscience. *Developmental Psychology, 30*, 20–22.

Kochanska, G., & Aksan, N. (2006). Children's conscience and self-regulation. *Journal of Personality, 74*(6), 1587–1618.

Kochanska, G., Gross, J. N., Lin, M. H., & Nichols, K. E. (2002). Guilt in young children: Development, determinants, and relations with a broader system of standards. *Child Development, 73*(2), 461–482.

Kochanska, G., Murray, K. T., & Harlan, E. T. (2000). Effortful control in early childhood: Continuity and change, antecedents, and implications for social development. *Developmental Psychology, 36*(2), 220.

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development, 67*(2), 490–507.

Kochanska, G., Murray, K., & Coy, K. C. (1997). Inhibitory control as a contributor to conscience in childhood: From toddler to early school age. *Child Development, 68*(2), 263–277.

Kohlberg, L. (1969). *Stage and sequence: The cognitive-developmental approach to socialization*. New York: Rand McNally.

Kohlberg, L., & Candee, D. (1984). The relationship of moral judgment to moral action. *Morality, Moral Behavior, and Moral Development, 52*, 73.

Krueger, F., Parasuraman, R., Iyengar, V., Thornburg, M., Weel, J., Lin, M., et al. (2012). Oxytocin receptor genetic variation promotes human trust behavior. *Frontiers in Human Neuroscience, 6*.

Kurtines, W. M., Mayock, E., Pollard, S. R., Lanza, T., & Carlo, G. (1991). Social and moral development from the perspective of psychosocial theory. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Handbook of moral behavior and development, Vol. 1: Theory* (pp. 303–333). New York: Erlbaum.

Ladd, G. W., & Profilet, S. M. (1996). The Child Behavior Scale: A teacher-report measure of young children's aggressive, withdrawn, and prosocial behaviors. *Developmental Psychology, 32*(6), 1008.

shackelf@oakland.edu

Laible, D., Carlo, G., Panfile, T., Eye, J., & Parker, J. (2010). Negative emotionality and emotion regulation: A person-centered approach to predicting socioemotional adjustment in young adolescents. *Journal of Research in Personality, 44*(5), 621–629.

Laible, D., Eye, J., & Carlo, G. (2008). Dimensions of conscience in mid-adolescence: Links with social behavior, parenting, and temperament. *Journal of Youth and Adolescence, 37*(7), 875–887.

Laible, D., Carlo, G., Murphy, T., Augustine, M., & Roesch, S. (2014). Predicting children's prosocial and co-operative behavior from their temperamental profiles: A person-centered approach. *Social Development*.

Marsh, A. A., Crowe, S. L., Henry, H. Y., Gorodetsky, E. K., Goldman, D., & Blair, R. (2011). Serotonin transporter genotype (5-HTTLPR) predicts utilitarian moral judgments. *PLoS One, 6*(10), e25148.

Matthews, K. A., Batson, C. D., Horn, J., & Rosenman, R. H. (1981). "Principles in his nature which interest him in the fortune of others…": The heritability of empathic concern for others. *Journal of Personality, 49*(3), 237–247.

McAdams, D. P. (2009). The moral personality. *Personality, identity, and character: Explorations in moral psychology* (pp. 11–29).

McGinley, M., Opal, D., Richaud, M. C., & Mesurado, B. (2014). Cross-cultural evidence of multidimensional prosocial behaviors. In L. M. Padilla-Walker & G. Carlo (Eds.), *Prosocial development: A multidimensional approach* (pp. 258–278). New York: Oxford University Press.

Menesini, E., & Camodeca, M. (2008). Shame and guilt as behaviour regulators: Relationships with bullying, victimization and prosocial behaviour. *British Journal of Developmental Psychology, 26*(2), 183–196.

Mischel, W., & Mischel, H. N. (1976). A cognitive social-learning approach to morality and self-regulation. *Moral development and behavior: Theory, research, and social issues* (pp. 84–107).

Mrazek, A. J., Chiao, J. Y., Blizinsky, K. D., Lun, J., & Gelfand, M. J. (2013). The role of culture–gene coevolution in morality judgment: Examining the interplay between tightness–looseness and allelic variation of the serotonin transporter gene. *Culture and Brain, 1*(2–4), 100–117.

Ongley, S. F., Nola, M., & Malti, T. (2014). Children's giving: moral reasoning and moral emotions in the development of donation behaviors. *Frontiers in Psychology, 5*.

Padilla-Walker, L. M. (2007). Characteristics of mother–child interactions related to adolescents' positive values and behaviors. *Journal of Marriage and Family, 69*(3), 675–686.

Padilla-Walker, L. M., & Carlo, G. (2007). Personal values as a mediator between parent and peer expectations and adolescent behaviors. *Journal of Family Psychology, 21*(3), 538.

Padilla-Walker, L. M., & Carlo, G. (2014). Prosocial behavior: Past, present, and future. In L. M. Padilla-Walker & G. Carlo (Eds.), *Prosocial behavior: A multidimensional approach* (p. 3e16). New York: Oxford University Press.

Pfeiffer, U. (2013). Oxytocin—Not always a moral molecule. *Frontiers in Human Neuroscience, 7*, 10.

Piaget, J. (1965). *The moral judgment of the child (1932)*. New York: The Free.

Plomin, R., Kagan, J., Emde, R. N., Reznick, J. S., Braungart, J. M., Robinson, J., et al. (1993). Genetic change and continuity from fourteen to twenty months: The MacArthur Longitudinal Twin Study. *Child Development, 64*(5), 1354–1376.

Rest, J. (1983). Morality. In J. H. Flavell & E. M. Markman (Eds.), P. H. Mussen (Series Ed.), *Handbook of child psychology: Vol. 3. Cognitive development* (pp. 556–629). New York: Wiley.

Roberts, W., Strayer, J., & Denham, S. (2014). Empathy, anger, guilt: Emotions and prosocial behaviour. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 46*(4), 465.

Rotenberg, K. J., Fox, C., Green, S., Ruderman, L., Slater, K., Stevens, K., et al. (2005). Construction and validation of a children's interpersonal trust belief scale. *British Journal of Developmental Psychology, 23*(2), 271–293.

Rothbart, M. K., Ahadi, S. A., & Evans, D. E. (2000). Temperament and personality: Origins and outcomes. *Journal of Personality and Social Psychology, 78*(1), 122.

shackelf@oakland.edu

Rothbart, M. K., Ellis, L. K., Rosario Rueda, M., & Posner, M. I. (2003). Developing mechanisms of temperamental effortful control. *Journal of Personality, 71*(6), 1113–1144.

Rothbart, M. K., Ahadi, S. A., & Hershey, K. L. (1994). Temperament and social behavior in childhood. *Merrill-Palmer Quarterly (1982)*, 21–39.

Saudino, K. J. (2005). Behavioral genetics and child temperament. *Journal of Developmental and Behavioral Pediatrics, 26*(3), 214.

Scheele, D., Striepens, N., Kendrick, K. M., Schwering, C., Noelle, J., Wille, A., et al. (2014). Opposing effects of oxytocin on moral judgment in males and females. *Human Brain Mapping, 35*(12), 6067–6076.

Schreier, H. M., Schonert-Reichl, K. A., & Chen, E. (2013). Effect of volunteering on risk factors for cardiovascular disease in adolescents: A randomized controlled trial. *JAMA Pediatrics, 167*(4), 327–332.

Schwartz, S. H. (2010). Basic values: How they motivate and inhibit prosocial behavior. *Prosocial motives, emotions, and behavior: The better angels of our nature, 14*, 221–41.

Shalvi, S., & De Dreu, C. K. (2014). Oxytocin promotes group-serving dishonesty. *Proceedings of the National Academy of Sciences, 111*(15), 5503–5507.

Shaw, D. S., Gilliom, M., Ingoldsby, E. M., & Nagin, D. S. (2003). Trajectories leading to school-age conduct problems. *Developmental Psychology, 39*(2), 189.

Sober, E., & Wilson, D. S. (1998). *Unto others: The evolution and psychology of unselfish behavior*. Cambridge, MA: Harvard University Press.

Sommerville, J. A., Schmidt, M. F., Yun, J. E., & Burns, M. (2013). The development of fairness expectations and prosocial behavior in the second year of life. *Infancy, 18*(1), 40–66.

Staub, E. (1978). *Positive social behavior and morality: I. Social and personal influences*. Waltham, MA: Academic Press.

Stoltenberg, S. F., Christ, C. C., & Highland, K. B. (2012). Serotonin system gene polymorphisms are associated with impulsivity in a context dependent manner. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 39*(1), 182–191.

Streit, C., Carlo, G., Killoren, S. E., Len-Rios, M. E., Cooper, L., Deutsch, A., et al. (2015). Buffering effects of moral identity on the links between peer relationship quality and problem behaviors. *Drug & Alcohol Dependence, 146*, e110.

Subiaul, F., Vonk, J., Okamoto-Barth, S., & Barth, J. (2008). Do chimpanzees learn reputation by observation? Evidence from direct and indirect experience with generous and selfish strangers. *Animal Cognition, 11*(4), 611–623.

Szeto, A., Rossetti, M. A., Mendez, A. J., Noller, C. M., Herderick, E. E., Gonzales, J. A., et al. (2013). Oxytocin administration attenuates atherosclerosis and inflammation in Watanabe Heritable Hyperlipidemic rabbits. *Psychoneuroendocrinology, 38*(5), 685–693.

Tangney, J. P. (1996). Conceptual and methodological issues in the assessment of shame and guilt. *Behaviour Research and Therapy, 34*(9), 741–754.

Tangney, J. P., & Dearing, R. L. (2002). *Shame and guilt (emotions and social behavior)*. New York: Guilford Press.

Thomas, A., & Chess, S. (1977). *Temperament and development*. New York: Brunner/Mazel.

Thomas, A., Chess, S., & Birch, H. (1968). *Temperament and behavior disorders in children*. New York: New York University Press.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46*, 35–57.

Turiel, E. (2006). Thought, emotions, and social interactional processes in moral development. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 7–36). Mahwah, NJ: Erlbaum.

Van Aken, C., Junger, M., Verhoeven, M., Van Aken, M., & Deković, M. (2007). The interactive effects of temperament and maternal parenting on toddlers' externalizing behaviours. *Infant and Child Development, 16*(5), 553–572.

Walter, N. T., Montag, C., Markett, S., Felten, A., Voigt, G., & Reuter, M. (2012). Ignorance is no excuse: Moral judgments are influenced by a genetic variation on the oxytocin receptor gene. *Brain and Cognition, 78*(3), 268–273.

shackelf@oakland.edu

Warneken, F., & Tomasello, M. (2009). Varieties of altruism in children and chimpanzees. *Trends in Cognitive Sciences, 13*(9), 397–402.
Wentzel, K. R. (1993). Does being good make the grade? Social behavior and academic competence in middle school. *Journal of Educational Psychology, 85*(2), 357.
Wilson, E. O. (1978). *On human nature* (p. 175). New York: Bantam.
Wong, D. B. (1984). *Moral relativity*. Oakland, CA: University of California Press.
Wong, D. B. (2009). Emotion and the cognition of reasons in moral motivation. *Philosophical Issues, 19*(1), 343–367.
Young, S. K., Fox, N. A., & Zahn-Waxler, C. (1999). The relations between temperament and empathy in 2-year-olds. *Developmental Psychology, 35*(5), 1189.
Zahn-Waxler, C., Radke-Yarrow, M., & King, R. A. (1979). Child rearing and children's prosocial initiations toward victims of distress. *Child Development, 319–330*.
Zak, P. J., Stanton, A. A., & Ahmadi, S. (2007). Oxytocin increases generosity in humans. *PLoS One, 2*(11), e1128.

shackelf@oakland.edu

Part II
Philosophy and Ethics

shackelf@oakland.edu

Evolutionary Awareness: A Metacognitive Framework for Ethics

Gregory Gorelik and Todd K. Shackelford

Evolutionary Awareness: A Metacognitive Framework for Ethics

One of the most important foundations of modern education is the scientific framework of evolution by natural selection. Not only does this framework encompass all of biology, but it is also fundamental to psychology, anthropology, history, economics, political science, and even art and aesthetics. This does not mean that these disciplines are "reducible" to evolution by natural selection or that other theoretical frameworks cannot explain any of their aspects—only that a proper understanding of them is incomplete without an understanding of their dependence on processes that have evolved for billions of years, processes that unite us and other organisms in the same tree of life.

Prior to the publication of Darwin's *Origin of Species*, there was a coherent reason—if not an excuse—for humanity's stumbling in the dark with respect to the origins, the present state (at the time), and the future of humanity. How could humanity improve the short-term and long-term well-being of individuals and societies if it lacked both the tools and methods of modern science and the body of knowledge that could explain humanity's past and current functioning? After Darwin, however, this excuse is no longer tenable; by learning about our evolutionary origins and understanding the "universal acid" that is the theory of natural

G. Gorelik (✉)
Department of Psychology, Florida Atlantic University,
Boca Raton, FL 33431, USA
e-mail: gregory.gorelik33@gmail.com

T.K. Shackelford
Department of Psychology, Oakland University, Rochester, MI, USA
e-mail: shackelf@oakland.edu

© Springer International Publishing Switzerland 2016

T.K. Shackelford, R.D. Hansen (eds.), *The Evolution of Morality*,
Evolutionary Psychology, DOI 10.1007/978-3-319-19671-8_4

shackelf@oakland.edu

selection (Dennett, 1995), we have taken on the responsibility of applying this knowledge to the betterment of humans and, perhaps, sentient nonhumans. Alas, generations after Darwin's *Origin*, broad swaths of the public, the media, and the academy (at least when it comes to the evolution of human mental life and behavior) are reluctant to embrace the very idea that could better the lives of individuals and societies.

Our definition of "better" can be discussed at two (sometimes irreconcilable) levels: (1) the level of individual well-being and (2) the level of societal well-being. At both levels, a proper understanding of evolution by natural selection (which includes a sophisticated understanding of human physical and mental traits) can promote self-knowledge, enhance social and economic prospects, and engender nurturing and thriving relationships. In this chapter, we argue for the implementation of evolutionary knowledge as a metacognitive framework—that is, as a framework that can inform our individual and collective decisions by setting up an overarching perspective from which we can examine our thoughts, emotions, and behaviors. We term this framework "evolutionary awareness." First, we describe the moral foundations upon which our framework rests. Second, we present some possibilities for the application of our framework to several domains of human functioning. Although our discussion of such domains is not exhaustive, it is sufficient and general enough for others to widen this application to other domains. Finally, we discuss whether an evolutionary awareness may be reconciled with other fields of inquiry and propose the concept of *intergenerational extended phenotypes* to better understand the long-term effects of human decision-making.

Our approach rests on some of the latest findings in the evolutionary science of human behavior. These findings draw attention to several spheres of human interaction for which an evolutionary perspective can be used to identify the often "unconscious" roots of conflicts between individuals and societies. Our approach aims at a practical implementation of evolutionary principles for identifying some of the ultimate (in the historical sense) causes of conflict and lack of well-being. Unlike Wilson, Hayes, Biglan, and Embry's (2014) the framework of "intentional change," which focuses on bringing about positive individual and social changes by couching various tools of phenotypic plasticity (such as behavioral conditioning, cognitive therapy, and mindfulness meditation) within classic evolutionary theory, our approach seeks to identify the evolved, recurring patterns of human behavior—patterns that should be identified before any attempt is made to modify them by cognitive-behavioral or mindfulness approaches. It is important that we recognize our evolved abilities and limitations before we try to modify them. Indeed, totalitarian philosophies often fail because of their attempts to radically change human nature at the cost of human beings. In this sense, our discussion should be used to supplement Wilson et al.'s by highlighting those contexts of human interaction in which change is possible (or desired) and those in which it is not. We may come to find that it is sometimes easier to enact prosocial change by redirecting our evolved proclivities rather than resisting them.

shackelf@oakland.edu

The Ethical Dimension of an Evolutionary Awareness

We are always straddling the ethical divide—a ubiquitous dimension along which our actions are judged by others and ourselves to be either ethical or unethical. This divide is not always apparent. Indeed, it is often hidden beneath layers of selfishness, group favoritism, religious morality, legal obfuscation, consumerism, and worldly practicality. We often must actively search for the ethical dimension in all our experiences. Navigating the world of selfishness and group favoritism is much easier because we evolved to be sensitive to these phenomena. Likewise, religious triangulations and politically charged alliances—circumstances whose ethics is under the influence of an "evolved awareness"—are also easier to heed than is the call of the ethical dimension of an "evolutionary awareness." Once one acknowledges the relative supremacy of the naturalistic worldview and, more specifically, acknowledges one's status as a biological being that possesses strong tendencies toward self-preservation and reproduction, one is forced to face the ethical repercussions of such an acknowledgment. Although we have some misgivings about employing such a dualistic framework, it is almost as if every one of our actions—from the pouring of coffee (e.g., where were those beans procured and at what human and environmental cost?), to the enactment of life-goals (e.g., how will I be able to benefit my fellow humans if I devote myself to scientific writing as opposed to law or civil service?)—can be judged as contributing or not contributing to the well-being of other humans. It is up to us to develop an intellectual understanding that can tackle these ethical dilemmas in a scientifically cogent manner.

Because individual choices lead to cultural movements and social patterns (Kenrick, Li, & Butner, 2003), it is up to every individual to accept the responsibility of an evolutionarily informed ethics. Although we are confident that an evolutionary perspective can increase human well-being (both in an evolutionary sense of reproductive fitness and in a more general sense of human happiness, fulfillment, and meaning), we do not deny that this may come at some cost to our evolved reproductive interests, if only because of the time-consuming task of having to reexamine many of our automatically enacted cognitions and behaviors from such an overarching perspective as biological evolution (see Vallacher and Wegner, 1989) for a discussion of "action identification theory"). However, as with any intellectually motivated course of action, developing an evolutionarily informed ethics entails an intellectual sacrifice: are we willing to forego certain reproductive benefits or personal pleasures for the sake of building a more ethical community? Such an intellectual endeavor is not just relevant to academic debates but is also of great practical and ethical importance. To apply the paleontologist Simpson's (1951) ethical standard of knowledge and responsibility, evolutionary scientists have the responsibility of ensuring that their findings are disseminated as widely as possible. In addition, evolutionarily minded researchers should expand their disciplinary boundaries to include the application of an evolutionary awareness to problems of ethical and practical importance. Although deciphering the ethical dimension of life's varying circumstances is difficult, the fact that there are physical consequences

shackelf@oakland.edu

for every one of our actions—consequences on other beings and on the environment—means that, for better or worse, we are all players in constructing the future of our society and that all our actions, be they microscopic or macroscopic, are reflected in the emergent properties of our society (Kenrick et al., 2003).

Admittedly, the misuse of evolutionary thought by individuals and groups is a possibility, as is the likelihood that evolutionary research might be used to justify the "naturalness" of violence or infidelity. The preponderance of evidence for the existence of manipulation and deception in nature—from parasites and their hosts, to behavior-altering flukes, to ingeniously devious predatory spiders (Aidoo, Terlouw, & Kolczak, 2002; Carius, Little, & Ebert, 2001; Jackson, Pollard, Li, & Fijn, 2002; Libersat, Delago, & Gal, 2009; Ridley, 1993)—has made it clear that nature does not owe us either morality or immorality. That an understanding of evolved predispositions, motivations, and modes of behavior can be used by manipulative individuals to benefit themselves at others' cost is, of course, a possibility. In addition to the benefits of cooperative kin-based and tribal coalitions, human intelligence and sociality may have evolved because of the reproductive benefits of social deception and manipulation (Bailey & Geary, 2009; Brüne & Brüne-Cohrs, 2006). We know that our great-ape cousins engage in deliberate distortions of the truth. Lower-ranking chimpanzees, for example, will hide food from higher-ranking males and deliberately avoid being in the vicinity of the hiding places when in view of the higher-ranking males (Hare, Call, Agnetta, & Tomasello, 2000; Hare, Call, & Tomasello, 2001). Likewise, although our ability to place ourselves in others' shoes—a necessary precondition for the implanting of false beliefs in others—develops at around 4 years of age, our ability to spot deceivers develops much earlier (as would be expected if we accept that early-developing deceiver-detection abilities were necessary in an ancestral environment that was rife with liars and manipulators; see Flynn, 2006; Mascaro & Sperber, 2009). Beginning in preschool, we become progressively better at telling self-serving lies—i.e., lies that serve our evolved interests, be they avoiding punishment (Talwar & Lee, 2002, 2011) or acquiring a resource (e.g., procuring cards in a card game; Smith & LaFreniere, 2013). That these juvenile lies can develop into the prevarication and sophistry of politicians or the corporate doublespeak of business executives suggests that the misuse of scientific theories and findings is a real possibility (and, lest we forget, the use of science for nefarious ends such as systematic killing is attested to by history). Our hope is that by creating the conditions for the betterment of modern education in the realm of evolutionary science, a widespread realization that cooperation and reciprocity are preferable to manipulation and exploitation can take hold (Axelrod, 2006; Stewart & Plotkin, 2013). In what follows, we make a case for an ethical approach to the study of human evolution and its application to human behavior. To better understand what an evolutionary approach to ethics and morality entails, we first discuss the evolutionary history of human morality.

shackelf@oakland.edu

The Evolution of Human Morality

Haidt (2012) proposed the existence of six psychological mechanisms (or "foundations"), each with its own evolutionary history and upon which most of human morality is based. These foundations are universal, although there are individual and cultural differences in the degree to which they are expressed. One such foundation—the one that is most salient to secular liberals—is what Haidt referred to as the "Care/Harm" foundation. This foundation is responsible for the emergence of such sentiments as the preservation of happiness, pleasure, and longevity and the elimination of pain and suffering. In addition to the "Care/Harm" foundation, Haidt proposed the existence of a "Fairness/Cheating" foundation (i.e., proportionality between giving and taking) and a "Liberty/Oppression" foundation (i.e., the suspicion of selfish upstarts and despots). Although, collectively, these foundations have existed for millions of years of evolutionary history (could we raise children or have any type of intimate or cooperative relationship otherwise?), their valuation did not reach intellectual and political sophistication until the coming of Enlightenment thinkers such as Thomas Payne and Thomas Jefferson, and neither were they philosophically explicated until Jeremy Bentham's and John Stuart Mill's advancement of utilitarian systems of ethics (although these had their antecedents in Epicureanism and other philosophies).

Haidt also proposed the existence of a "Loyalty/Betrayal" foundation and an "Authority/Subversion" foundation, each responsible for the moralistic sentiments associated with adherence to one's in-group and its leaders, respectively. Politically, these foundations have been active in extolling conservative and reactionary ideologies—e.g., exemplified by phrases such as "One nation under God" and "God save the queen." Because our ancestors depended on loyalty to their in-groups and their leaders to survive and reproduce, we all possess these foundations. Conservatives, however, tend to espouse ideologies that are in line with these group and authority-related foundations (often centered on "sacred values"; Atran & Axelrod, 2008) to a greater extent than do liberals (although there is some overlap when it comes to the "Fairness/Cheating" foundation and its embracement by classical liberals and conservative libertarians).

Our ancestors have been battling parasites and pathogens such as viruses, bacteria, lice, ticks, mites, and vermin for millions of years. Indeed, it is the perpetual arms races between viruses and hosts that may have led to the evolution of sexual reproduction as an antiviral adaptation in our early ancestors (Hamilton, Axelrod, & Tanese, 1990; Ridley, 1993). In addition to our biological defenses against viruses and bacteria (e.g., B cells, antibodies, macrophages, etc.), we possess physiological adaptations that respond to threats of parasitism and contamination, such as disgust and the emetic response. We also possess many psychological adaptations that function to deter us from people, places, and situations that are associated with contamination (or, rather, would have been associated with contamination for our ancestors; Miller & Maner, 2011, 2012).

shackelf@oakland.edu

It makes sense that much of what falls within the realm of human morality—particularly, the moralizing of sexuality—stems from evolved mechanisms that motivate avoiding contamination. Accordingly, these evolved mechanisms are included in what Haidt refers to as the "Sanctity/Degradation" foundation of morality—usually manifested as an avoidance of out-group members and the strict imposition of sexual mores. In support of this, studies by Fincher, Thornhill, and colleagues have highlighted the extent to which a geographical region's pathogen load may explain the social and political circumstances of that region (Fincher, Thornhill, Murray, & Schaller, 2008; Letendre, Fincher, & Thornhill, 2010). Regions with a higher prevalence of parasites and pathogens—usually tropical and subtropical regions—also display the highest rates of ethnocentrism and xenophobia (often manifested as group-centered collectivist and conservative political ideologies; for alternative explanations of these findings, see Hackman & Hruschka, 2013).

Haidt's approach to the study of human morality is nonjudgmental. He argues that the Western, cosmopolitan mindset—morally centered on the Care/Harm foundation—is limited because it is not capable of processing the many "moralities" of non-Western peoples. We disagree with this sentiment. For example, is Haidt really willing to support the expansion of the "Sanctity/Degradation" foundation (and its concomitant increase in ethnocentrism and out-group hostility)? As Pinker (2011) noted, "…right or wrong, retracting the moral sense from its traditional spheres of community, authority, and purity entails a reduction of violence" (p. 637). We would be falling prey to the naturalistic fallacy if we failed to critically examine the (reproductive and nonreproductive) utility of our evolved morality in today's world. It is for this reason that we advance a model that is centered on self-criticism and the occasional circumvention of our evolved psychology. Our model is compatible with Haidt's model, but with the coming of medical and technological innovation and internationalist sentiments, moral foundations based on notions such as "sanctity," "purity," and blind commitments to sectarian interests and charismatic leaders are harming us more than they are helping (as instances of ethnic cleansing, genocide, and the waging of group-perpetrated religious and patriotic wars can attest). Thus, we admit to embracing our secular and liberal biases and of valuing some of our evolved moral foundations—namely, the utilitarianism, anti-totalitarianism, and libertarianism of the "Care/Harm," "Liberty/Oppression," and "Fairness/Cheating" foundations—over others. The model that we propose aims to upregulate the activity of these humanitarian (and humanistic) moral foundations while downregulating all others. That, at least, is the goal; how we are to go about regulating these aspects of evolved human morality is the hard part, and, as previously mentioned, handling nature requires the utmost caution.

We stress the distinction between our advocacy of the use of evolutionary science to inform the ethically fraught decisions of individuals and the use of evolutionary findings to structure society via collective action. Whereas the former approach helps individuals to clarify the nature of humanity's evolved biases and their consequences on other individuals and on society, the latter can potentially further the worst of human nature if enacted too hastily. This is because the structuring of

shackelf@oakland.edu

human morality based on evolutionary principles is subject to misuse by those who have their own conception of human "nature." Even if our conception of this nature is scientifically accurate, we would be succumbing to the naturalistic fallacy if we were to unreflectively advance our evolved predispositions as social policy. Thus, our concept of "evolutionary awareness" invokes only some of humanity's evolved moral foundations. Our ethical stance is somewhat divorced from simple considerations of survival and reproduction and is, instead, a continuously evolving combination of both biological and cultural concerns. No doubt, our ethical system is partly based on the application of certain Enlightenment values to ensure that individuals have the right to survive and reproduce (as opposed to forcing individuals to live according to various religious or political dogmas). Nevertheless, individual survival and reproduction is often furthered by the violent intrusion by some individuals onto the lives of others. As discussed by Pat Shipman in *The Evolution of Racism* (1994), the misuses of Darwinism in the nineteenth and twentieth centuries to further racist colonial policies and to legitimize eugenic programs of forced sterilization of immigrants and the mentally ill should not be downplayed. Nor, however, should our fear of the misuse of science hinder us from unflinchingly examining ourselves as the biological beings that we are and using the fruits of such an examination to formulate courses of action aimed at bringing about individual and collective well-being. We admit to not having any clear-cut answers to the ongoing debates between individual and collective interests; nor do we have any hard-and-fast rules for applying deontological versus consequentialist systems of ethics to situations possessing contradictory moral alternatives. Instead, we hope that our subsequent discussion of the moral ramifications of our evolved nature might help individuals to make their own decisions in a context-dependent manner, knowing full well that many of our prescriptions for conduct are subject to change as situations themselves change and as more knowledge is gained.

Applications of Evolutionary Awareness

The "evolutionary awareness" concept has had a number of precedents in science and academia, reaching, perhaps, as far back as the Enlightenment. What distinguishes it from the giants upon whose shoulders it stands, however, is its unprecedented potential to extend the biological and psychological sciences to the individual realm—a realm with social consequences. For the first time in history, we are capable of understanding our biological and psychological nature and using this understanding to inform our ethical deliberations on personal and societal levels. Unlike the failed utopias of state-instituted eugenics and Social Darwinism, evolutionary awareness—and other scientific/naturalistic approaches to ethics—empowers *individuals* to make informed choices from a perspective that is enlarged by deep history and ecological awareness.

Below, we discuss the utility of applying evolutionary thinking to domains such as sexual maturation, mate choice, intrasexual competition, and human culture to

shackelf@oakland.edu

(1) argue for the relevance of evolutionarily informed research on human psychology and behavior and (2) identify some of the ways in which an evolutionary awareness can be used to navigate the ethically charged atmosphere of individuals and societies.

Evolutionary Awareness and Sexual Maturation Sexual maturation is a chaotic and emotionally stressful period in development. We are suddenly plunged into a world of novel drives, desires, thoughts, and social interactions. Navigating this world was no doubt stressful for our maturing hunter-gatherer ancestors. For the most part, their sexual maturation involved struggles with efficiently appraising their own and others' sexual attractiveness, battling sexual competitors, learning the courtship rituals of their culture, and enacting reproductive strategies that might have included acquiring short-term reproductive partners, investing in one long-term partner, or engaging in a dual mating strategy (see Pillsworth & Haselton, 2006) of securing a stable and nurturing long-term partner who could help to rear offspring while also pursuing short-term liaisons.

Although sexual maturation involves less intrasexual violence over mating rights in the Western societies of today than in ancestral and some modern societies (Kruger, 2010; Kruger & Nesse, 2007), the preponderance of evolutionarily novel cultural and technological tools and innovations has made the pubertal transition no less confusing, even if less dangerous. Most of us grew up in such modern environments and, aside from the few of us who were endowed with traffic-arresting good looks or with a precociously developing intuition about sexual and romantic matters, were left to navigate the labyrinth of sociosexual development without help from parents or teachers. In addition, most of us were bombarded with mixed messages during development, including the sexualized images permeating the media, the propaganda of anti-sex religious conservatives, and the biology-denying political correctness of gender feminists. It is thus no surprise that many took a longer time to reach psychological maturation in the mating arena than others. With the aid of an evolutionary understanding of human sexuality (and its dependence on developmental milestones as tracked by biologists and psychologists studying life history transitions; Del Giudice, Angeleri, & Manera, 2009; Ellis, 2004; Hill & Kaplan, 1999), we can better equip adolescents and adults to engage in physically and psychologically healthier, more emotionally fulfilling, more intellectually stimulating, more self-awareness inducing, and, ultimately, more ethical sexual and romantic practices.

When not assaulted by conservative activists calling for "parental rights" in matters of sex education, the educators of safe-sex practices and contraception use have been effective in reducing the spread of sexually transmitted diseases and unwanted teenage pregnancies (Chin et al., 2012; Kirby & Laris, 2009). However, sex education in public schools could be supplemented with evolutionarily informed, age-appropriate information pertaining to the psychological aspects of human sexuality. Below, we provide a sketch of what such an educational environment might entail.

First, we must recognize that to impart evolutionarily cogent knowledge about human sexuality to juveniles and adolescents, we must understand the evolution-

shackelf@oakland.edu

imposed opportunities and limitations to the teaching and learning of any subject. Geary (2007), one of the foremost proponents of an evolutionarily informed approach to education, argues that children are biologically primed to acquire some forms of knowledge more easily than others. For example, although even human infants and some nonhuman primates may be equipped with a nascent or rudimentary understanding of quantity and simple arithmetic (Beran, 2004; Wynn, 1992), extensive and explicit schooling is required for the development of more complex mathematical abilities. Geary terms these contrasting abilities as "biologically primary" and "biologically secondary," respectively, and argues that educational practices must take into account the relative ease with which children can acquire knowledge that invokes the former as opposed to the latter.

Similarly, sex educators should be cognizant of the evolved opportunities and limitations of juveniles and adolescents with regard to understanding certain topics. Fortunately, sexuality and romantic life invoke many of our biologically primary abilities; much of the information, however, may not make sense unless children have reached certain developmental milestones (i.e., life history transitions such as puberty or menarche). Educators should anticipate these transitions and, if possible, develop systems by which these transitions can be made less turbulent and involve less ignorance than is currently experienced by coming-of-age youths. Furthermore, although sexuality is perhaps more interesting than differential calculus, adults who try to direct children's dating and sexual behavior might be met with resistance or hostility. To counteract such possibilities, educators might need to implement a mixture of implicit, student-directed, and group-oriented learning practices with explicit instruction, as children are more accepting of information and influence coming from their peers than from their parents or teachers.

It is during the primary and secondary school years that children undergo the major physiological and psychological upheavals that are associated with puberty. As it stands now, parents, teachers, and school administrators might not be well equipped to deal with such changes. Being as how children spend many of their waking hours in classrooms and school playgrounds and lunchrooms, biological ignorance on the part of well-meaning adult caregivers is inexcusable. Preventing boys' playground tussles over social status, averting their distractedness by their physically developing female peers' bodies, trying to stifle girls' indirect aggression, and curbing their naïve manipulation of others via their newly discovered sexuality are as useless as trying to prevent the anatomical and physiological changes associated with puberty. Indeed, the hormonal effusions associated with such biological transitions do not stop at the blood–brain barrier. Parents, teachers, and school administrators must therefore anticipate and properly educate children about these psychological changes, in addition to their focus on reading and writing—and perhaps these traditional subjects might be combined with evolutionarily informed lessons on biological and psychological development in new and creative ways. In what follows, we discuss two arenas where such an approach might prove successful: boys' rise in aggressiveness following adrenarche and girls' susceptibility to early puberty following paternal neglect.

shackelf@oakland.edu

According to life history theory (Del Giudice, 2009; Del Giudice, Angeleri, & Manera, 2009; Ellis, 2004, 2005), organisms have evolved to apportion their metabolic resources in an adaptive manner across the lifespan. For example, some species (e.g., elephants and humans) apportion most of their metabolic resources toward the production of only a few offspring that receive a great deal of parental care, whereas other species (e.g., mice and cockroaches) apportion most of their metabolic resources toward the production of many offspring that receive little parental care. Neither strategy is better or worse than the other, as each is associated with its own reproductive advantages and costs. Similarly, there is variation in life history strategies within individual species (though not to the same extent as the variation between species). Finally, organisms are often sensitive to the social and ecological cues of their developmental environments, cues which may be used to gauge the prospects of success for a number of different reproductive strategies. Such "conditional" adaptations are often influenced by factors such as the degree of nutritional resources or parental care encountered by the organism (Ellis, 2005).

Different sexes within a species often exhibit different life history strategies. For example, human males' reproductive success is limited by human females' mandatory investment in pregnancy and childcare, which has led to the evolution of men's greater propensity to engage in competition with other men over access to women (Trivers, 1972). Evidence suggests that this propensity appears early in development; specifically, it is with the influx of adrenaline during the juvenile transition (a period roughly corresponding to 7–10 years of age, referred to as "adrenarche") that children's behavior becomes sexually differentiated—e.g., boys become more interested in hierarchical competition with other boys and girls become more interested in childcare (Del Giudice, Angeleri, & Manera, 2009). More generally, such hormone-mediated developmental shifts are what West-Eberhard (2003) refers to as "switch points" that correspond to the major transitions in morphology and behavior that organisms undergo as part of their life history. Anticipating this rise in boys' aggressiveness levels might help to moderate the adverse effects of this life history transition for the boys themselves and for other children who might be adversely affected by such changes in boys' behavior. For instance, steps might be taken to introduce more healthy forms of competition, and for boys who may end up on the losing end of such bouts, alternative domains of competitive success might be suggested. Likewise, nonviolent team-building exercises (be they sports or video game related) might be used as positive outlets for boys' evolved need to belong to competitive coalitions with other boys. The key, again, is to redirect children's evolved tendencies into more prosocial forms of behavioral expression rather than trying to stifle these evolved tendencies whenever they emerge as socially undesirable behaviors.

As they approach puberty, girls experience similar hormonal transitions as do boys; however, there is evidence for a unique influence of paternal care (or lack thereof) on girls' pubertal timing, or menarche (Draper & Harpending, 1982; Ellis, 2004, 2005). Specifically, although nutritional stressors seem to delay the onset of menarche in girls, psychosocial stressors such as father absence seem to expedite it. Girls seem to be adaptively sensitive to the amount of paternal investment they

shackelf@oakland.edu

experience in their early environments because the level of paternal investment experienced might cue them as to how likely they are to experience commitment and childrearing assistance from men in general. That is, if a girl grows up without a father (or with a non-investing father), it might be to her advantage to forgo investing in the prolonged development of attributes that are helpful in securing a stable long-term mate and co-parent (attributes such as conscientiousness, self-control, and educational attainment), if the men in her environment have the same commitment and parenting deficiencies as her father. Instead, signals from her psychosocial environment might initiate a conditional hormonal response aimed at an earlier pubertal onset and, hence, an earlier and more frequent reproductive career. Such a shift to a faster life history strategy might be just as biologically adaptive as an investment in long-term mating and parenting, though in the modern environments of today, teenage pregnancy and a lack of education might subsequently lead to a life of ill health and poverty for the girl and her potential offspring.

The question of what—if anything—can or should be done to alter the course of girls' sexual maturation is one that is fraught with social and political implications. First, there is the question of whether early sexual maturation is indeed a problem. Although teenage pregnancy and a lack of long-term investment in educational and career opportunities are undesirable, infringing on a young girl's sexual development or a young woman's evolved sexual strategy might be a recipe for totalitarianism. Who is the government or the public school system to decide on the social desirability of one sexual strategy over another? Second, there is the possibility that girls from some racial or socioeconomic groups might be more affected by such developmental interventions than others. Just as the early 1990s Violence Initiative was met with misunderstanding and hostility stemming from the racially charged nature of its supposed focus on young, mostly Black inner-city youths (Shipman, 1994), any biologically driven approach to social policy is bound to be passionately opposed by political interests that detect a racial imbalance in its implementation. Third, and perhaps most importantly, there is no clear method for effectively altering the pubertal timing of girls. Although policymakers might be tempted to substitute male role models and father figures for absentee biological fathers, it is possible that the presence of a biologically unrelated male (e.g., a stepfather or a mother's boyfriend) might actually speed up girls' sexual maturation (Ellis, 2005), not to mention increase the risk of abuse or neglect (Tooley, Karakis, Stokes, & Ozanne-Smith, 2006). Economic incentives aimed at preventing fathers from abandoning their daughters might hold some potential, though debates over who pays for such programs might keep them from ever being implemented in today's politically polarized environment. Similarly, social or cultural movements focused on instilling values such as paternal commitment and investment might encounter political opposition over the singling out of some racial or socioeconomic groups over others—and it is not like such attempts at cultural change would be entirely new.

Although recent estimates show a substantial decline in teenage pregnancies (Hamilton & Ventura, 2012), their toll on young women's health, socioeconomic status, and educational opportunities necessitates that paternal investment and other biological influences on girls' pubertal timing are not ignored. We believe that a

shackelf@oakland.edu

middle ground should be sought between respecting young women's developmental and sexual autonomy and making sure that they have the educational and economic opportunities that they otherwise would not have if they were to get pregnant. As previously mentioned, a comprehensive sexual education that stresses both contraception and evolutionarily grounded biological and psychological instruction might be able to offer some solutions.

Any evolutionarily informed sex education course should include (but not be limited to) topics such as physical and psychological sex differences, violent male–male competition, relational aggression between females, female choosiness and male readiness to engage in sexual behavior, the evolution of sexually transmitted parasites and pathogens, within-sex and between-sex differences in life history strategies, and the social interactions involved in parenting. Assuming the development of appropriate educational techniques, educators may also target students who are lacking in self-esteem or confidence—or who may be bullied by their more dominant peers into subordination—for programs of skill and confidence enhancement in the dating arena (e.g., students may be encouraged to develop a talent or to engage in better hygiene or sartorial practices as opposed to engaging in violence or sexual coercion). Far from being a carte blanche for permissive sexual practices, such an educational program may be effective in battling the ignorance that comes with adolescent violence and unsafe, impulse-driven sexual behavior and may give students the confidence to resist maladaptive peer and cultural influences.

Evolutionary Awareness and Mate Choice The realm of mate choice presents opportunities for the enlightened extension of evolutionary understanding to problems of ethical concern. There is now a preponderance of research on why we find some individuals more physically attractive than others. Bilateral symmetry (i.e., a close correspondence between the left and right sides of the body and facial structure), for example, is valued as a mate characteristic by animals as diverse as humans, sparrows, and fruit flies (Brown et al., 2008; Gangestad, Thornhill, & Yeo, 1994; Hunt, Crean, Wood, & Gilburn, 1998; Moller, 1992). This is because symmetrical structures communicate a developmental history unperturbed by harmful genetic mutations, parasites, and other developmental stressors. Indeed, symmetry is related to a host of practical benefits such as balance, efficiency of movement, and—especially in humans—intelligence (Bates, 2007; Furlow, Armijo-Prewitt, Gangestad, & Thornhill, 1997; Prokosch, Yeo, & Miller, 2005). For similar reasons, "averageness" of features (i.e., not possessing an appearance that deviates overwhelmingly from the population average) is another characteristic that is correlated with attractiveness (Apicella, Little, & Marlowe, 2007). We might also do well to acknowledge the effects that hormones and other physiological processes have on our cognition and behavior. For instance, women's mate preferences change as a function of their ovulatory cycle status (see Gangestad, Thornhill, & Garver-Apgar, 2005, for review). More to the point, women prefer dominant "hunks" when they are at peak fertility (presumably because such men are genetically well endowed), but prefer men who promise long-term commitment, empathy, and paternal care when they are in the lower fertility phase. Acknowledging these influences on our

shackelf@oakland.edu

mating decisions (in addition to their influence on how we treat friends and colleagues) should cause us to reexamine some of our implicit judgments, attitudes, and actions.

Characteristics such as physical attractiveness and dominance not only influence the treatment of individuals in our immediate vicinity but also affect large-scale societal outcomes such as political elections. There is a growing body of research on the effects of physical characteristics such as candidate attractiveness, dominance, and voice pitch on the likelihood of election. For example, Little, Burriss, Jones, and Roberts (2007) found that people are more likely to cast hypothetical votes for masculine-faced candidates during wartime but were more likely to vote for feminine-faced candidates during peace time. The authors propose that this is due to an evolved propensity for humans to choose group leaders who are differentially suited to various group-related circumstances. In line with this prediction, candidates with lower-pitched voices and facial characteristics indicative of greater height are more likely to be chosen as leaders during wartime scenarios over candidates with higher-pitched voices and faces indicative of shorter stature (Re, DeBruine, Jones, & Perrett, 2013; Tigue, Borak, O'Connor, Schandl, & Feinberg, 2012).

These findings suggest that humans may be prone to making suboptimal decisions when choosing their leaders and that an evolutionary awareness might motivate us to reexamine our political decisions—decisions that can mean the difference between life and death on a global scale. Although Little et al. (2007) state that there may be "'a kernel of truth' in judgements made based on appearance and the belief that faces do provide valid guides to character" (p. 26), they also contend that because our decision-making psychology evolved in ancestral—rather than current—environments, "[i]ndividuals appear to not consider aspects of large-scale technology-driven warfare and [instead] make the best choice for small-scale intergroup conflict" (p. 26).

Evolutionary Awareness and Intrasexual Competition Intrasexual competition includes many of the defining elements of ethically charged human behavior. From individual-level violence to warfare, deception to exploitation, male–male competition is rife with the gravest of personal and social issues. This is evident not only in cases of physical violence between men who are competing for women (such as Aché tribesmen banging each other over the head with clubs) but also in cases of economic and resource-based competition. Much of socioeconomic inequality, for instance, can be explained as the indirect effect of men's competition over resources. Men who participate in a lab-based resource allocation task, for instance, are less likely to apportion resources to other men than they are to women, whereas women do not sexually discriminate in their generosity (Buunk & Massar, 2012). Although this type of competition is usually nonviolent, it can lead to the sequestration of environmental and monetary resources into the hands of the few, to the detriment of the many. Economic imbalance per se is not unethical (but see Daly, 2010) for the effects of economic inequality on societal violence), but a man aspiring to ascend the social-status hierarchy of the corporate world must choose whether he wants to

shackelf@oakland.edu

pursue the course of honesty, cooperation, and philanthropy or cheating, backstabbing, and white-collar crime. Of course, considerations of corporate decision-making have long had an ethical dimension, but what is missing from most discussions of business ethics is an evolutionary dimension—a dimension that not only informs individuals about the origins of psychological mechanisms responsible for unethical business practices, such as various forms of cheating (Cosmides & Tooby, 1992), but also suggests avenues for its reduction. Thus, although men may still be tempted to fudge tax documents and bribe politicians, understanding that these behaviors are produced by evolved motivations to survive and reproduce may serve as a guide toward more ethically sound survival and mate-acquisition tactics.

Male–male competition, whether a barroom brawl or a corporate standoff, may lead to significant social consequences. Researchers have tracked changes in men's testosterone (a hormone that is positively associated with male pugnacity and sex drive; Bernhardt, 1997; Rupp & Wallen, 2007) as a function of male–male competition. Men exhibit an increase in testosterone when they engage in competitive activities—especially athletic activities (Trumble et al., 2012). Moreover, male winners (be they chess players, basketball players, or the supporters of a winning sports team or political candidate) exhibit an increase in testosterone (especially if they feel that they personally contributed to the victory), whereas male losers exhibit a decrease (Bernhardt, Dabbs, Fielden, & Lutter, 1998; Gonzalez-Bono, Salvador, Ricarte, Serrano, & Arnedo, 2000; Mazur, Booth, & Dabbs, 1992; Stanton, Beehner, Saini, Kuhn, & LaBar, 2009). Because social status is an important criterion on which women judge men's mate value, male winners may upregulate their production of testosterone—and, hence, their sex-drive—in expectation of increased mating opportunities. This phenomenon is highlighted by the finding that men who are randomly assigned to drive luxury automobiles experience an increase in testosterone, whereas men who are given dilapidated clunkers experience a decrease (Saad & Vongas, 2009)—a physiological effect that may be reflective of the finding that women are more attracted to drivers of luxury cars than they are to drivers of more modestly priced vehicles (Dunn & Searle, 2010). A series of studies that tracked the Internet behavior of US voters likewise found that US states that overwhelmingly voted for winning candidates experienced a greater increase in Internet pornography searches, whereas states that voted for losing candidates experienced a decrease (Markey & Markey, 2010, 2011). Because most consumers of pornography are male, this finding would suggest that men who are the vicarious winners of political contests may upregulate their sex drive in anticipation of increased mating opportunities—even if those opportunities do not materialize.

There are at least three reasons why men do not maintain a state of high testosterone indefinitely (which would presumably be an optimal strategy in a world where sexual opportunities are always present): (1) the dangers associated with competing with men of greater social status, (2) the possibility of squandering one's time and resources on women who are uninterested in low-status men, and (3) the health-related costs resulting from the immunosuppressive effects of testosterone (Alonso-Alvarez, Bertrand, Faivre, Chastel, & Sorci, 2007). As suggested by the "challenge hypothesis" (a concept initially applied to an evolutionary biological

shackelf@oakland.edu

study of bird behavior but later extended to mammals and primates; see Archer, 2006), males experience an increase in testosterone and other androgenic hormones during the breeding season or—as in many primate species—during exposure to sexually receptive females. Because androgens are associated with male pugnacity and sex drive, it is thought that these typically "male" hormones are responsible for physiologically and behaviorally readying males for increases in both competition with other males and sexual access to estrous females (Sobolewski, Brown, & Mitani, 2012). For example, in males of the African cichlid *Astatotilapia burtoni*, hypothalamic neurons responsible for the release of gonadotropin-releasing hormone (GnRH)—a necessary step in the production of testosterone by the gonads—undergo a period of growth upon a male's rise in social status (Fernald, 2007). However, because of the costs associated with violent intrasexual competition and a decrease in the opportunity to maintain a territory and, hence, to attract mates, GnRH neurons decrease in size upon a male's descent in the dominance hierarchy (Francis, Soma, & Fernald, 1993; White, Nguyen, & Fernald, 2002).

The historical costs of competition between men are well documented and include, for example, homicide, genocide, war, and economic inequality (the latter of which cannot be doubted given the preponderance of men in executive-level positions; see Daly & Wilson, 1988). A particularly insidious cost, however, involves men's propensity to link dominance-related aggression with sex drive. For a number of scientific and political reasons, academics can spend decades arguing about whether rape is produced by specialized adaptations. Its occurrence, however, remains a problem. That Nazi war crimes often were perpetrated with sexualized excitement (Theweleit, 1977/1987, cited in Pinker, 2011), or that Russia's 1944 March across Poland often included the rape of local women (Smith, 2007), suggests that a metacognitive evolutionary perspective may be important for preventing such atrocities in the future. Understanding the influence of male–male competition on individuals and societies is a start.

Evolutionary Awareness and Culture Prior to our discussion of how an evolutionary awareness can be applied to human culture, we must clarify what we mean by human "modularity" and "culture." Note that when we use the term "module" (or "evolved psychological adaptation," "mental process," "mental mechanism," "neurocognitive network," etc.), we are referring to any task-specific psychological system that is an evolved product of natural selection. We will not address the debate between massive modularity and domain generality except to say that many of our modules can perform a variety of different functions, are shaped by associative learning and experience (i.e., they are malleable), are composed of sub-modules and embedded in larger modules, may exhibit distributed network processes that are not topologically cogent across the neocortex, and depend on contextual and cultural influences for the development of their evolved functions (see Barrett, 2009; Fuster, 2003; Kurzban, 2010, for further discussion). By "culture," we refer to any material product, mode of behavior, or unit of knowledge that is not genetically inherited, can be transferred both within and between generations, and can progress by building on previously acquired products in a cumulative manner (see, e.g., Boyd & Richerson, 1996; Caldwell & Millen, 2008).

shackelf@oakland.edu

Human culture often reflects the dynamics of evolved modules. This reflection is not isomorphic; some modules may be inexpressible in human language or via other symbolic systems, and culture may exhibit emergent properties that have no analogues in the human mind. Even here, however, culturally "disconnected" modules may indirectly shape cultural products and information (e.g., color perception and its concomitant neuronal connections influencing art and fashion, sound perception influencing musical tastes, and gustatory and hunger-satiety mechanisms influencing cuisines). Whatever the limits of the correlation between evolved modular processes and culture, across evolutionary time, many of our modular processes were shaped by cultural processes (Boyd & Richerson, 2008; Herrmann, Call, Hernandez-Lloreda, Hare, & Tomasello, 2007), and conversely, the imprint of our modular processes is embedded in cultural history and in today's cultural environment (Sperber & Hirschfeld, 1999). Below, we present examples of how human culture may reflect the functioning of our evolved psychological processes.

Primates are among the most socially complex animals on the planet. Within the primate order, the great apes are the most socially complex family, and within the great-ape family, humans are the most socially complex species. Primates are known for their fixation on social coalitions, status hierarchies, and in-group–out-group dynamics. Subordinate chimpanzees, for example, can build and maintain coalitions with other subordinates for prolonged periods of time and use these coalitions to depose reigning dominant males in planned attacks of sometimes lethal violence (de Waal, 1996; Goodall, 1990). Likewise, chimpanzees are known to live in hierarchical communities that occasionally wage war on other communities over mates and verdant feeding patches (Wilson & Wrangham, 2003). For the most part, these battles involve raids by a coalition of about six males on another community's territory. These coalitions mostly attack lone, unprotected individuals, often by kicking, punching, biting, battering with weapons such as sticks and projectiles, and, occasionally, dismembering and mutilating the unlucky stragglers. With enough of these attacks, one group will extinguish another group and incorporate some of the defeated group's fertile females into its own community. Humans inherited these traits from our common ancestor with chimpanzees (who lived 5–7 million years ago) and have been annihilating out-groups of competitor hominins and impregnating their females ever since (our mating with and possible annihilation of the Neanderthals is one such example; Hortola & Martinez-Navarro, 2013; Yotova et al., 2011).

At the cultural level, humanity shows no dearth of coalitional markers and group identities—from religious, political, and corporate affiliations to sports team and rock-star fanbases. In modern environments, individuals may belong to social coalitions without ever interacting directly with other coalition members, as in online communities and other social media. If our coalitional nature is a product of millions of years of natural selection on our ancestors' brains and, conversely, if our brains are a product of millions of years of coalition building, it is unsurprising that coalitional markers are so important in the history and current milieu of human culture. That is, our evolved coalition modules may culturally express themselves in multiple ways.

shackelf@oakland.edu

Male proprietariness of female sexuality, for instance, may be culturally expressed via patriarchal religious beliefs and practices such as opposition to abortion; the covering up of the bodies of wives, daughters, and sisters; and lethal violence that is referred to as "honor killing." Many of these practices stem from evolved mental mechanisms that motivate men to act proprietarily toward their wives, daughters, and sisters. One such mechanism, which is directed toward wives and lovers, leads to the deployment of mate guarding tactics such as monitoring of women's whereabouts and associations and more severe behaviors such as cloistering, physical and sexual violence, and even homicide (Wilson & Daly, 1998). There is evidence that such behaviors are produced by the evolved motivation of men to prevent cuckoldry and the associated reproductive costs of raising another man's offspring. Similar mechanisms associated with male proprietariness can be directed toward the sexual control of daughters and sisters, as men's genetic interests may be threatened if their female kin—i.e., individuals with whom they share genes—mate with men who are disapproved of by their fathers or brothers. Cultural and religious beliefs and practices may exacerbate such expressions of proprietariness and coercion by making it a "duty" to defend one's own or one's family's "honor" by punishing the "impurity" of one's wife, lover, daughter, or sister. Such motivations are often highly moralistic and may additionally invoke the "Sanctity/Degradation" foundation of morality. That the "Sanctity/Degradation" foundation is invoked may explain the coalitional nature of patriarchal cultures and religions whose moralistic disgust unifies them against the threat of sexually impure, pathogen-carrying outsiders wishing to impregnate their wives, daughters, and sisters.

On the other side of the sexual divide, women also may coalesce into cultural groups that collectively represent their reproductive goals. Much like the female coalitions of our bonobo cousins, human females may compensate for their lack of physical and political strength by congregating in all-female coalitions that further their evolved interests. For example, it is hypothesized that because of the incessant threat of sexual coercion and rape across our evolutionary history, women may have evolved rape-avoidance mechanisms, such as avoidance of strange men, avoidance of appearing sexually receptive, avoidance of being alone, and being aware of one's surroundings and engaging in defensive behaviors (McKibbin et al., 2009). Such a psychologically salient awareness of the possibility of rape may be partly responsible for the feminist movement's struggle against "rape culture" and its concomitant expressions of male misogyny and hypersexuality. Thus, certain forms of feminism may be the cultural manifestations of women's rape-avoidance modules. The evolutionary history of these modules is rooted in women's greater parental investment in gestation, lactation, and child rearing—factors that contribute to women's greater need for physical safety and sexual choosiness compared to men.

Women may find that their vigilance with regard to the threat of rape is physically, psychologically, and economically taxing and, indeed, some feminist movements are attempting to put the onus on men to avoid raping women, rather than focusing on women's rape-avoidance tactics (Williams, 2013). Although educating men about women's interests in avoiding rape and sexual coercion is important, to be effective, this education must be informed by evolutionary concepts such as

shackelf@oakland.edu

differential parental investment and sexual conflict (Shackelford & Goetz, 2012). The latter concept describes a coevolutionary arms race between the sexes whereby an adaptation in one sex (e.g., men's sexual aggressiveness) evolves at a cost to the other sex (e.g., women's vulnerability to men's sexual aggressiveness). The other sex, in turn, must evolve counter-adaptations (e.g., women's rape-avoidance mechanisms) to defend its reproductive interests. Thus, the cultural struggle between patriarchal institutions and feminism may be a macroscopic manifestation of individual-level sexual conflict. Furthermore, in addition to its value in educating men about their aggressive impulses, the concept of sexual conflict suggests that an abandonment of focus on women's rape-avoidance tactics—as advocated by some radical feminists (MacDonald, 2008)—may endanger women by causing them to ignore their evolved defenses against rape and sexual exploitation. Although we should never excuse the blaming of women for their victimhood, neither should we expect that all men will suddenly stop raping and sexually coercing women. Therefore, it is equally inexcusable for politically correct interests to eviscerate women's last line of defense against rape and sexual coercion.

Evolutionary Awareness and Postmodernism Not to wax ecumenical, but perhaps some of the critical tenets of postmodernism can at last be reconciled with the core tenets of evolutionary science. From the postmodern perspective, Darwinian thinking is (wrongly) considered as just another narrative—perhaps a "Eurocentric" structure with its own signs and symbols, often espoused by White, well-to-do men seeking to maintain their privileged status. Many Darwinian thinkers, in contrast, view postmodernism as unscientific, academically legitimized obscurantism, often with intellectual and moral relativism as motivating factors. Despite the animosity within the ivory tower and the likelihood of there being irreconcilable modes of approaching knowledge in these disciplines, the evolutionary biological study of animal signaling and communication may help to forge an interdisciplinary bridge between evolutionary science and the empirically cogent aspects of postmodernism.

According to animal signaling theory, signals such as dogs' barks, lions' growls, birds' songs, or chimpanzees' hoots evolve only if they contribute to an organism's survival and reproduction. Organisms that exhibited such signals must have survived and reproduced more successfully than organisms that did not exhibit them. In turn, these signal exhibitors bequeathed their signaling abilities to their offspring. Over time, a population of non-signalers became a population of signalers within which each individual organism used its signaling abilities to enhance its survival and reproduction. For example, vervet and colobus monkeys, along with other primate and bird species, have evolved the ability to produce alarm signals that warn their groups about dangers (Leavesley & Magrath, 2005; Schel, Tranquilli, & Zuberbuhler, 2009). In the case of vervets (*Cercopithecus aethiops*), these signals are specific to the type of predator observed (e.g., eagle, leopard, or snake) (Seyfarth, Cheney, & Marler, 1980). The effect of such signals is, in turn, predator specific. For example, if the alarm call forebodes an eagle attack, the monkeys seek shelter in tangled bushes while continuously scanning the sky for signs of the intruder. It is

shackelf@oakland.edu

easy to imagine the evolutionary benefits of such signaling; because most primate troops are composed of extended kin (i.e., individuals who share genes with the signaler), signalers are helping to spread the signaling behavior by benefiting the very individuals who will likely pass this behavior on to future generations, even if the signaler may be imperiled by drawing the predator's attention to itself. Another benefit of signaling may stem from the diffusion of costs among unrelated signalers—that is, if the ratio of signaling costs to signaling benefits is such that the risks of producing an occasional signal are outweighed by the benefits of shared signaling among unrelated group members (an example of delayed reciprocity), signaling adaptations can be favored.

Of course, if we acknowledge nature's duplicity, then we should not be surprised that such signaling behaviors are often wielded to nefarious ends. Subordinate tufted capuchins (*Cebus apella*), for instance, produce false alarm calls when dominants are on the verge of monopolizing a food resource (Wheeler, 2009), causing the latter to scurry away in fear of a nonexistent snake or ravenous feline. In a similar manner, male Formosan squirrels (*Callosciurus erythraeus thaiwanensis*) emit deceptive alarm signals (usually meant to signal the presence of terrestrial predators such as feral cats) after copulating with females (Tamura, 1995). Because such signals have the effect of freezing other males in their tracks, signalers thereby increase their reproductive success by prolonging their access to the females which, in turn, reduces the risk of sperm competition from rival males. It is easy to see the parallels between these deceptive animal signals and the fear-inducing political messages coming from the politicians and media demagogues of our own species. Falsely claiming that others pose a threat to one's existence may indeed be an effective method of societal control. The dehumanization of Jews as disease-carrying vermin by Nazi propagandists (Smith, 2011) and the scare tactic of using nonexistent weapons of mass destruction as a pretext for the 2003 US invasion of Iraq are handy examples of such deceptive fear mongering in humans.

Do humans produce deceptive signals? The answer is a resounding yes!—not only because of the prevalence of deception in childhood, as previously mentioned, but also because of the high likelihood that our propensity for language creates innumerable opportunities for manipulation and exploitation. For example, Mercier and Sperber (2011) proposed that human reason did not necessarily evolve because our ancestors used it to arrive at accurate knowledge or to make good decisions but because it helped them to win arguments. By winning arguments, our ancestors could rise in social status and enjoy the reproductive benefits that came with this increase in prestige and influence. Thus, what are normally considered to be maladaptive errors in reasoning (such as the confirmation bias) might make sense if reason were deployed to support prior attitudes and motivations rather than to arrive at accurate knowledge or decisions. One can already imagine what some of the questionable consequences of the deployment of reason in collective settings such as politics, religion, science, academia, and in individual settings such as romantic relationships and friendships, might be (the Bay of Pigs and the Challenger and Columbia disasters come to mind). However, humans also have the capacity to evaluate others' use of reason, and although we are bound to make mistakes, collective

shackelf@oakland.edu

skepticism in science and philosophy may be reason's self-corrective antidote (Haidt, 2012).

Sperber (2010) likewise proposed the existence of the "guru effect"—a psychological propensity to believe that someone's obscurity or complexity in speech or ideas may be indicative of great wisdom or insight. This may mean that if executed correctly, incomprehensibility—especially alongside seemingly "reasoned" argumentation—might lead to an increase in one's social status and reputation for intelligence and profundity which, in turn, could lead to a self-reinforcing cycle whereby one's reputation increases as one's incomprehensibility deepens. Thus, a politician, pundit, priest, or professor may be esteemed by thousands—and sometimes billions, if we discount the professor—of people, even if none of them can understand or evaluate any of his or her beliefs or attitudes. Such a phenomenon is referred to as "pluralistic ignorance," a collective delusion that the king's obscurantism is covering his nakedness.

These examples highlight the ever-present danger of deceptive and manipulative signaling in the human species. Especially in postmodernist and poststructuralist circles (where obscurantism seems to enjoy a laudatory pedestal), but also in many scientific disciplines (it cannot be denied that some scientific publications are accepted by respected journals due to the author's prestige and the journal's deference to authority), it is important to reexamine previously held ideas and the sources for embracing them. To reexamine the content and sources of our signals, symbols, and ideas—what more could a postmodernist, critical-theory espousing semiotician wish for?

Haphazard criticism is not enough, however; it must be supplemented with a Darwinian framework that is sensitive to the as-yet-unexplored, evolutionarily derived sources of deception and misrepresentation. Much of the misinformation in the culture war between misogynistic patriarchal institutions on the one hand and scientifically uninformed gender feminists on the other, for example, may be exposed if we evaluate each side's arguments with the aid of evolutionary concepts such as parental investment and sexual conflict.

Intergenerational Extended Phenotypes

Organisms invest in reproductive resources cross-generationally. Indeed, parental investment is a "future-directed," intergenerational investment of reproductive resources in copies of one's genes that reside in one's offspring. Is it possible that humans might invest their reproductive resources in more than just one or two generations of children and grandchildren? Because of the human capacity to travel mentally in time (i.e., retrieving memories of the past and imagining the future; see Suddendorf & Corballis, 2007), humans are capable of waging evolutionary conflicts across more than just the immediate two or three generations (the actual number would have to be investigated by field or by modeling data). It is possible that intergenerational evolutionary conflicts may have been one of the sets of selection

shackelf@oakland.edu

pressures that drove the evolution of human intelligence, as evidenced by the gradual expansion of our ancestors' craniums over the past 2 million years. Thus, much like the distinction between fast and slow life history strategists (and perhaps somewhat reflective of them), there may be differences in the extent to which individuals are informed by the imagined prospect of the future when making reproductively relevant calculations. If such calculations were heritable, and provided they were conducive to an organism's reproductive success, the ability to perform them may have been naturally selected into the human lineage.

There are several hypotheses regarding the adaptive value of mental time travel. If, like us, our ancestors spent much of their time escaping the present moment, then such exhaustion of cognitive resources and inattention to one's current environment had to have given them a survival or reproductive advantage—otherwise, it would have been selected out of our species' repertoire (Bjorklund & Sellers, 2013). One possibility is that invoking neurocognitive networks associated with episodic memories helps humans to learn from past social and nonsocial mistakes to reach reproductively optimum decisions in the present and future (we are here excluding classically and behaviorally conditioned "memories" because, unlike our ability to mentally travel through time, they are not unique to humans; Suddendorf & Corballis, 2007). This is often accomplished via the formation of a "narrative self," perceived as an inner voice that is persistent through time and many of whose experiences can be verbalized (i.e., one can verbally describe one's past experiences and associate those experiences with one's stable self-identity). To prepare for the future, humans evolved the ability to imagine future possible worlds and outcomes. Such self-initiated "memories"[1] of the future may have helped our ancestors to anticipate climactic conditions, movements of prey and predators, and outcomes of familial and social relationships. Doing so could have helped them to avoid imagined futures that were unfavorable while working to bring about the ones that were favorable.

It would not be surprising if our ancestors, also like us, used their capacity for mental time travel to imagine the future successes and failures of their children—the very individuals who held their reproductive future on the line. When wedded to their propensity for the creation and preservation of cultural products such as beliefs, rituals, traditions, and modes of behavior, our ancestors' ability to remember the past and imagine the future of their family line may have been greatly expanded. For example, it is not unusual for hunter-gatherers to worship their distant ancestors—individuals who were thought to be present at the mythical time of creation. Examples of culturally expanded intergenerational thinking are likewise ubiquitous throughout the historical period. For example, the Bible is replete with long genealogical tracts of who "begat" whom and exhortations to be "fruitful and multiply." Likewise, the modern period has been held hostage to the Nazi dream of the

[1] This is not just metaphorical; humans may be employing some of the same neuronal pathways during both their experience of past memories and their imagining of possible futures (Botzung, Denkova, & Manning, 2008; Schacter, Addis, & Buckner, 2007).

shackelf@oakland.edu

"Thousand-Year Reich," the "historical materialism" of Marxist utopians, various millennial and messianic movements, second comings, and mythical caliphates.

Humans use cultural products as *extended phenotypes*. According to Dawkins (1982), an organism's phenotype is not confined to its bodily frame but can encompass any adaptive structure or environmental alteration (e.g., anthills, termite mounds, beaver dams) that was naturally selected to aid the organism's (or, rather, its genes') reproductive success. Such extended phenotypes can even include other organisms—of the same or of a different species—that are manipulated into benefiting the organism that uses them for its own reproductive interests. An example of such extended-phenotypic manipulation is the behavioral manipulation of an ant by a parasitic lancet fluke that causes the ant (by chemically altering its nervous system) to position itself atop vegetation that is vulnerable to being eaten by a grazing animal—the fluke's ultimate host. Is it possible that humans have evolved to use cultural products such as religious beliefs and group identities as intergenerational extended phenotypes? That is, could the human propensity for the manipulation of cumulatively acquired cultural products—say, a religiously motivated belief to be "fruitful and multiply"—be enacted as a way to ensure that one's progeny fruitfully multiply? Although the evidence for this is still lacking, there are some suggestions that such a dynamic may be at play.

Regardless of how conscious our ancestors were of the effects of cultural products, whether material or immaterial, on their and their offspring's reproductive success, there is evidence that many cultural trends throughout history were associated with changes in reproductive outcomes. (As an aside, such investigations suggest a mechanism for falsifying the present hypothesis; specifically, if it can be shown that individuals' or groups' cultural traditions preceded their genetic spread, such evidence would support—although it would not be decisive with regard to—the hypothesis that humans use culture to bring about their long-term reproductive success.) The first line of evidence comes from the advent of early Christianity. Scholars of the early Christian church believe that the historical Jesus was probably more provincial and Judeo-centric than he is characterized in the New Testament (Wright, 2009). Jesus' exhortation to "love thy neighbor as thyself," for example, is thought to have been originally restricted to fellow *Jewish* neighbors and that it was the Apostle Paul who was responsible for giving Christianity its universal appeal by spreading the gospel to downtrodden and impoverished Romans. Whatever the cause of its success, its adoption spread like wildfire across the Roman world, partly as a result of the close-knit ties and cooperative networks generated among its adherents (Wilson, 2002; Wright, 2009).

Although, to our knowledge, there has not been a systematic analysis of gene spread across early Christendom, it is interesting to speculate about whether beliefs in universal salvation and in-group cooperation—perhaps wedded to apocalyptic, future-oriented beliefs such as Christ's second coming—could have helped the early Christians to spread their genes in addition to their religion, much like the cultural innovation of keeping livestock helped to spread the gene for lactose tolerance alongside herding cultural practices. The sixteenth century rise of Calvinism and its concomitant beliefs in preordained salvation and humble industriousness—

shackelf@oakland.edu

beliefs associated with the "Protestant work ethic"—may have had a similar effect on the economic (and genetic) success of the American colonists (Wilson, 2002). Christianity, however, may not have always been a positive influence on the genetic success of populations. For example, Christianity may have indirectly led to the fall of the Roman Empire by pacifying its population into submission to the Vandals (Frost, 2010), as well as the fall of the early Viking settlers in Greenland to "pagan" Inuit invaders (Diamond, 2005)—two outcomes that collectively highlight the occasional inefficiency (from a gene's perspective)—of cultural evolution.

The effect of evolved reproductive strategies on political and religious beliefs and behaviors is a vibrant field of investigation. Single women in the fertile phase of their ovulatory cycle, for example, are more likely to endorse political and religious liberalism. This effect is reversed for married women—that is, they tend to espouse more conservative and religious positions when they are at their most fertile (Durante, Rae, & Griskevicius, 2013). Durante et al. further showed that these individual and context-depended changes in political and religious orientation lead to changes in political behavior; specifically, ovulating single women tended to vote for and donate money to the campaign of the more liberal presidential candidate (i.e., Barack Obama), while ovulating married women tended to vote for and donate to the more conservative candidate (i.e., Mitt Romney). The researchers suggested that the differential effects on political behavior lie in the different reproductive strategies employed by ovulating and non-ovulating women. Specifically, because single women who are fertile would be more likely to acquire high-status sexual partners by endorsing more liberal sexual norms, their voting behavior would tend to swing to the left. The researchers also suggested that because married women have much to lose by engaging in extra-pair copulations when they are ovulating (e.g., the loss of resources from spouses and the incurred shame from families and communities), they reinforce their own marital commitments by publicly endorsing more conservative sexual mores.

As suggested by Weeden, Cohen, and Kenrick's (2008) research, however, monogamy-prone men and women (many of whom are married) might endorse more conservative religious beliefs and practices because those beliefs and practices might help them to promote high fertility within their monogamous relationships and to deter their partners' infidelity (see also Weeden & Kurzban, 2013). If that is the case, then Weeden et al.'s findings may provide two additional explanations as to why the ovulating married women in Durante et al.'s study tended to vote for and donate to the more conservative candidate: (1) these women may have been endorsing the monogamous cultural practice of raising large families, as advocated by the religious injunction to "be fruitful and multiply," and (2) by voting for a conservative candidate who was more committed to "family values" and "pro-life" positions than his opponent, monogamous women may have sought to make promiscuous reproductive strategies—strategies that threaten their monogamous relationships with the constant threat of their partners' infidelity—more costly (Weeden & Kurzban, 2014). In other words, political and religious positions are, in part, tools by which some individuals manipulate the reproductive behaviors of other individuals in an extended phenotypic manner. That fast and slow life history strategies (i.e.,

shackelf@oakland.edu

strategies that differ on investment in mating vs. investment in parenting, respectively) may differentially affect political and social positions is further attested by findings showing that sexually promiscuous individuals tend to endorse less restrictive drug laws, whereas sexually monogamous individuals tend to endorse more restrictive drug laws (Kurzban, Dukes, & Weeden, 2010; Quintelier, Ishii, Weeden, Kurzban, & Braeckman, 2013). If drug use is associated with sexual promiscuity (see Müller & Schumann, 2011, for a discussion of the aphrodisiac role of drugs), then individuals may be using religious and political positions on drug use to differentially manipulate others' social and sexual behavior for their own reproductive benefit—that is, using other humans as extended phenotypes via the espousal of specific religious and political positions.

Taken together, these findings highlight the need for an increased acknowledgment, both inside and outside of academia, of how evolved interests (moderated by dispositional and contextual factors) influence national and international developments via the use of political and religious positions as extended phenotypic tools of social and cultural manipulation. We must also gain a fuller understanding of how our political and cultural beliefs may be used as intergenerational extended phenotypes. This concept shares similarities with the "cultural niche construction" model of Laland, Odling-Smee, and Feldman (2001) and, specifically, Lehmann's (2008) "posthumous extended phenotype" model, which states that niche-constructing extended phenotypes (such as beaver dams) are selected because of the benefits they confer to an organism's offspring in addition to the organism itself. Lehmann suggested that if organisms are benefited by their ancestors' extended phenotypic effects on their current environments, then, assuming they inherited their ancestors' genes, those niche-constructing extended phenotypes will be under positive selection—even if the immediate generation is not benefited by the extended phenotypes it produces. Lehmann also suggested that extended phenotypic effects can lead populations to extinction while benefiting the lineage that possesses them. According to Lehmann, the fact that agriculture enabled humans to build sedentary communities (which are the most optimal communities for the inheritance of extended phenotypes across generations), "…raises the intriguing question of the extent to which humans have been shaped by natural selection to behave in accordance to their impact on future generations, be it at a local or at a more global scale" (p. 560).

It is not clear what the effects of being evolutionarily aware of our political and social behaviors will be. At the least, we can raise the level of individual and societal self-awareness by shining the light of evolutionary awareness onto our religious, political, and cultural beliefs. Better still, by examining our ability to mentally time travel from an evolutionarily aware perspective, we might envision more humane futures rather than using this ability to further our own and our offspring's reproductive interests. In this way, we may be able to monitor our individual and societal outcomes and direct them to a more ethical and well-being-enhancing direction for ourselves, for other species, for our—often fragile—environment, and for the future of all three.

shackelf@oakland.edu

Toward a Consilient, Metacognitive Evolutionary Paradigm

Is it worth gambling away my long-term commitment to a loved one for the opportunity to have a good time with an attractive liaison? Am I genuinely impressed by the new job candidate's credentials, or is his attractiveness preventing me from forming a more realistic assessment? Do I choose friends based on how caring, intelligent, or loyal they are, or am I swayed into befriending only the more attractive individuals? Is my testosterone-fueled sexual aggression the result of my winning a competition and thereby rising in social status and prestige? Do I vote for political candidates because of their credentials and society-benefiting policies or because of their appearance, manipulative speeches, demagoguery, or alignment with my selfish reproductive interests (characteristics that are nowadays exaggerated by the propaganda machine of corporate-funded political campaigns)? These are ethically charged questions whose immediate urgency was not made apparent until the—relatively—recent investigations into our biological workings caused us to start questioning the Standard Social Science Model and its cultural determinism (Tooby & Cosmides, 1992). This academic soul-searching was in some ways a reflection of a more personal struggle that some academics had with the ethical implications of evolutionary science. Of course, there are still some scholars who are afraid to acknowledge the influence of Darwinian processes on human psychology and behavior—though their ranks are dwindling. Whether it is due to intellectual laziness or an unfounded association between evolutionary science and Social Darwinism, these scholars' ethically blinkered views of humanity have failed to keep pace with the enlarged scope of ethical thinking that this revolution in academia has inspired.

Though there is as yet little empirical investigation of the effects of evolutionary education on intra- and interhuman dynamics, evolutionary thinking—whether about sexual maturation, mate selection, intrasexual competition, or the use of cultural beliefs as extended phenotypes—may help to enhance the quality of human relationships and to promote individual and social well-being. We hope that this discussion will contribute to the advancement of a metacognitive approach whose aim is the theoretical and empirical analysis of how individuals are physiologically, psychologically, and behaviorally affected by various forms of scientific and cultural ideas, be they evolution by natural selection or postmodernism. Such an analysis requires consilience across disciplinary boundaries that separate the sciences from the humanities (Wilson, 1998). For this endeavor to succeed, however, we may need to achieve an individual-level consilience of heretofore irreconcilable cognitive processes (modules?). An individually enacted evolutionary awareness may be the first step toward such an enlightened synthesis of knowledge and interdisciplinary cooperation.

shackelf@oakland.edu

Unresolved Issues

Without a systematic analysis of specific beliefs and forms of knowledge (e.g., religious, political, scientific, philosophical, moral, and aesthetic) on an individual's psychological state and social behavior, it would be difficult to predict the utility of an evolution-based approach to individual and social improvement such as the one advocated here. A study by Dar-Nimrod, Heine, Cheung, and Schaller (2011), however, has made some headway in this endeavor. The researchers found that, after exposure to a social-constructionist interpretation of male–female behavioral differences (i.e., one that was based on "gender roles"), male participants imposed harsher punishments on other men accused of either rape or engaging in prostitution. Male participants who were exposed to a biological interpretation of sex differences (i.e., one that was based on differential parental investment) did not exhibit this effect. This suggests that ideological narratives within the cultural sphere might influence collective behavior in pursuit of the reproductive interests of those whose narratives predominate. For example, feminist women and their supporters might be pursuing their reproductive interests by promoting a social-constructionist belief system that is more punitive toward sexually aggressive men. Conversely, "men's rights" activists and their supporters might seek to challenge such a belief system if it is as likely to punish the falsely accused as the guilty, or to be more punishing of male sexuality, in general.

Dar-Nimrod and colleagues suggest that a social-constructivist interpretation might be associated with an increase in *perceived* individual autonomy and, hence, a greater attribution of guilt to individuals accused of sex crimes. What needs explaining is why a social-constructivist interpretation of reality would lead to a greater belief in individual autonomy and, hence, culpability for sexual offenses. Could it be that if something is seen as a product of a rapidly changing culture, it is (either rightly or wrongly) assumed to be easier to modify and, hence, one is more morally responsible for acts of culturally influenced sexual violence? Whatever the explanation, it is clear that the cultural definitions and interpretations that individuals impose on the male–female relationship can have wide-ranging social consequences. As such, it is important to understand the consequences and the potential points of opposition that an evolutionary awareness might engender.

We believe that an evolutionary perspective is a perspective without which ethical discussions are incomplete, at best, and disastrous, at worst. However, evolutionary approaches to ethics must be part of a broader, consilient framework (both inside and outside of academia) if these approaches are to assist us in bringing about greater well-being. The reason for this is that evolutionary awareness—and science, in general—is a system of informing humans about the way the world works. As Sam Harris suggested in *The Moral Landscape* (2010), science has the power not only to describe reality but also to inform us as to what is moral and what is immoral (provided that we accept certain utilitarian ethical foundations such as the promotion of happiness, flourishing, and well-being—all of which fall into Haidt's (2012) "Care/Harm" foundation of morality). Likewise, Simpson (1951) proposed that if

shackelf@oakland.edu

there is any ethical lesson that we can learn from the evolutionary process, it is that of the acquisition, dissemination, and the responsible stewardship of knowledge. Harris's and Simpson's ethical stances both rely on knowledge and the methodology of science to illuminate our perspective and guide our decisions in a manner that is ethically sound.

Scientific knowledge, however, can be parasitized by selfish human tendencies to gain power over others or to inflict pain. Harris posits that because science often shows us the very steps that are needed to bring about human well-being, it can be said to "determine" human values. The counterargument is that science can also show us how to increase pain and suffering, as is exemplified by the ingenious scientist torturers of the Middle Ages, devising the most innovative inventions for inflicting cruelty on heretics, apostates, and blasphemers (Pinker, 2011), the considerable talent that went into the design of Zyklon B, the gas that would later be used in Nazi gas chambers, and, of course, the design of the atom and hydrogen bombs. This is why the scientific mindset is not inherently moral or immoral. We agree with Harris that, as an institution, science should be allowed—and, indeed, it should be exhorted—to facilitate the flourishing of human and nonhuman sentient beings. However, at least semantically, the various steps needed to be taken toward that betterment—as discovered or discoverable by science—are not values in themselves. Instead, they are the implementations of values—in this case, the increase of human and nonhuman well-being, pleasure, happiness, longevity, and contentment.

Because scientific knowledge is necessary—but not sufficient—for the advancement of human and nonhuman well-being, academic fields on the border between science and ethics are required. We hope that the framework we have advanced here may give rise to such a discipline in the evolutionary sciences—namely, a discipline that synthesizes knowledge from evolutionary, biological, and philosophical fields. The potential danger of such investigations is that the flip side of knowing how to better our lives also means knowing how to make them worse. For this reason, ethically minded academic fields must not neglect to share knowledge and collaborate with various social-workers' groups, nursing organizations, and charitable foundations to occasionally get up from the armchair of philosophical ethics and work on "real-world" problems.

Acknowledgement This chapter is based on Gorelik and Shackelford (2014).

References

Aidoo, M., Terlouw, D. J., & Kolczak, M. S. (2002). Protective effects of the sickle cell gene against malaria morbidity and mortality. *The Lancet, 359*, 1311–1312.

Alonso-Alvarez, C., Bertrand, S., Faivre, B., Chastel, O., & Sorci, G. (2007). Testosterone and oxidative stress: The oxidation handicap hypothesis. *Proceedings of the Royal Society of London B, 274*, 819–825.

Apicella, C. L., Little, A. C., & Marlowe, F. W. (2007). Facial averageness and attractiveness in an isolated population of hunter-gatherers. *Perception, 36*, 1813–1820.

shackelf@oakland.edu

Archer, J. (2006). Testosterone and human aggression: An evaluation of the challenge hypothesis. *Neuroscience and Biobehavioral Reviews, 30*, 319–345.
Atran, S., & Axelrod, R. (2008). Reframing sacred values. *Negotiation Journal, 24*, 221–246.
Axelrod, R. (2006). *The evolution of cooperation*. New York: Basic Books.
Bailey, D. H., & Geary, D. C. (2009). Hominid brain evolution. *Human Nature, 20*, 67–79.
Barrett, H. C. (2009). Where there is an adaptation, there is a domain: The form-function fit in information processing. In S. M. Platek & T. K. Shackelford (Eds.), *Foundations in evolutionary cognitive neuroscience*. Cambridge, UK: Cambridge University Press.
Bates, T. C. (2007). Fluctuating asymmetry and intelligence. *Intelligence, 35*, 41–46.
Beran, M. J. (2004). Chimpanzees (Pan troglodytes) respond to nonvisible sets after one-by-one addition and removal of items. *Journal of Comparative Psychology, 118*, 25–36.
Bernhardt, P. C. (1997). Influences of serotonin and testosterone in aggression and dominance: Convergence with social psychology. *Current Directions in Psychological Science, 6*, 44–48.
Bernhardt, P. C., Dabbs, J. M., Fielden, J. A., & Lutter, C. D. (1998). Testosterone changes during vicarious experiences of winning and losing among fans at sporting events. *Physiology & Behavior, 65*, 59–62.
Bjorklund, D. F., and Sellers, P. D. (2013). Memory development in evolutionary perspective. In P. Bauer and R. Fivush (Eds.). The Wiley handbook on the development of children's memory, Vol. 1 (pp. 126–150). Boston: Wiley-Blackwell.
Botzung, A., Denkova, E., & Manning, L. (2008). Experiencing past and future personal events: Functional neuroimaging evidence on the neural bases of mental time travel. *Brain and Cognition, 66*, 202–212.
Boyd, R., & Richerson, P. J. (1996). Why culture is common, but cultural evolution is rare. *Proceedings of the British Academy, 88*, 77–93.
Boyd, R., & Richerson, P. J. (2008). Gene-culture coevolution and the evolution of social institutions. In C. Engel & W. Singer (Eds.), *Better than conscious? Decision making, the human mind, and implications for institutions* (pp. 305–324). Cambridge: MIT Press.
Brown, W. M., Price, M. E., Kang, J., Pound, N., Zhao, Y., & Yu, H. (2008). Fluctuating asymmetry and preferences for sex-typical bodily characteristics. *Proceedings of the National Academy of Sciences, 105*, 12938–12943.
Brüne, M., & Brüne-Cohrs, U. (2006). Theory of mind: Evolution, ontogeny, brain mechanisms and psychopathology. *Neuroscience & Biobehavioral Reviews, 30*, 437–455.
Buunk, A. P., & Massar, K. (2012). Intrasexual competition among males: Competitive towards men, prosocial towards women. *Personality and Individual Differences, 52*, 818–821.
Caldwell, C. A., & Millen, A. E. (2008). Experimental models for testing hypotheses about cumulative cultural evolution. *Evolution and Human Behavior, 29*, 165–171.
Carius, H. J., Little, T. J., & Ebert, D. (2001). Genetic variation in a host-parasite association: Potential for coevolution and frequency-dependent selection. *Evolution, 55*, 1136–1145.
Chin, H. B., Sipe, T. A., Elder, R., Mercer, S. L., Chattopadhyay, S. K., Jacob, V., et al. (2012). Comprehensive risk-reduction and abstinence education interventions to prevent or reduce the risk of adolescent pregnancy, human immunodeficiency virus, and sexually transmitted infections. *American Journal of Preventive Medicine, 42*, 272–294.
Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163–228). New York: Oxford University Press.
Daly, M. (2010). Income inequality (still) rules in explaining variations in homicide rates. In Y. Muirhead & J. P. Jarvis (Eds.), *Proceedings of the 2010 meeting of the homicide research working group* (pp. 70–77). Washington, DC: National Institute of Justice.
Daly, M., & Wilson, M. I. (1988). *Homicide*. New York: Aldine de Gruyter.
Dar-Nimrod, I., Heine, S. J., Cheung, B. Y., & Schaller, M. (2011). Do scientific theories affect men's evaluations of sex crimes? *Aggressive Behavior, 37*, 440–449.
Dawkins, R. (1982). *The extended phenotype*. Oxford, England: Oxford University Press.
de Waal, F. (1996). *Good natured: The origins of right and wrong in humans and other animals*. Cambridge, MA: Harvard University Press.

shackelf@oakland.edu

Del Giudice, M. (2009). Sex, attachment, and the development of reproductive strategies. *Behavioral and Brain Sciences, 32*, 1–67.
Del Giudice, M., Angeleri, R., & Manera, V. (2009). The juvenile transition: A developmental switch point in human life history. *Developmental Review, 29*, 1–31.
Dennett, D. (1995). *Darwins dangerous idea*. New York: Simon and Schuster.
Diamond, J. (2005). *Collapse: How societies choose to fail or succeed*. New York: Viking.
Draper, P., & Harpending, H. (1982). Father absence and reproductive strategy: An evolutionary perspective. *Journal of Anthropological Research, 38*, 255–273.
Dunn, M. J., & Searle, R. (2010). Effect of manipulated prestige-car ownership on both sex attractiveness ratings. *British Journal of Psychology, 101*, 69–80.
Durante, K. M., Rae, A., & Griskevicius, V. (2013). The fluctuating female vote: Politics, religion, and the ovulatory cycle. *Psychological Science, 24*(6), 1007–1016.
Ellis, B. J. (2004). Timing of pubertal maturation in girls: An integrated life history approach. *Psychological Bulletin, 130*, 920–958.
Ellis, B. J. (2005). Determinants of pubertal timing: An evolutionary developmental approach. In B. J. Ellis & D. F. Bjorklund (Eds.), *Origins of the social mind: Evolutionary psychology and child development* (pp. 164–188). New York: Guilford Press.
Fernald, R. D. (2007). The social control of reproduction: Physiological, cellular, and molecular consequences of social status. In S. M. Platek, J. P. Keenan, & T. K. Shackelford (Eds.), *Evolutionary cognitive neuroscience* (pp. 197–220). Cambridge, MA: MIT Press.
Fincher, C. L., Thornhill, R., Murray, D. R., & Schaller, M. (2008). Pathogen prevalence predicts human cross-cultural variability in individualism/collectivism. *Proceedings of the Royal Society of London B, 275*, 1279–1285.
Flynn, E. (2006). A microgenetic investigation of stability and continuity in theory of mind development. *British Journal of Developmental Psychology, 24*, 631–654.
Francis, R. C., Soma, K., & Fernald, R. D. (1993). Social regulation of the brain-pituitary-gonadal axis. *Proceedings of the National Academy Sciences, 90*, 7794–7798.
Frost, P. (2010). The Roman state and genetic pacification. *Evolutionary Psychology, 8*, 376–389.
Furlow, F. B., Armijo-Prewitt, T., Gangestad, S. W., & Thornhill, R. (1997). Fluctuating asymmetry and psychometric intelligence. *Proceedings of the Royal Society of London B, 264*, 823–829.
Fuster, J. M. (2003). *Cortex and mind: Unifying cognition*. New York: Oxford University Press.
Gangestad, S. W., Thornhill, R., & Garver-Apgar, C. E. (2005). Adaptations to ovulation. *Current Directions in Psychological Science, 14*, 312–316.
Gangestad, S. W., Thornhill, R., & Yeo, R. A. (1994). Facial attractiveness, developmental stability, and fluctuating asymmetry. *Ethology and Sociobiology, 15*, 73–85.
Geary, D. C. (2007). Educating the evolved mind: Conceptual foundations for an evolutionary educational psychology. In J. S. Carlson & J. R. Levin (Eds.), *Educating the evolved mind* (pp. 1–100). Greenwich, CT: Information Age.
Gonzalez-Bono, E., Salvador, A., Ricarte, J., Serrano, M. A., & Arnedo, M. (2000). Testosterone and attribution of successful competition. *Aggressive Behavior, 3*, 235–240.
Goodall, J. (1990). *Through a window: My thirty years with the chimpanzees of Gombe*. New York: Houghton Mifflin.
Gorelik, G., & Shackelford, T. K. (2014). Evolutionary awareness. *Evolutionary Psychology, 12*, 783–813.
Hackman, J., & Hruschka, D. (2013). Fast life histories, not pathogens, account for state-level variation in homicide, child maltreatment, and family ties in the U.S. *Evolution and Human Behavior, 34*, 118–124.
Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York: Pantheon Books.
Hamilton, B. E., & Ventura, S. J. (2012). Birth rates for U.S. teenagers reach historic lows for all age and ethnic groups: Hyattsville, MD: National Center for Health Statistics. *NCHS Data Brief, 89*.

shackelf@oakland.edu

Hamilton, W. D., Axelrod, R., & Tanese, R. (1990). Sexual reproduction as an adaptation to resist parasites (a review). *Proceedings of the National Academy of Sciences, 87*, 2566–3573.
Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behavior, 61*, 139–151.
Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour, 59*, 771–785.
Harris, S. (2010). *The moral landscape: How science can determine human values*. New York: Free Press.
Herrmann, E., Call, J., Hernandez-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science, 317*, 1360–1366.
Hill, K., & Kaplan, H. (1999). Life history traits in humans: Theory and empirical studies. *Annual Review of Anthropology, 28*, 397–430.
Hortola, P., & Martinez-Navarro, B. (2013). The quaternary megafaunal extinction and the fate of Neanderthals: An integrative working hypothesis. *Quaternary International, 295*, 69–72.
Hunt, M. K., Crean, C. S., Wood, R. J., & Gilburn, A. S. (1998). Fluctuating asymmetry and sexual selection in the Mediterranean fruitfly (Diptera, Tephritidae). *Biological Journal of the Linnean Society, 64*, 385–396.
Jackson, R. R., Pollard, S. D., Li, D., & Fijn, N. (2002). Interpopulation variation in the risk-related decisions of Portia labiata, an araneophagic jumping spider (Araneae, Salticidae), during predatory sequences with spitting spiders. *Animal Cognition, 5*, 215–223.
Kenrick, D. T., Li, N. P., & Butner, J. (2003). Dynamical evolutionary psychology: Individual decision rules and emergent social norms. *Psychological Review, 110*, 3–28.
Kirby, D., & Laris, B. A. (2009). Effective curriculum-based sex and STD/HIV education programs for adolescents. *Child Development Perspectives, 3*, 21–29.
Kruger, D. J. (2010). Socio-demographic factors intensifying male mating competition exacerbate male mortality rates. *Evolutionary Psychology, 8*, 194–204.
Kruger, D. J., & Nesse, R. M. (2007). Economic transition, male competition, and sex differences in mortality rates. *Evolutionary Psychology, 5*, 411–427.
Kurzban, R. (2010). *Why everyone (else) is a hypocrite*. Princeton, NJ: Princeton University Press.
Kurzban, R., Dukes, A., & Weeden, J. (2010). Sex, drugs and moral goals: Reproductive strategies and views about recreational drugs. *Proceedings of the Royal Society of London B, 277*, 3501–3508.
Laland, K. N., Odling-Smee, J., & Feldman, M. W. (2001). Cultural niche construction and human evolution. *Journal of Evolutionary Biology, 14*, 22–33.
Leavesley, A. J., & Magrath, R. D. (2005). Communicating about danger: Urgency alarm calling in a bird. *Animal Behaviour, 70*, 365–373.
Lehmann, L. (2008). The adaptive dynamics of niche constructing traits in spatially subdivided populations: Evolving posthumous extended phenotypes. *Evolution, 62*, 549–566.
Letendre, K., Fincher, C. L., & Thornhill, R. (2010). Does infectious disease cause global variation in the frequency of intrastate armed conflict and civil war? *Biological Reviews, 85*, 669–683.
Libersat, F., Delago, A., & Gal, R. (2009). Manipulation of host behavior by parasitic insects and insect parasites. *Annual Review of Entomology, 54*, 189–207.
Little, A. C., Burriss, R. P., Jones, B. C., & Roberts, C. (2007). Facial appearance affects voting decisions. *Evolution and Human Behavior, 28*, 18–27.
MacDonald, H. (2008). The campus rape myth. *City Journal, 18*(1). Retrieved from http://www.city-journal.org/2008/18_1_campus_rape.html.
Markey, P. M., & Markey, C. N. (2010). Changes in pornography-seeking behaviors following political elections: An examination of the challenge hypothesis. *Evolution and Human Behavior, 31*, 442–446.
Markey, P., & Markey, C. (2011). Pornography-seeking behaviors following midterm political elections in the United States: A replication of the challenge hypothesis. *Computers in Human Behavior, 27*, 1262–1264.

shackelf@oakland.edu

Mascaro, O., & Sperber, D. (2009). The moral, epistemic, and mindreading components of children's vigilance towards deception. *Cognition, 112*, 367–380.
Mazur, A., Booth, A., & Dabbs, J. M. (1992). Testosterone and chess competition. *Social Psychology Quarterly, 55*, 70–77.
McKibbin, W. F., Shackelford, T. K., Goetz, A. T., Bates, V. M., Starrat, V. G., & Miner, E. J. (2009). Development and initial psychometric assessment of the rape avoidance inventory. *Personality and Individual Differences, 46*, 336–340.
Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences, 34*, 57–111.
Miller, S. L., & Maner, J. K. (2011). Sick body, vigilant mind: The biological immune system activates the behavioral immune system. *Psychological Science, 22*, 1467–1471.
Miller, S. L., & Maner, J. K. (2012). Overperceiving disease cues: The basic cognition of the behavioral immune system. *Journal of Personality and Social Psychology, 102*, 1198–1213.
Moller, A. P. (1992). Female swallow preference for symmetrical male sexual ornaments. *Nature, 357*, 238–240.
Müller, C. P., & Schumann, G. (2011). Drugs as instruments: A new framework for non addictive psychoactive drug use. *Behavioral and Brain Sciences, 34*, 293–347.
Pillsworth, E. G., & Haselton, M. G. (2006). Male sexual attractiveness predicts differential ovulatory shifts in female extra-pair attraction and male mate retention. *Evolution and Human Behavior, 27*, 247–258.
Pinker, S. (2011). *The better angels of our nature: Why violence has declined*. New York: Penguin.
Prokosch, M. D., Yeo, R. A., & Miller, G. F. (2005). Intelligence tests with higher g-loadings show higher correlations with body symmetry: Evidence for a general fitness factor mediated by developmental stability. *Intelligence, 33*, 203–213.
Quintelier, K. J., Ishii, K., Weeden, J., Kurzban, R., & Braeckman, J. (2013). Individual differences in reproductive strategy are related to views about recreational drug use in Belgium, the Netherlands, and Japan. *Human Nature, 24*(2), 196–217.
Re, D. E., DeBruine, L. M., Jones, B. C., & Perrett, D. I. (2013). Facial cues to perceived height influence leadership choices in simulated war and peace contexts. *Evolutionary Psychology, 11*, 89–103.
Ridley, M. (1993). *The red queen: Sex and the evolution of human nature*. New York: Viking.
Rupp, H. A., & Wallen, K. (2007). Relationship between testosterone and interest in sexual stimuli: The effect of experience. *Hormones and Behavior, 52*, 581–589.
Saad, G., & Vongas, J. G. (2009). The effect of conspicuous consumption on men's testosterone levels. *Organizational Behavior and Human Decision Processes, 110*, 80–92.
Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: The prospective brain. *Nature Reviews: Neuroscience, 8*, 657–661.
Schel, A. M., Tranquilli, S., & Zuberbuhler, K. (2009). The alarm call system of two species of black-and-white colobus monkeys (Colobus polykomos and Colobus guereza). *Journal of Comparative Psychology, 123*, 136–150.
Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science, 210*, 801–803.
Shackelford, T. K., & Goetz, A. T. (Eds.). (2012). *The oxford handbook of sexual conflict in humans*. New York: Oxford University Press.
Shipman, P. (1994). *The evolution of racism: Human differences and the use and abuse of science*. New York: Simon and Schuster.
Simpson, G. G. (1951). *The meaning of evolution*. New York: Yale University Press.
Smith, D. L. (2007). *The most dangerous animal*. New York: St. Martin's Press.
Smith, D. L. (2011). *Less than human: Why we demean, enslave, and exterminate others*. New York: St. Martin's Press.
Smith, R. M., & LaFreniere, P. (2013). Development of tactical deception from 4 to 8 years of age. *British Journal of Developmental Psychology, 31*, 30–41.

shackelf@oakland.edu

Sobolewski, M. E., Brown, J. L., & Mitani, J. C. (2012). Female parity, male aggression, and the challenge hypothesis in wild chimpanzees. *Primates, 54*, 81–88.
Sperber, D. (2010). The guru effect. *Review of Philosophy and Psychology, 1*, 583–592.
Sperber, D., & Hirschfeld, L. (1999). Culture, cognition, and evolution. In R. Wilson & F. Keil (Eds.), *MIT encyclopedia of the cognitive sciences* (pp. 111–132). Cambridge: MIT Press.
Stanton, S. J., Beehner, J. C., Saini, E. K., Kuhn, C. M., & LaBar, K. S. (2009). Dominance, politics, and physiology: Voters' testosterone changes on the night of the 2008 United States presidential election. *PloS One, 4*, 1–6.
Stewart, A. J., & Plotkin, J. B. (2013). From extortion to generosity, evolution in the iterated prisoner's dilemma. *Proceedings of the National Academy of Sciences, 110*, 1–6.
Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel and is it unique to humans? *Behavioral and Brain Sciences, 30*, 299–351.
Talwar, V., & Lee, K. (2002). Development of lying to conceal a transgression: Children's control of expressive behaviour during verbal deception. *International Journal of Behavioral Development, 26*, 436–444.
Talwar, V., & Lee, K. (2011). A punitive environment fosters children's dishonesty: A natural experiment. *Child Development, 82*, 1751–1758.
Tamura, N. (1995). Postcopulatory mate guarding by vocalization in the Formosan Squirrel. *Behavioral Ecology and Sociobiology, 36*, 377–386.
Tigue, C. C., Borak, D. J., O'Connor, J. J., Schandl, C., & Feinberg, D. R. (2012). Voice pitch influences voting behavior. *Evolution and Human Behavior, 33*, 210–216.
Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In L. C. J. Barkow (Ed.), *The adapted mind: Evolutionary psychology and the generation of culture*. New York: Oxford University Press.
Tooley, G. A., Karakis, M., Stokes, M., & Ozanne-Smith, J. (2006). Generalising the Cinderella effect to unintentional childhood fatalities. *Evolution and Human Behavior, 27*, 224–230.
Trivers, R. L. (1972). Parental investment and sexual selection. In B. Campbell (Ed.), *Sexual selection and the descent of man 1871–1971* (pp. 136–179). Chicago: Aldine.
Trumble, B. C., Cummings, D., von Rueden, C., O'Connor, K. A., Smith, E. A., Gurven, M., et al. (2012). Physical competition increases testosterone among Amazonian forager-horticulturalists: A test of the 'challenge hypothesis'. *Proceedings of the Royal Society of London B, 279*, 2907–2912.
Vallacher, R. R., & Wegner, D. M. (1989). Levels of personal agency: Individual variation in action identification. *Personality Processes and Individual Differences, 57*, 660–671.
Weeden, J., Cohen, A. B., & Kenrick, D. T. (2008). Religious participation and reproductive strategies. *Evolution and Human Behavior, 29*, 327–334.
Weeden, J., & Kurzban, R. (2013). What predicts religiosity? A multinational analysis of reproductive and cooperative morals. *Evolution and Human Behavior, 34*, 440–445.
Weeden, J., & Kurzban, R. (2014). *The hidden agenda of the political mind: How self- interest shapes our opinions and why we won't admit it*. Princeton, NJ: Princeton University Press.
West-Eberhard, M. J. (2003). *Developmental plasticity and evolution*. New York: Oxford University Press.
Wheeler, B. C. (2009). Monkeys crying wolf? Tufted capuchin monkeys use anti-predator calls to usurp resources from conspecifics. *Proceedings of the Royal Society of London B, 276*, 3013–3018.
White, S. A., Nguyen, T., & Fernald, R. D. (2002). Social regulation of gene expression during changes in male and female social status. *Experimental Biology, 205*, 2567–2581.
Williams, M. E. (2013, March 8). Can men be taught not to rape? *Salon*. Retrieved from http://www.salon.com/2013/03/08/can_men_be_taught_not_to_rape/.
Wilson, D. S. (2002). *Darwin's cathedral: Evolution, religion, and the nature of society*. Chicago, IL: University of Chicago Press.
Wilson, E. O. (1998). *Consilience: The unity of knowledge*. New York: Knopf.

shackelf@oakland.edu

Wilson, M., & Daly, M. (1998). Lethal and nonlethal violence against wives and the evolutionary psychology of male sexual proprietariness. In *Rethinking violence against women* (pp. 199–230). Thousand Oaks, CA: Sage.

Wilson, M. L., & Wrangham, R. W. (2003). Intergroup relations in chimpanzees. *Annual Review of Anthropology, 32*, 363–392.

Wilson, D. S., Hayes, S. C., Biglan, A., & Embry, D. D. (2014). Evolving the future: Toward a science of intentional change. *Behavioral and Brain Sciences, 37*(4), 395–416.

Wright, R. (2009). *The evolution of god*. New York: Little, Brown and Company.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358*, 749–750.

Yotova, V., Lefebvre, J. F., Moreau, C., Gbeha, E., Hovhannesyan, K., Bourgeois, S., et al. (2011). An X-linked haplotype of Neandertal origin is present among all non-African populations. *Molecular Biology and Evolution, 28*, 1957–1962.

shackelf@oakland.edu

The Containment Problem and the Evolutionary Debunking of Morality

Tyler Millhouse, Lance S. Bush, and David Moss

Introduction

Recent years have seen a dramatic increase in interest in evolutionary debunking arguments (or EDAs).[1] EDAs follow a long tradition of concern about the implications of evolution for morality. Since the rise of evolutionary biology, many have worried that morality may be undermined by the discovery that our moral faculties are, in some sense, a product of evolution. Even Darwin's conviction of purpose in the universe was shaken by "the horrid doubt… whether the convictions of man's mind, which has been developed from the mind of the lower animals, are of any value or at all trustworthy" (Darwin, 1887, p. 316). Contemporary concerns typically do not center on the ultimate purpose of nature (some philosophers have even attempted to exploit biological accounts of function to support ethical theories, e.g., Foot, 2001).

[1] For example, Behrends (2013), Brosnan (2011), Carruthers and James (2008), Clarke-Doane (2012), Cline (2014), Copp (2008), De Cruz and De Smedt (2012), De Lazari-Radek and Singer (2012), Enoch (2010, 2011), FitzPatrick (2013, 2014a, 2014b), Fraser (2014), Griffiths and Wilkins (2010), Jong and Visala (2014), Joyce (2006, 2013, 2014), Kahane (2011), Mason (2010), Peters (2012), Schafer (2010), Shafer-Landau (2012), Skarsaune (2011), Street (2006, 2008), Talbott (2014), Toner (2011), Vavova (2014), Visala (2011), Wielenberg (2010), and Wilkins and Griffiths (2012).

T. Millhouse (✉)
University of Arizona, 213 Social Sciences, 1145 E. South Campus Drive,
Tucson, AZ 85721, USA
e-mail: tylermillhouse@email.arizona.edu

L.S. Bush
Tufts University, Miner Hall, Medford, MA 02155, USA
e-mail: lance.bush@tufts.edu

D. Moss
Department of Media & Communication, University of Leicester,
Bankfield House, 132 New Walk, Leicester LE1 7JA, UK
e-mail: dm301@le.ac.uk

© Springer International Publishing Switzerland 2016

T.K. Shackelford, R.D. Hansen (eds.), *The Evolution of Morality*,
Evolutionary Psychology, DOI 10.1007/978-3-319-19671-8_5

shackelf@oakland.edu

They do, however, concern the reliability of human cognition and, more specifically, normative cognition. As we outline below, EDAs can take several forms. Some argue that if our moral beliefs are satisfactorily explained as a product of evolution, then this obviates an account of moral knowledge as grounded in objective, mind-independent truths. Others simply argue that whether there are moral truths or not, if evolution has pervasively influenced our moral beliefs, then we have reason to be skeptical that the moral beliefs we do have are correct.

Machery and Mallon (2010) note that there are at least three ways one might interpret the claim that "morality evolved" which each lends varying degrees of support to EDAs. While EDAs differ in their particulars, Machery and Mallon argue that they tend to be (at least implicitly) committed to the strongest of these interpretations, namely, that "moral cognition, understood as a special sort of normative cognition," evolved (p. 4). Alternatively, it might be true that "morality evolved" only insofar as a general capacity for "*normative cognition* – that is, the capacity to grasp norms and to make normative judgments" evolved (p. 4). Finally, it may be the case that "morality evolved" in the sense that merely "some components of moral psychology evolved" (p. 4). Machery and Mallon contend that only the latter two interpretations are supported by the evidence.

If Machery and Mallon (2010) are right, then this has profound implications for research into the evolution of morality. It would require both philosophers and scientists to pay closer attention to the metaethical implications of their claims, being careful not to take our present understanding of moral discourse for granted. We do not have the time or resources to fully evaluate the empirical case Machery and Mallon advance in support of their claim. Indeed, Machery and Mallon are careful to present their case as tentative and as dependent on the present state of evidence in an ongoing area of research. As they admit, much important work remains.

Accepting their conclusion for the sake of argument, we evaluate the alleged implications of their thesis for EDAs. Machery and Mallon (2010) hold that if it is not the case that morality evolved in the strong sense that a specifically *moral* form of cognition evolved, then there is little hope for EDAs. One of the central problems for EDAs is the possibility that they prove too much. After all, there is no doubt that our capacities to reason about matters far beyond the moral domain have been shaped by natural selection. Are all these capacities thereby debunked? How could one frame an EDA that exempts these capacities? We call this challenge "the containment problem."

Machery and Mallon (2010) extend the containment problem by arguing that, deprived of the strong claim that "morality evolved," debunkers are left with the more moderate claims: that normative cognition evolved and that components of moral cognition evolved. If debunkers adopt the latter claim, it's not clear that debunkers can demonstrate a pervasive evolutionary influence on morality or that they can obviate truth-tracking accounts. If debunkers adopt the former claim, they risk undermining our broader capacity for normative cognition. This "unpalatable" consequence, Machery and Mallon argue, constitutes an informal reductio ad absurdum of evolutionary debunking.

In the light of Machery and Mallon's (2010) challenge, we propose five alternative strategies for constructing EDAs. If these alternatives are successful, it would

shackelf@oakland.edu

no longer be necessary to show that a specific capacity for moral cognition evolved. Rather, research into the details of how evolution has influenced our normative cognition (broadly understood) may suffice to support a strong challenge to many of our moral beliefs. For example, we may be able to ground EDAs in the influence of natural selection on the content of our normative judgments without reference to the moral domain or, indeed, any other normative domain.[2]

As a guide to the reader, the structure of this chapter is as follows: In section "Evolutionary Metaethics," we outline in more detail what is at stake in philosophical discussions concerning how to describe, delineate, and divide the normative domain. In section "Varieties of Debunking Arguments," we describe how existing EDAs attempt, in various ways, to debunk morality. In section "The Containment Problem: Debunking Debunked?," we outline in greater detail the challenge raised by Machery and Mallon (2010) and the containment problem. This brings us to our more detailed responses to Machery and Mallon's challenge to debunking in section "Alternative Debunking Strategies." Among the responses we propose are (1) "Revised Domains," (2) "Content Domain Debunking," (3) "Meta-Normative Debunking," (4) "Functional Debunking," and (5) "Concede and Redeem." Whether these strategies ultimately succeed is beyond the scope of this chapter. We only intend to evaluate whether they could in principle avoid (or even resolve) the containment problem. Naturally, whether these strategies *ultimately* succeed depends on further *empirical* questions about how our cognitive faculties evolved and on the resolution of other philosophical objections. Also, depending on the strength of these alternatives, the success of only one may be sufficient to seriously undermine our moral beliefs as well as Machery and Mallon's reductio. Finally, in section "Some Implications for Evolutionary Psychology," we reflect on the implications of these arguments for evolutionary psychology, suggest areas for future research, and emphasize the need for an Evolutionary Metaethics.

Evolutionary Metaethics

Meta-Normative Properties and the Moral Domain

Machery and Mallon (2010) argue that there is no evolved capacity for distinctively moral cognition (p. 20).[3] On their view, moral cognition concerns a particular domain of norms (i.e., moral norms). However, what distinguishes moral norms from other kinds of norms (e.g., social, conventional, religious, and legal norms)? Moral philosophers, psychologists, and neuroscientists have put forward various

[2] Throughout the paper, we equivocate between the influence of evolution and the influence of natural selection. Without offering a full-fledged defense of adaptationism, we assume that the influence of evolutionary mechanisms other than natural selection (e.g., genetic drift) can only enhance EDAs since none of these mechanisms seem suited to producing truth-tracking capacities.

[3] See also, Machery (2012,2013).

shackelf@oakland.edu

accounts of the distinctive features of moral norms. Hare (1981) maintains that moral norms differ from other norms in that they provide reasons for action that supersede any competing reasons for action. Bicchieri (2006) proposes that moral/personal norms are unconditional, whereas social norms are conditioned on the behavior and expectations of others. Gert (2005) claims that moral norms distinctively concern harms to others. Baumard and colleagues argue that moral cognition is fundamentally about fairness (Andre & Baumard, 2011; Baumard, André, & Sperber, 2013; Baumard, Mascaro, & Chevallier, 2012; Baumard & Sheskin, in press). Still others attempt to identify the moral domain with a distinctive pattern of neural activity (Moll, Zahn, de Oliveira-Souza, Krueger, & Grafman, 2005), or a cluster of specialized psychological mechanisms (Haidt, 2012). Finally, Nucci (2001) and Turiel (1983) maintain that humans exhibit a pancultural tendency to reliably distinguish between two distinct classes of norms, one of which corresponds to stereotypical cases of moral norms (e.g., murder, theft) and the other to stereotypical conventional norms (e.g., appropriate dress, table manners).

Which account, if any, is correct is not at issue here. The important point is that different domains of norms are distinguished by their *meta-normative properties*. These properties simply describe the various ways in which norms can differ from each other. For example, norms can differ in their content, justification, adaptive role, scope of application, authority contingency, seriousness, phenomenology, or social function (Sinnott-Armstrong & Wheatley, 2012). Now that we are clearer on what it would mean to talk about specifically moral cognition, we can ask what it would mean to say that moral cognition evolved. Simply put, moral cognition evolved just in case humans possess an evolved capacity to reason about a certain class of norms (i.e., moral norms).

While we disagree that Machery and Mallon's (2010) conclusion seriously undermines debunking arguments, we enthusiastically endorse their efforts to clarify the nature of moral cognition and to distinguish the different ways in which one might interpret the claim that morality evolved. Thus, our goal in this section is to elaborate on Machery and Mallon's argument that moral cognition, interpreted as a distinctive form of normative cognition, did not evolve. We do not intend to endorse this claim, but rather to defend its plausibility. Like Machery and Mallon, we regard as uncontroversial the claim that some of the components involved in moral cognition evolved.

Machery and Mallon's (2010) critique of the claim that moral cognition evolved rests on three key points. First, they address the claim that moral norms are universal and that their universality is evidence that they evolved (Dwyer, 2006, p. 237; Hauser, 2006, p. 53; Joyce, 2006, p. 134; Prinz, 2009). Joyce (2006), for instance, claims that the capacity to make moral judgments is present in "all human societies we have ever heard of" and cites the presence of norms in ancient works such as the Epic of Gilgamesh and the Egyptian Book of the Dead as evidence that moral norms seem to appear in every known culture (p. 134). Indeed, even critics of moral nativism such as Prinz (2009) claim that "Moral norms are found in almost every recorded human society. Like language, religion, and art, morality seems to be a human universal" (p. 167). Yet such assertions run the risk of conflating the uncontroversial

shackelf@oakland.edu

claim that every society possesses *norms* with the much more controversial claim that every society possesses a set of norms that fit the researcher's particular account of what constitutes specifically *moral* norms (Machery & Mallon, 2010, p. 30). For instance, Prinz (2009) argues that to establish a robust form of moral nativism,[4] One approach would be to (1) "identify some moral norms that can be found in all (or almost all) cultures" and then (2) "show that innate, domain specific mechanisms are the best explanation of how those norms are acquired" (p. 168). Yet Prinz moves too quickly here. The problem is that this evidence does not show that these cultures possess moral norms, in particular. In order to demonstrate this, researchers would have to show that the norms found in these cultures share the same metanormative properties, properties that would allow us to distinguish these norms from nonmoral norms, and, as Machery and Mallon note, it is unclear that anyone has conducted research that would support such claims.

Second, Machery and Mallon (2010) critique one of the most prominent and well-supported lines of evidence that purports to distinguish moral norms from nonmoral norms: Turiel and colleagues' research on the moral/conventional distinction (Huebner, Lee, & Hauser, 2010; Nisan, 1987; Nucci, 2001; Nucci & Nucci, 1982; Nucci & Turiel, 1978; Nucci, Turiel, & Encarnacion-Gawrych, 1983; Smetana & Braeges, 1990; Turiel, 1979, 1983). Naturally, Machery and Mallon (2010) could not respond to every account of the moral domain, but they choose a formidable target as an exemplar. According to proponents of the moral/conventional distinction such as Nucci (2001) and Turiel (1983), humans exhibit a pancultural tendency to reliably distinguish between moral norms and conventional norms. Machery and Mallon (2010, p. 32) note that *moral* norms are supposedly judged to be independent from authorities, to hold universally, and to be grounded in harm to others, infringement of rights, or justice, whereas *conventional* norms depend on authority, are locally applicable, and are grounded in local practice. There are two questions raised by this model: (1) is the distinction innate and (2) is the distinction reflected in people's actual judgments?

Much of the work in favor of the moral/conventional distinction is developmental. The reliable and early emergence of the distinction has prompted some to conclude that it is innate (e.g., Dwyer, 2006, 1999; Joyce, 2006; Mikhail, 2000; Wilson, 1993). These arguments, like many nativist arguments, typically appeal to a "poverty of the stimulus." Interestingly, Nucci (2001) explicitly rejects nativist accounts, arguing that children learn the moral/conventional distinction by attending to whether the consequences of actions are fixed (e.g., hitting and pain) or whether they are socially/culturally mediated (e.g., vulgar language and offense).[5] Machery and Mallon cite additional work which questions whether children do reliably draw the distinction (e.g., Carter & Patterson, 1982; Gabennesch, 1990).

[4] Prinz (2009) only considers these the steps necessary to establish a particular form of moral nativism he calls "immodest moral nativism" (p. 168). Thus, Prinz acknowledges that there are other means by which one could establish some form of moral nativism.

[5] Turiel (1983) also disputes whether the developmental evidence supports nativism.

shackelf@oakland.edu

Machery and Mallon (2010) further criticize the distinction by citing work that calls into question whether adult judgment conforms to the moral/conventional distinction. For example, Kelly et al. (2007) report that judgments about several harm norms appear to display authority dependence and cultural relativity. In other cases, individuals will judge an act to be seriously wrong and authority independent, but fail to justify their conclusion by appeal to harm, rights, or justice (Haidt et al., 1993). There are reasons to doubt these studies (e.g., Rosas, 2012), but combined they do suggest that the moral domain may be less distinct and less unified than previously supposed.

Third, Machery and Mallon (2010) criticize other developmental evidence that has been cited in support of the claim that specifically moral cognition evolved. For example, they note that it is tempting to conclude from the evidence that young children seem able to understand deontic conditionals (e.g., "It's not safe outside for the squeaky mouse, so all squeaky mice must stay in the house") earlier and more easily than indicative conditionals (e.g., "It's not safe outside for the squeaky mouse, so all squeaky mice are in the house") that *moral* cognition specifically evolved (2010, p. 52). Yet they argue that this simply suggests that a special capacity for *normative* cognition appears early and has evolved, not that a tendency for *moral* cognition specifically evolved (since the *deontological* conditionals in question are not distinctively *moral*, merely *normative*) (2010, pp. 52–53).

Similarly, they criticize (2010, p. 53) the inference made by Dwyer (2007) from the fact that infants seem to develop a tendency for empathic responses to others' suffering from an early age, to the conclusion that this is compelling evidence that *moral* capacities develop at an early age. Again, they note that empathic responses to suffering need not be taken to be distinctively *moral*, even if empathy is often thought to be morally significant (2010, p. 53). Such capacities (for empathic responses) might be evolved psychological traits and might be recruited by moral cognition, but it does not follow that moral cognition specifically evolved.

Metaethical Variability and Indeterminacy

A second line of evidence that provides indirect support for Machery and Mallon's (2010) rejection of the claim that distinctively moral cognition evolved comes from a growing body of research on the psychology of folk metaethics. In particular, experimental metaethics suggests that folk intuitions about the meta-normative properties of norms exhibit both interpersonal and intrapersonal variability. Insofar as such evidence suggests that putatively moral norms do not share a distinctive set of meta-normative properties, this evidence lends some indirect support to Machery and Mallon's claim that specifically moral cognition did not evolve. The reason for this is that if there is no distinct moral domain, then this tends to undermine the notion that there is a dedicated, innate faculty for reasoning about moral norms.

Of course, the question of the unity of the moral domain is strictly independent of the question of its distinctiveness. For example, despite variations in people's

shackelf@oakland.edu

meta-normative judgments about moral norms, there may nevertheless be relatively sharp distinctions between moral norms and other norms. For example, moral norms might all possess at least three of four meta-normative properties. This would imply that there are at most five types of moral norms (i.e., 1, 2, 3, 4; 2, 3, 4; 1, 3, 4; 1, 2, 4; 1, 2, 3). In this case, there would be no property that all moral norms share. However, we could still distinguish these norms from norms that have less than three of these properties. That said, disabusing ourselves of the notion that all moral norms share a distinctive set of properties should make us less convinced that the boundaries between moral norms and other kinds of norms are adequately sharp. If even some meta-normative properties are shared between different kinds of norms, then we must ask whether reasoning about those norms is subserved by the same cognitive mechanisms.

Gill (2009) refers to the assumption that moral discourse can be captured by a single, unified account as the *uniformity determinacy* (*UD*) *assumption*. According to Gill, the UD assumption may be false. Rather than being uniform and determinate, ordinary moral discourse could be variable and indeterminate. Gill refers to this as the *metaethical indeterminacy variability* (*IV*) *thesis*. Variation may occur within or between individuals. *Intrapersonal variability* occurs when the same individual makes moral judgments in a way that fits one account in some cases but a different account in others. For example, the same individual might presuppose one set of meta-normative properties when reasoning about some moral issues but a different set when reasoning about other moral issues (e.g., by assuming relativism when thinking about sexual ethics, but assuming absolutism when thinking about physical harm.). *Interpersonal variability* or *intergroup variability* occurs when different individuals or groups presuppose different meta-normative properties when making moral judgments. For example, one cultural group might understand morality as authority independent, whereas another might understand morality to derive from divine commands.

The IV thesis derives much of its plausibility from anecdotal observations that people report an adherence to conflicting metaethical standards or speak and act in ways that suggest that they do. However, Nichols (2004b) has also found empirical support for such observations. Across five studies, most of Nichols's subjects treated moral claims as objective, yet a third or more gave responses that suggested that they were moral relativists. Subjects were also asked to explain their answers and often gave explicitly relativist rationales. For example, one subject stated that "Morality is subjective to culture," but that whether the earth is flat is "a cold, unwavering scientific fact" (as quoted in Nichols, 2004b, p. 4). Other studies provide evidence of context variability and content variability (Beebe, 2014; Goodwin & Darley, 2008; 2012; Sarkissian, Park, Tien, Wright, & Knobe, 2011; Wright, Grandjean, & McWhite, 2013). Sarkissian et al. (2011) found that the greater the cultural difference between two hypothetical people who disagreed about a moral issue, the more subjects agreed that both of these people could be right. As Beebe (2014) observes, these findings suggest that "folk intuitions about metaethical objectivity vary as a function of cultural distance, with increased cultural distance between disagreeing parties leading to decreased attributions of metaethical objectivity" (p. 167).

shackelf@oakland.edu

In addition, researchers have found that the perceived objectivity of moral norms is in part a function of their content, with some norms such as robbing a bank exhibiting a high degree of perceived objectivity, whereas others, such as abortion and assisted suicide, are perceived as less objective (Goodwin & Darley, 2008, 2012). These findings are far from decisive evidence of the variability thesis, but they at least shift some of the burden of proof back on proponents of the UD assumption (Gill, 2009). Once again, it is possible that, however disunified, moral norms are distinct from other kinds of norms. Nevertheless, as morality is revealed to be less unified, it seems increasingly likely that morality will be shown to share important features and even cognitive mechanisms with other kinds of norms.

Varieties of Debunking Arguments

As stated before, our primary project in this chapter is not to fully evaluate Machery and Mallon's (2010) conclusion, but to evaluate what their conclusion implies for evolutionary debunking arguments (and, ultimately, for evolutionary psychology). EDAs attempt to use the evolutionary origins of moral cognition to undermine *moral realism*. Moral realism is the view that moral claims attempt to report mind-independent facts (as opposed to, say, emotions or attitudes) and that at least some of these claims are true (Sayre-McCord, 2009).[6] There are several ways in which evolutionary theory might undermine realism, and some are more controversial than others. Typically, debunkers do not attempt to demonstrate that there are no objective moral truths—demonstrating nonexistence is notoriously difficult. Rather, debunkers typically attempt to *epistemically* undermine moral realism, that is, they attempt to undermine evidence in favor of the realist position.

The most basic form of debunking involves identifying and checking the empirical claims of moral philosophers. Just which of a philosopher's claims should be understood as empirical is not always obvious, but once these claims have been identified, there is generally no question as to whether one may raise a scientific challenge. Insofar as certain empirical claims are central to an ethical theory *and concern evolution*, evolutionary science may support debunking arguments against that theory. Less directly, some varieties of moral realism (e.g., Foot, 2001) make use of notions ostensibly drawn from evolutionary biology (e.g., "function," "species," or "fitness"). Insofar as these theories depend on the currency of these notions in evolutionary science, we can assess how faithfully ethicists have interpreted the science and whether the concepts on which they rely actually enjoy the

[6] By mind-independent, we don't mean to say that moral facts cannot mention mental states. For example, causing unnecessary pain might be morally wrong, but it is not wrong *in the realist sense* if its truth depends on how people regard that prohibition. For example, whether evolution occurred does not depend on how people regard the theory. On the other hand, whether something is money depends crucially on whether people regard that thing as a medium of exchange.

shackelf@oakland.edu

support of the best scientific theories. If the conceptual foundation of an ethical theory is sufficiently undermined, that theory is debunked.[7]

By contrast, most realist theories make no essential reference to evolutionary theory, and the most important (and controversial) debunking arguments go beyond mere fact-checking. Unless otherwise noted, all references to "EDAs" in this chapter refer to these kinds of arguments. Within EDAs proper, there are roughly two approaches, arguments from influence and arguments from sufficient nonmoral explanation.

Influence

Influence-based arguments contend that evolution has pervasively influenced the content of our moral judgments[8] by processes unconcerned with moral truth (e.g., Street, 2006). This argument is meant to undermine our justification for accepting the conclusions of our moral judgments. By way of analogy, suppose we learned that a private interest in maximizing publications motivates many researchers. This should undermine our confidence in the content of academic publications, ceteris paribus. We trust these sources because there are constraints that help make epistemically virtuous research practices an *instrumental* goal of the selfish researcher (e.g., peer review and severe consequences for plagiarism and fabrication). The problem in this case is not that a desire to publish is strictly opposed to arriving at true findings. Rather, the problem is that this desire is *indifferent* to true findings, except insofar as such findings further the goal of publication. By a similar argument, debunkers hold that if our capacity for moral judgment has been pervasively shaped by natural selection, this should undermine our trust in the content of our moral beliefs, ceteris paribus. Natural selection is unconcerned with producing true beliefs except insofar as true beliefs enhanced ancestral fitness. Of course, if one could demonstrate either (1) that we have no reason to believe that evolution has significantly shaped our moral capacities or (2) that there are reasonable grounds for expecting our evolved moral capacities to track the truth instrumentally, then this brand of debunking can be defeated.

[7] An illustrative analogy is the pseudo-scientific use of scientific concepts. When new age healers talk about quantum uncertainty or the vibration of strings, the healers intend to support their theories by illustrating their coherence with or basis in established science. Closer inspection reveals (of course) that the notions employed by healers bear little resemblance to their scientific counterparts. Once one has substituted the genuine notions for the fakes, the theories of pseudo-scientists generally lose whatever apparent plausibility they once enjoyed. Of course, this is an extreme example, but similar problems may exist in philosophy. Ladyman et al. (2007) levy a similar critique against the dubious use of physical concepts by metaphysicians (pp. 25–27).

[8] In general, when we speak about moral judgments, we intend to include what philosophers call "moral intuitions." In ordinary language, "intuition" often refers to a gut decision or feeling. While moral intuitions may frequently take this form, philosophers typically see intuitions as including judgments that involve more explicit reasoning or consideration.

shackelf@oakland.edu

Sufficient Nonmoral Explanations

Sufficient nonmoral explanation arguments contend that truth-tracking theories of moral judgment are superfluous (e.g., Joyce, 2006). One influential philosophical theory (Quine, 1948) holds that we should be committed to the existence of just those kinds of things that the best scientific theories tell us exist. Regardless of how strictly one adheres to this theory of *ontological commitment*, it's hard to deny that a simpler, more explanatory theory should supplant a more complicated, less explanatory one. Importantly, we generally have no problem concluding (if tentatively and probabilistically) that belief in the entities and processes proposed by superseded theories is unwarranted.

Some varieties of moral realism hold that moral facts are true in virtue of natural ones (*ethical naturalism*). Others hold that moral facts are true in virtue of nonnatural ones (*ethical nonnaturalism*). In either case, we can imagine different kinds of explanations for our moral judgments—some that explain moral judgment as tracking (or attempting to track) a domain of moral facts and some that explain moral judgment without reference to a domain of moral facts. If the best theories turn out to be of the latter sort, there seem to be good grounds for dropping our commitment (if we had one) to a domain of moral facts, since the notion of such a domain belongs to a superseded explanation. Insofar as evolutionary theory provides a superior non-truth-tracking theory, it can offer grounds for rejecting moral realism.

There are several objections one might levy against these approaches, either in their schematic form or as applied in particular arguments. Many of these objections are beyond the scope of our present discussion. However, a new objection to debunking arguments suggests serious implications, not only for debunking arguments but for broader work on the evolution of morality.

The Containment Problem: Debunking Debunked?

The Containment Problem

A popular argumentative strategy across disciplines is reductio ad absurdum. Formally, a reductio is a proof that a proposition is false because it entails a contradiction. Informally, a reductio is an argument to the effect that a proposition is false because it entails an absurd or unlikely conclusion. For example, if an ethical principle requires one to dine nightly on human infants, one would (quite reasonably) count that as a strong reason to reject the principle in question. The primary benefit of reductios is that they are indirect. Rather than directly rebutting the reasons one's opponent has offered in favor of some proposition, p, one can cast doubt on those reasons by showing that if they support p, they also support some further, untenable conclusion. Unlike formal reductios which involve strict contradictions, informal reductios always allow one to accept the unsavory implications in the name of

shackelf@oakland.edu

preserving one's original conclusion.[9] One such argument against evolutionary debunking attempts to show that the central premise in debunking arguments has undesirable implications beyond morality. Plantinga, for rather different purposes, concisely summarizes this point:

> Evolution is directly interested (so to speak) only in adaptive behavior (in a broad sense including physical functioning), not in true belief. Natural selection doesn't care what you believe; it is interested only in how you behave. It selects for certain kinds of behavior: those that enhance fitness, which is a measure of the chances that one's genes will be widely represented in the next and subsequent generations. It doesn't select for belief, except insofar as the latter is appropriately related to behavior. But then the fact that we have evolved guarantees at most that we behave in certain ways… The objective probability that our cognitive faculties are reliable, given naturalism and given that we have been cobbled together by the processes to which contemporary theory calls our attention, is low. (2009, p. 302)

The important point here is that the evolutionary debunker must provide some grounds for thinking that the evolutionary influence on moral judgments is importantly different (in either character or strength) from the evolutionary influence on other judgments which the debunker considers reliable. The challenge of drawing these distinctions we call "the containment problem."

One traditional solution to the problem is to suggest that our capacity for moral judgment contributed to ancestral fitness in a different way than, say, our capacity for perceptual judgments. For example, we have good reasons to expect biases in perception which favor the *least costly error* (as assessed in the EEA) (Haselton & Nettle, 2006). Why aren't these biases more pervasive? What if they are and we cannot tell? We cannot (of course) give a complete response to these worries, but we can identify some relevant constraints. First, there are costs associated with responding to false alarms (e.g., in time and energy). Second, there are cognitive costs arising from biases. A bias toward detecting certain objects necessarily makes other perceptions less likely. For example, if one saw a bear in *every* cave or a tiger in *every* shadow, one couldn't identify cases where danger was absent even when doing so would have been ancestrally adaptive. While these constraints still permit substantial bias, they make it harder to imagine the evolution of a *pervasively* misleading perceptual system.[10] While plausible truth-tracking theories are easy to generate for perception, they are harder to generate for moral judgment. Whether these moves are ultimately successful for debunkers, the key point here is that they are attempting to resolve the containment problem by reference to relevant differences between moral cognition and our other capacities.

[9] Kant (in)famously argues that one should not lie to a murderer to protect the innocent. This is an unsettling consequence of his deontological ethics, but one he appears willing to accept. Others have not been as willing to follow Kant in accepting this implication, with some insisting that Kant's own principles do not require that one always tell the truth (Korsgaard, 1986) and others accepting that it does (Constant, 1776) and concluding that this is a deeply problematic implication for Kant's ethical system.

[10] Obviously more needs to be said here, but a lengthy digression into the evolution of perception seems inappropriate.

shackelf@oakland.edu

The Metaethical Containment Problem

As detailed in our section on "Evolutionary Metaethics," Machery and Mallon (2010) argue that the claim "morality evolved," when interpreted in its strongest form (as the claim that specifically moral cognition evolved), overreaches the available evidence. As a substitute for this claim, they contend that the less specific claim "normative cognition evolved" is more defensible. The idea here is that however we come to make moral judgments, our competence is not the result of an evolved system for moral cognition specifically. Rather, moral reasoning is the result of a general system for normative cognition that can be exploited to subserve reasoning about a contingent and culturally variable moral domain.

One of the philosophical implications suggested by Machery and Mallon (2010) is for EDAs (detailed above). They contend that insofar as debunking relies on the claim that "morality evolved" (i.e., moral cognition evolved), it is undermined by their conclusion. It's hard to see at first glance how this is so. After all, the claim that normative cognition evolved appears to be *broader* than the claim that moral cognition evolved, since moral cognition is a type of normative cognition. Thus, the conclusion that normative cognition in general evolved seems to license (if anything) a broader critique of normative reasoning. However, this is precisely the problem. If evolution undermines moral realism, Machery and Mallon argue, then it must do so via the evolution of this broadly normative system (since that's all that *really* evolved). This makes it hard to see how other normative domains can survive unscathed. They contend:

> We have focused on a second interpretation of the claim that morality evolved: normative cognition—the capacity to grasp and apply norms—evolved…this conclusion is cold comfort to those philosophers who want to get some philosophical mileage out of evolutionary findings. This is particularly clear when one focuses on the argument that the evolution of morality would undermine the authority of moral norms […]. Suppose that this argument from the evolution of morality is meant to hang on the reading of the claim that morality evolved considered in this section: normative cognition in general evolved…If the evolution of normative cognition really undermines the authority of moral norms, then it should also undermine the authority of any kind of norms (including epistemic norms), for there is no reason why only the authority of moral norms would be undermined by the evolution of the capacity to grasp norms *tout court.* (p. 19)

If Machery and Mallon are correct, then the strategy proposed in the previous section fails since it depends on the distinctive evolutionary history of *moral* cognition. Of course, they do assume that debunkers care about defending realism in other normative domains. In particular, Machery and Mallon cite epistemic norms as a point of special concern. To undermine those, it seems, would undermine any grounds one might have for endorsing a debunking argument in the first place. Going forward, then, we will generally assume (1) that distinctively moral cognition did not evolve and (2) that some of our normative judgments (especially epistemic norms) are worth defending.

shackelf@oakland.edu

Alternative Debunking Strategies

Put briefly, our conclusion is that Machery and Mallon's (2010) evolutionary claim is merely *cool* comfort to the evolutionary debunker. It certainly rules out naïve approaches to debunking which take for granted a determinate, invariable, and innate target for debunking (i.e., moral cognition). However, we believe that a number of powerful strategies remain open to the debunker. Our purpose here is not to pass judgment on the ultimate success of these strategies against the realist but rather to evaluate them in light of the metaethical containment problem.

(1) Revised Domains

The debunker needn't defend evolved, containment-supporting distinctions between every domain of normative cognition. Suppose, for instance, that the moral/conventional distinction is not innate/evolved but varies with culture, education, or class. This needn't trouble the debunker unless he or she is concerned with defending realism in the conventional domain. The only distinctions the debunker needs to defend are those between the targeted domains and the defended domains. For example, we might imagine a distinction between behavioral norms and norms of rationality. "Behavioral norms" might encompass what we commonly consider moral, social, and conventional norms, with "norms of rationality" encompassing epistemic norms and the norms of instrumental reasoning. If the debunkers can construct a successful debunking argument against moral realism and aren't concerned with defending realism in the social and conventional domains, they should not be troubled if their argument debunks the entire domain of behavioral norms.

This kind of distinction would hardly be unprecedented in the philosophical literature. On Hume's sentimentalist ethical theory, reason alone is not a source of motivation. Reason, for Hume, is a "slave of the passions," and while it can determine *how* to achieve what one desires, it cannot determine *what* one desires (Cohon, 2010; Hume, 2007). While both norms of rationality and behavioral norms prescribe behaviors (in a sense), it is only behavioral norms that noninstrumentally guide motivation. For example, the norm *update your beliefs according to Bayes' theorem* is neutral with respect to an agent's utility function, but the norm *do not murder* is not. It is interesting that many criteria for rationality presuppose a system of preferences or desires on the part of an agent (Russell, 1997). Much more needs to be said about this to fully flesh out the behavioral/rational distinction, but it does serve as a useful example of one possible approach to revising our normative domains. Ultimately, further empirical work is needed to determine whether Hume's conjecture about the motivational roles of sentiment and reason is psychologically plausible.

shackelf@oakland.edu

(2) Content Domain Debunking

The claim that moral reasoning per se did not evolve is consistent with the claim that the content of our normative beliefs has been influenced by evolution. For example, suppose we found that some cultural groups identified the prohibition of incest as a conventional norm and others identified it as a moral norm. In addition, suppose that having some kind of norm against incest is nearly universal. One interpretation of this finding is that while we have an evolved, innate bias toward condemning incest, we have no innate bias concerning the normative domain under which we condemn it. Indeed, we might imagine whole content domains (fairness seems a likely candidate) that exhibit pervasive evolutionary influence. These content domains might primarily encompass normative judgments we consider moral but also encompass normative judgments we regard as conventional, personal, social, or legal. Perhaps these content domains could serve as a target for debunking.

Private property intuitions provide a felicitous example. Property rights appear to straddle several normative domains (at least as we conceive them). Issues of property rights appear in moral, social, and legal norms. These norms emerge fairly early in human development (e.g., in preschool children) (Bakeman & Brownlee, 1982; Weigel, 1984). In addition, as Gintis (2007) observes, property norms are also present, mutatis mutandis, in a wide range of nonhuman species, from butterflies to apes. These norms often employ some standard of prior possession or occupation.

Gintis (2007) and many others (e.g., Dawkins, 2006; Maynard Smith & Parker, 1976) suggest that evolution exploited these salient local features (e.g., occupation/possession) to support a novel strategy for handling resource conflicts. To see why this is the case, we can begin by considering only aggressive (i.e., "hawk") and submissive (i.e., "dove") strategies in a resource conflict game. The result is a mixed strategy equilibrium. Once individuals can recognize occupation/possession, it can then be used as a signal in a correlated equilibrium (in this case a "property equilibrium," Gintis, 2007, p. 3), and deference to possessors/occupiers can invade the population. Strictly speaking, deference to invaders could also be stable, but this is relatively rare in actual populations (Dawkins, 2006).

Given these considerations, the observational and theoretical case for an evolved basis for our property intuitions appears quite strong. Of course, nothing in the account above invokes moral truth. All this plausible history does is estimate the payoffs of different strategies in terms of individual reproductive fitness and use these payoffs to explain population-level phenomena (i.e., the observed frequencies of different strategies).

Let's stipulate that this selective history of property rights judgments undermines or obviates a truth-tracking account of these intuitions (i.e., it debunks them[11]).

[11] It's important to note that this debunking account would only undermine those property-related normative judgments that depend (directly or indirectly) on the relevance of possession or occupation to ownership (e.g., approving of adverse possession as a means of acquiring property).

shackelf@oakland.edu

Is this a *moral* debunking argument? Not as such, since it doesn't target norms according to their particular normative domain (e.g., the moral domain). As we noted earlier, our intuitions about property seem to shape norms across domains. It would be better to see this as *content domain debunking*. As such, it evades the problems raised by Machery and Mallon (2010) for *normative domain debunking*.

By using our normative intuitions about property as an example, we hope to show just how ambitious content-based debunking could be. As it happens, debunking arguments which selectively target other content domains have already been offered. De Lazari-Radek and Singer (2012) argue that one can selectively debunk partialist ethical intuitions. Partialism is the view that (in principle) it is morally justified or even obligatory to prefer those one knows or with whom one has a particular social or genetic relationship. Impartialism does not deny that we may have *instrumental* reasons to act in the interest of those with whom we have greater familiarity. After all, we might often be in a better position to aid those whose needs we know best. However, impartialism denies that we have any noninstrumental (i.e., principled) reasons to prefer those with whom we happen to be acquainted. In any case, it seems that most people have strong partialist intuitions (especially in regard to parental obligations). Lazari-Radek and Singer argue that normative intuitions in favor of partialism are quite plausibly a product of natural selection (e.g., via kin selection), whereas impartialist intuitions defy evolutionary explanation. Hence, only partialist intuitions are subject to evolutionary debunking. There are reasons to believe that evolution can account for key aspects of our impartialist judgments (Kahane, 2014), but the debunking of partialism can proceed regardless.

(3) Meta-Normative Debunking

As discussed earlier, some researchers have attempted to individuate normative domains (in part) by identifying regularities in the attribution of different meta-normative properties to different subsets of norms. Suppose (following Machery & Mallon, 2010) that these projects will fail to identify a universal moral domain. Nevertheless, there may still be a universal, evolved capacity to make certain kinds of meta-normative judgments.

For example, suppose there were plausible reasons to believe we have a specially evolved capacity to reason about authority-independent norms. We might be able to investigate these reasons without invoking a normative domain within which independence judgments tend to occur. Indeed, the common factor eliciting such judgments may not be anything proper to the norm in question. For example, perhaps such judgments arise in circumstances of conflict with social superiors (regardless of the nature of the particular normative judgment under dispute). The capacity to invoke an authority-independent answer to normative questions may confer a social or rhetorical advantage. This is purely speculative, of course, but the point is that our capacities to reason about authority contingency, generalizability, etc., may have selective histories in their own right. Depending on the details, these histories might

shackelf@oakland.edu

support debunking. For example, if one could substantiate the claim that our capacity for reasoning about authority independence evolved and cannot identify any plausible truth-tracking account of authority independence judgments, then perhaps such judgments should not be trusted.

(4) Functional Debunking

There is no reason that the distinctions found in folk morality need to track the joints between underlying cognitive mechanisms. For example, if the dual process theory of moral judgment is correct, what many regarded as "moral judgment" encompasses two cognitive processes (Cushman & Greene, 2012; Cushman, Young, & Greene, 2010; Greene, 2007) which tend to produce different kinds of moral judgments. Similarly, we might discover (as seems plausible) that many of the mechanisms underlying epistemic judgment are functionally distinct (in at least some important ways) from the mechanisms underlying other forms of normative cognition. If this turns out to be the case, then regardless of whether epistemic norms are distinct from moral norms in folk metaethics, there are plausible grounds for thinking about them differently with respect to their reliability since they depend on distinct cognitive mechanisms. Also, a functional distinction between epistemic and moral judgment processes would suggest separate evolutionary histories for each, which may be debunking (or not) depending on their individual elements.

Of course, if only some of the mechanisms that contribute to a moral judgment are evolved, this may be insufficient to support debunking (as Machery & Mallon, 2010, suggest). This kind of debunking would depend on a class of normative judgments being *substantially* influenced by a particular cognitive mechanism. For example, on Greene's view, automatic/affective processing determines moral judgments that differ from moral judgments made under cool(er) controlled processing. For Greene, each of these dual processes depends to some degree on affect. Only the goals of controlled processing are set via affect (similar to the Humean picture), but in the case of automatic processing, the immediate practical content of the judgment (i.e., what you ought to do) is determined by affect. If the system that governs our affective responses to moral situations is innate and evolved, one might be able to run a debunking argument against the output of automatic processing, but not against the output of controlled processing.

(5) Concede and Redeem

The debunker might simply accept that no normative domain is spared debunking. This is less radical than it first appears. Let's suppose that the evolution of normative cognition does provide some reason to doubt the conclusions of our normative reasoning *tout court*. That's far from the only data we have. There might be other good

shackelf@oakland.edu

reasons to think that aspects of our normative reasoning are sound. After all, it is not that a process of evolution cannot produce a truth-tracking system; it is merely that it is strictly indifferent to doing so. For example, we can ask with Wigner why mathematics is so "unreasonably" effective (1960). We can ask why science has been such a remarkable success (Putnam, 1975). Our best explanation in such cases may be that these disciplines (with the aid of the normative judgments they require) are tracking something deep about the world.

Of course, in adopting this approach, the debunker must commit to redeeming whatever parts of normative reasoning he or she intends to protect from debunking. These *redemption arguments* can follow the same strategies outlined for debunking. For example, the debunker might be able to selectively redeem our reasoning in a particular content domain, in a stipulated normative domain, or in a revised (but natural) normative domain. The debunker might also selectively redeem normative reasoning that depends on reliable cognitive mechanisms. Of course, the realist might offer similar defenses of those aspects of normative reasoning targeted by the debunker. Whether any of these defenses are ultimately successful is a separate matter; we merely intend to point out another path open to debunkers.

Some Implications for Evolutionary Psychology

Talking About the Evolution of Morality

So far, we have disputed the philosophical implications of a hypothesis in evolutionary psychology. In addressing those implications, we have outlined strategies for philosophers interested in evolutionary debunking. However, Machery and Mallon's (2010) hypothesis has wide ranging implications for evolutionary psychologists working on the evolution of morality. We believe that many of the strategies outlined for the debunker can be adapted for use in evolutionary psychology. The reason for this is straightforward. Both debunkers and evolutionary psychologists are interested in characterizing the influence of evolution on our normative judgments. While we highlight the relevance of "functional debunking" and "content domain debunking," other strategies suggest similar ways of characterizing the evolutionary origins of normative cognition. For example, one might talk about the evolution of our capacity for specific meta-normative judgments (suggested by "meta-normative debunking"), or one might propose new normative domains that more plausibly share a common evolutionary origin (suggested by "revised domains").

Functional Debunking

The third strategy is perhaps the most relevant to researchers in evolutionary psychology. In fact, modular approaches to cognitive processes are especially suited to this strategy. Essentially, instead of talking about the output of innate

shackelf@oakland.edu

mechanisms as *moral* judgments, one might refer to these judgments as simply *normative*, allowing the modules to individuate their outputs. For example, a researcher working within Greene's (2007) dual process model of moral cognition might distinguish between judgments that arise from each process and talk separately about each.

This approach also invites greater specificity about the contribution of different cognitive processes to our meta-normative judgments. For example, on Nichols' (2004a) sentimental rules account, representations of rules are distinct from relevant affective responses. Certain affective responses to rule violations (e.g., anger about harm) mediate judgments about the generality of the rule (i.e., how widely it applies). For example, he hypothesizes that strong negative affect disposes individuals to more broadly generalize about their judgments (e.g., that action is always and everywhere wrong). Nichols' model explains why strong affect (e.g., disgust at the sight of a surgery) does not automatically trigger condemnation and why condemnation is not always severe or absolute. In any case, the functional division between rule representation and affective processing would allow researchers to study rule representation without the metaethical baggage of invoking specifically *moral* rules. Hence, when meta-normative judgments are at issue, (e.g., in assessing the influence of affect), researchers could more precisely indicate which meta-normative judgments are produced by which systems if they drop references to morality.

Content Domain Debunking

Nativism about moral cognition can be separated from nativism about the content of normative judgments. In other words, it is possible to argue for biases in normative judgment that concern which norms are endorsed, but not how those norm are endorsed (e.g., as authority independent, or culturally relative, etc.). There might be any number of content biases in ordinary judgment. As discussed in the context of content domain debunking, one might argue that people are innately predisposed to endorse norms that require care for family. These norms might be viewed as social norms in one society and moral norms in another. They might even be represented as a kind of norm that doesn't fit into any of the normative domains we've so far discussed. However, in nearly all cultures, there may be norms (of one kind or another) that require caring for one's family. The best explanation for this universality may be a nativist one. Such an explanation need not invoke morality in particular, only a bias in broader normative cognition.

Areas for Future Research

All of the suggestions we've made for the debunker depend crucially on the outcome of relevant empirical research. While there are some extant evolutionary accounts that support these approaches, there are some particular areas for future research

shackelf@oakland.edu

that would be relevant to debunkers (and those who wish to question their empirical presuppositions). The most obvious area for future research is on the particular questions raised and conclusions drawn by Machery and Mallon (2010). However, we also attempt to highlight a more specific area for future research.

The Evolution of Meta-Normative Reasoning

Even if Machery and Mallon (2010) are correct that there is no set of meta-normative properties that universally typifies the moral domain, there remain important questions about how and why humans evolved the capacity to attribute these properties to norms and normative domains. Consider the notion of categorical ought claims (contrasted with a hypothetical ought claims):

Hypothetical ought claim: *If you do not want to feel guilty, you ought not steal.*
Categorical ought claim: *You ought not steal.*

It might be easy to take this difference for granted and imagine the acquisition of the categorical notion by simply dropping the antecedent. However, it's not clear either that people would be inclined to do this or that the result would be comprehensible. Note, we don't consider it acceptable or meaningful to say "If you do not want to feel guilty," so why should the reverse be any different? Indeed, some philosophers suggest that it is not (Neurath, 1983, p. 54). In any case, the why and how of our facility with categorical ought claims presents interesting questions for researchers. To give a more concrete example, let's continue with the example given earlier—Nichol's suggestion that strong negative affect mediates the generalization of moral norms. The idea here is that strong negative affect might incline people to universalize their normative judgments. Perhaps there are features of circumstances that give rise to negative affect that also made a tendency to generalize ancestrally adaptive. For example, if endorsing a norm as true for everyone enhances others' tendency to comply with that norm, then endorsing and defending a general reading of certain norms may be advantageous. For example, it might not be wise to conditionalize murder norms,[12] since a higher risk of being murdered would not have been ancestrally adaptive ceteris paribus. Applied more broadly, the idea might be that the tendency to generalize tracks actions (via negative affect) that were especially serious risks in the EEA.

Again, the main idea here is that evolutionary researchers can shine light on our narrower metaethical competences. This project avoids questions about the unity and universality of the moral domain by instead investigating meta-normative properties apart from the domains they traditionally characterize. Naturally, this will be a difficult project, and conclusive evolutionary accounts are unlikely, but that doesn't preclude significant advances in our understanding.

[12] Exactly what murder means would have to be fleshed out. It would be crucial to see which forms of killing are permitted and under what circumstances.

shackelf@oakland.edu

Conclusion

In this chapter, we have presented Machery and Mallon's (2010) argument in favor of the claim that morality, as such, did not evolve. In addition, we have explored the implications of this conclusion for EDAs and offered alternative debunking strategies that do not rely on the contested claim. In so doing, we have attempted to identify new ways of talking about the evolution of morality that both avoid controversial metaethical commitments and encourage a more precise and explicit discussion of our capacity for normative reasoning. Ultimately, we hope to encourage additional work in empirical evolutionary metaethics that illuminates the origins and structure of normative cognition.

References

Andre, J. B., & Baumard, N. (2011). Social opportunities and the evolution of fairness. *Journal of Theoretical Biology, 289*, 128–135.

Bakeman, R., & Brownlee, J. R. (1982). Social rules governing object conflicts in toddlers and preschoolers. In *Peer relationships and social skills in childhood* (pp. 99–111). New York: Springer.

Baumard, N., André, J. B., & Sperber, D. (2013). A mutualistic approach to morality: The evolution of fairness by partner choice. *Behavioral and Brain Sciences, 36*(01), 59–78.

Baumard, N., Mascaro, O., & Chevallier, C. (2012). Preschoolers are able to take merit into account when distributing goods. *Developmental Psychology, 48*(2), 492.

Baumard, N., & Sheskin, M. (2015). Partner choice and the evolution of a contractualist morality. In J. Decety & T. Wheatley (Eds.), *The moral brain*. Cambridge, MA: MIT Press.

Beebe, J. (2014). How different kinds of disagreement impact folk metaethical judgments. In J. C. Wright & H. Sarkissian (Eds.), *Advances in experimental moral psychology: Affect, character, and commitment* (pp. 167–187) London: Continuum.

Behrends, J. (2013). Meta-normative realism, evolution, and our reasons to survive. *Pacific Philosophical Quarterly, 94*(4), 486–502.

Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms*. Cambridge: Cambridge University Press.

Brosnan, K. (2011). Do the evolutionary origins of our moral beliefs undermine moral knowledge? *Biology & Philosophy, 26*(1), 51–64.

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

Carruthers, P., & James, S. M. (2008). Evolution and the possibility of moral realism. *Philosophy and Phenomenological Research, 77*(1), 237–244.

Clarke-Doane, J. (2012). Morality and mathematics: The evolutionary challenge. *Ethics, 122*(2), 313–340.

Cline, B. (2014). Nativism and the evolutionary debunking of morality. *Review of Philosophy and Psychology*, 1–23.

Cohon. R. (2010). Hume's moral philosophy. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2011 ed.). Retrieved from http://plato.stanford.edu/entries/moral-realism/.

Constant, B. (1776). *Des réactions politiques*. Retrieved from http://classiques.uqac.ca/classiques/constant_benjamin/des_reactions_politiques/des_reactions.html.

Copp, D. (2008). Darwinian skepticism about moral realism. *Philosophical Issues, 18*(1), 186–206.

shackelf@oakland.edu

Cushman, F., & Greene, J. D. (2012). Finding faults: How moral dilemmas illuminate cognitive structure. *Social Neuroscience, 7*(3), 269–279.
Cushman, F., Young, L., & Greene, J. D. (2010). Our multi-system moral psychology: Towards a consensus view. *The Oxford handbook of moral psychology*, 47–71.
Darwin, C. (1887). Letter to William Graham. In F. Darwin (Ed.), *The life and letters of Charles Darwin including an autobiographical note* (Vol. 1, pp. 315–317). London: John Murray.
Dawkins, R. (2006). *The selfish gene* (3rd ed.). New York, NY: Oxford University Press.
De Cruz, H., & De Smedt, J. (2012). Evolved cognitive biases and the epistemic status of scientific beliefs. *Philosophical Studies, 157*(3), 411–429.
De Lazari-Radek, K., & Singer, P. (2012). The objectivity of ethics and the unity of practical reason. *Ethics, 123*(1), 9–31.
Dwyer, S. (1999). Moral competence. In K. Murasugi & R. Stainton (Eds.), *Philosophy and Linguistics* (pp. 169–190). Boulder, CO: Westview Press.
Dwyer, S. (2006). How good is the linguistic analogy? In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind: Culture and cognition* (pp. 237–257). Oxford: Oxford University Press.
Dwyer, S. (2007) How not to argue that morality isn't innate: Comment on Prinz. In W. Sinnott-Armstrong (Ed.), Moral psychology (Vol. 1). Cambridge, MA: MIT Press. 894–914.
Enoch, D. (2011). *Taking morality seriously: A defense of robust realism*. Oxford: Oxford University Press.
Enoch, D. (2010). The epistemological challenge to meta-normative realism: How best to understand it, and how to cope with it. *Philosophical Studies, 148*(3), 413–438.
FitzPatrick, W. J. (2013). Debunking evolutionary debunking of ethical realism. *Philosophical Studies*, 1–22.
FitzPatrick, W. J. (2014a). Evolutionary theory and morality: Why the science doesn't settle the philosophical questions. *Philosophic Exchange*, *44*(1), 2.
FitzPatrick, W. J. (2014b). Morality and evolutionary biology. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2014 ed.). Retrieved from http://plato.stanford.edu/archives/fall2014/entries/morality-biology/.
Foot, P. (2001). *Natural goodness*. Oxford: Oxford University Press.
Fraser, B. J. (2014). Evolutionary debunking arguments and the reliability of moral cognition. *Philosophical Studies, 168*(2), 457–473.
Gabennesch, H. (1990). The perception of social conventionality by children and adults. *Child Development, 61*(6), 2047–2059.
Gert, B. (2005). *Morality: Its nature and justification*. New York: Oxford University Press.
Gill, M. B. (2009). Indeterminacy and variability in meta-ethics. *Philosophical Studies, 145*(2), 215–234.
Gintis, H. (2007). The evolution of private property. *Journal of Economic Behavior & Organization, 64*(1), 1–16.
Goodwin, G. P., & Darley, J. M. (2008). The psychology of meta-ethics: Exploring objectivism. *Cognition, 106*(3), 1339–1366.
Goodwin, G. P., & Darley, J. M. (2012). Why are some moral beliefs perceived to be more objective than others? *Journal of Experimental Social Psychology, 48*(1), 250–256.
Greene, J. D. (2007). The secret joke of Kant's soul. *Moral Psychology: Historical and Contemporary Readings*, 359–372.
Griffiths, P., & Wilkins, J. (2010). When do evolutionary explanations of belief debunk belief?. In P. Sloan (ed.), *Darwin in the 21st century: Nature, humanity, and god*. Notre Dame, IN: Notre Dame University Press.
Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog?. *Journal of personality and social psychology, 65*(4), 613.
Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York: Pantheon.
Hare, R. M. (1981). *Moral thinking: Its levels, method, and point*. New York: Oxford University Press.

shackelf@oakland.edu

Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and Social Psychology Review, 10*(1), 47–66.
Hauser, M. D. (2006). *Moral minds: How nature designed our universal sense of right and wrong*. New York: Ecco.
Huebner, B., Lee, J. J., & Hauser, M. D. (2010). The moral-conventional distinction in mature moral competence. *Journal of Cognition and Culture, 10*(1), 1–26.
Hume, D. (2007). A treatise of human nature: A critical edition. In D. F. Norton & M. J. Norton (Eds.). Oxford: Clarendon Press.
Jong, J., & Visala, A. (2014). Evolutionary debunking arguments against theism, reconsidered. *International Journal for Philosophy of Religion*, 1–16.
Joyce, R. (2006). *The evolution of morality*. Cambridge, MA: MIT Press.
Joyce, R. (2013). Irrealism and the genealogy of morals. *Ratio, 26*(4), 351–372.
Joyce, R. (2014). Evolution, truth-tracking, and moral skepticism. *Problems of Goodness: New Essays on Metaethics*.
Kahane, G. (2011). Evolutionary debunking arguments. *Noûs, 45*(1), 103–125.
Kahane, G. (2014). Evolution and impartiality. *Ethics, 124*(2), 327.
Kelly, D., Stich, S., Haley, K. J., Eng, S. J., & Fessler, D. M. (2007). Harm, affect, and the moral/conventional distinction. *Mind & Language, 22*(2), 117–131.
Korsgaard, C. M. (1986). The right to lie: Kant on dealing with evil. *Philosophy & Public Affairs*, 325–349.
Ladyman, J. A. C., Presnell, S. M., Short, A. J. & Groisman, B. (2007). *Every thing must go: Metaphysics naturalized*. Oxford: Oxford University Press.
Machery, E. (2012). Delineating the moral domain. *Baltic International Yearbook of Cognition, Logic and Communication*, *7*(1), 6.
Machery, E. (2013). Morality did not evolve. *Boston Colloquium for Philosophy of Science.* Lecture conducted from Boston University, Boston, MA.
Machery, E., & Mallon, R. (2010). Evolution of morality. In J. M. Doris et al. (Eds.), *The moral psychology handbook* (pp. 3–47). New York, NY: Oxford University Press.
Mason, K. (2010). Debunking arguments and the genealogy of religion and morality. *Philosophy Compass, 5*(9), 770–778.
Maynard Smith, J., & Parker, G. A. (1976). The logic of asymmetric contests. *Animal Behaviour, 24*(1), 159–175.
Mikhail, J. (2000). Rawls' linguistic analogy: A study of the 'generative grammar' model of moral theory described by John Rawls in 'A Theory of Justice.' (Unpublished doctoral dissertation). Ithaca, NY: Cornell University.
Miller, G. F. (2007). Sexual selection for moral virtues. *The Quarterly Review of Biology, 82*(2).
Moll, J., Zahn, R., de Oliveira-Souza, R., Krueger, F., & Grafman, J. (2005). The neural basis of human moral cognition. *Nature Reviews Neuroscience, 6*(10), 799–809.
Neurath, O. (1983). In R. S. Cohen & M. Neurath (Eds.), *Philosophical papers. 1913–1946.* Dordrecht: Reidel.
Nichols, S. (2004a). *Sentimental rules: On the natural foundations of moral judgment*. Oxford: Oxford University Press.
Nichols, S. (2004b). After objectivity: An empirical study of moral judgment. *Philosophical Psychology*, *17*(1), 3–26.
Nisan, M. (1987). Moral norms and social conventions: A cross-cultural comparison. *Developmental Psychology, 23*(5), 719.
Nucci, L. P. (2001). *Education in the moral domain*. Cambridge, UK: Cambridge University Press.
Nucci, L., & Nucci, M. (1982). Children's social interactions in the context of moral and conventional transgressions. *Child Development, 53*, 403–412.
Nucci, L., & Turiel, E. (1978). Social interactions and the development of social concepts in preschool children. *Child Development, 49*, 400–407.
Nucci, L., Turiel, E., & Encarnacion-Gawrych, G. (1983). Children's social interactions and social concepts analyses of morality and convention in the Virgin Islands. *Journal of Cross-Cultural Psychology, 14*(4), 469–487.

shackelf@oakland.edu

Peters, U. (2012). Evolution, moral justification, and moral realism. *Rivista Italiana di Filosofia Analitica-Junior, 3*(1), 8–18.
Plantinga, A. (2009). The evolutionary argument against naturalism: An initial statement of the argument. *Philosophy After Darwin: Classic and Contemporary Readings*, 301–309.
Prinz, J. (2009). Against moral nativism. In D. Murphy & M. Bishop (Eds.), *Stich and his critics* (pp. 167–189). Oxford: Wiley-Blackwell.
Putnam, H. (1975). *Mathematics, matter and method*. Cambridge: Cambridge University Press.
Quine, W. V. (1948). On what there is. *The Review of Metaphysics, 2*(1), 21–38.
Rosas, A. (2012). Mistakes to avoid in attacking the moral/conventional distinction. *Baltic International Yearbook of Cognition, Logic and Communication, 7*(1), 8.
Russell, S. J. (1997). Rationality and intelligence. *Artificial Intelligence, 94*(1), 57–77.
Sarkissian, H., Park, J., Tien, D., Wright, J. C., & Knobe, J. (2011). Folk moral relativism. *Mind & Language, 26*(4), 482–505.
Sayre-McCord, G. (2009). Moral realism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2011 ed.). Retrieved from http://plato.stanford.edu/entries/moral-realism/.
Schafer, K. (2010). Evolution and normative scepticism. *Australasian Journal of Philosophy, 88*(3), 471–488.
Shafer-Landau, R. (2012). Evolutionary debunking, moral realism and moral knowledge.
Sinnott-Armstrong, W., & Wheatley, T. (2012). The disunity of morality and why it matters to philosophy. *The Monist, 95*(3), 355–377.
Skarsaune, K. O. (2011). Darwin and moral realism: Survival of the iffiest. *Philosophical Studies, 152*(2), 229–243.
Smetana, J., & Braeges, J. L. (1990). The development of toddlers' moral and conventional judgements. *Merrill-Palmer Quarterly, 36*, 329–346.
Street, S. (2006). A Darwinian dilemma for realist theories of value. *Philosophical Studies, 127*(1), 109–166.
Street, S. (2008). Reply to Copp: Naturalism, normativity, and the varieties of realism worth worrying about. *Philosophical Issues, 18*(1), 207–228.
Talbott, W. J. (2014). How could a "blind" evolutionary process have made human moral beliefs sensitive to strongly universal, objective moral standards? *Biology & Philosophy*, 1–18.
Toner, C. (2011). Evolution, naturalism, and the worthwhile: A critique of Richard Joyce's evolutionary debunking of morality. *Metaphilosophy, 42*(4), 520–546.
Turiel, E. (1979). Distinct conceptual and developmental domains: Social convention and morality. In H. Howe & C. Keasey (Eds.), *Nebraska symposium on motivation, 1977: Social cognitive development*. Lincoln: University of Nebraska Press.
Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge, UK: Cambridge University Press.
Vavova, K. (2014). Debunking evolutionary debunking. *Oxford Studies in Metaethics, 9*, 76–101.
Visala, A. (2011). *Naturalism, theism, and the cognitive study of religion: Religion explained?* Farnham, UK/Burlington, VT: Ashgate Publishing Ltd.
Weigel, R. M. (1984). The application of evolutionary models to the study of decisions made by children during object possession conflicts. *Ethology and Sociobiology, 5*(4), 229–238.
Wielenberg, E. J. (2010). On the evolutionary debunking of morality. *Ethics*, *120*(3), 441–464.
Wigner, E. P. (1960). The unreasonable effectiveness of mathematics in the natural sciences. Richard courant lecture in mathematical sciences delivered at New York University, May 11, 1959. *Communications on Pure and Applied Mathematics, 13*(1), 1–14.
Wilkins, J. S., & Griffiths, P. E. (2012). Evolutionary debunking arguments in three domains. *A New Science of Religion, 23*, 133.
Wilson, J. Q. (1993). *The moral sense*. New York: The Free Press.
Wright, J. C., Grandjean, P. T., & McWhite, C. B. (2013). The meta-ethical grounding of our moral beliefs: Evidence for meta-ethical pluralism. *Philosophical Psychology, 26*(3), 336–361.

shackelf@oakland.edu

Life Is Not Good

David Benatar

To say that some lives are better or worse than others—or that *a* life is better or worse than it might otherwise have been—is obviously to make a comparative claim. It says nothing about whether any lives are good enough to count as good lives or bad enough to count as bad ones. Most people, however, do make the further claim that there are both good and bad lives. In contrast to the widespread idea that some people have good lives and others bad, I think that while some lives are better than others, no lives are good enough to count as (non-comparatively) good.

One common and instant response to such a claim is indignation. How dare one claim that no lives are good when there are billions of people who say otherwise about their own lives? I dare to make such a claim partly because there is excellent empirical evidence for the conclusion that people's judgments cannot be trusted as a reliable indicator of how good their lives really are. For example, research psychologists have shown that people are prone to optimism and optimistic (that is, inaccurately positive) assessments of their own lives. There are many manifestations of this phenomenon. People are more prone to remember good experiences than bad ones; they have exaggerated views of how well things will go for them in the future; and most people think that the quality of their lives is above average. When it comes to assessing their own moral goodness, people also tend to be overly optimistic. Very few people think of themselves as bad. If we were to trust self-assessments, we would have to conclude that there are very few bad people and evil actions, which is patently false.

It has also been shown that people's judgments about their own lives are influenced by comparisons with the lives of others. One important effect of this is that those bad features of life that are shared by all people tend to go unnoticed in assessments of how well a person's life is going.

D. Benatar (✉)
Philosophy Department, University of Cape Town, Room 3.03,
Humanities Building, University Avenue, Rondebosch 7700, South Africa
e-mail: philosophy@uct.ac.za

© Springer International Publishing Switzerland 2016

T.K. Shackelford, R.D. Hansen (eds.), *The Evolution of Morality*,
Evolutionary Psychology, DOI 10.1007/978-3-319-19671-8_6

shackelf@oakland.edu

Given the volume of evidence for the existence of such psychological traits that affect people's judgment, it would be a kind of denialism to insist that people's self-assessments are reliable.

However, there is a difference between saying that people's self-assessments of their lives are unreliably positive and saying that people's lives are not good. After all, it is possible that although people exaggerate the quality of their lives, their lives are nonetheless good. Thus, further argument is required to support the conclusion that life is not good. Why should we think that this is the case?

The simple answer is that whatever view one might have about what makes a life good or bad, human lives fall short on the good things but abound with the bad. In support of this, both some general observations and some more specific ones can be offered.

Consider pleasures and pains. Most lives contain both (to varying degrees), but there is an unfortunate asymmetry between these that seems to apply to even the best of lives. The upshot of this is that there is much more pain than pleasure. For example, whereas the most intense pleasures (such as sexual or gustatory ones) are short-lived, the worst pains have the capacity to be much more enduring. Indeed, pleasures in general *tend* to be shorter-lived than pains. Chronic pain is common, whereas there is no such thing as chronic pleasure. Moreover, the worst pains seem to be worse than the best pleasures are good. Anybody who doubts this should consider what choice they would make if they were offered the option of securing an hour of the most sublime pleasures possible in exchange for suffering an hour of the worst pain possible. Almost everybody would put much more emphasis on the avoidance of this pain even if it entailed the forfeiture of the pleasure. (This is not to say that people are unwilling to endure some lesser pains for some greater pleasures. Instead, it shows only that the best pleasures do not offset the worst pains, at least of comparable duration.)

This asymmetry applies not only to pleasures and pains but also to goods and bads more generally. Consider how an injury can be incurred in a split second and the effects felt for life. While it is true that we can also *avoid* an injury in an instant, we do not *gain* benefits that are comparable in their magnitude and longevity in a mere moment. A lifetime of learning can be obliterated by a cerebral stroke, but there are no comparable events in which one acquires as much knowledge and understanding so speedily and easily. One can lose a limb or an eye in a few seconds, whereas gaining mobility or sight, where it is possible at all, never occurs so rapidly, effortlessly or completely. A life in which benefit came quickly and effortlessly and harm came only slowly and with effort would be a fantastically better life.

Next, consider the fulfilment of our desires or the satisfaction of our preferences. There are various reasons why there is more unfulfilment than fulfilment. First, many desires are never fulfilled. Second, even when desires are fulfilled, this usually occurs only after the exercise of effort. This means that there is a period of time in which the desire is not *yet* fulfilled. Finally, when desires are eventually fulfilled, the satisfaction is typically only transitory. Satisfied desires give way to new desires (e.g., one is hungry, eats to satiety, but then becomes hungry again). Thus, a relatively small proportion of life is spent satisfied.

shackelf@oakland.edu

On some views, the good life is constituted not only of pleasure and fulfilled desires but also of certain purportedly objective goods such as knowledge, understanding, aesthetic appreciation and virtue. It is noteworthy, however, that as advanced as some of these may be in some humans, they are only a fraction of what they could, in principle, be. Human knowledge and understanding are infinitesimal. What we do know and understand is only a tiny fraction of everything that there is to know and understand. Thus, there is a much greater difference between what we know and what there is to be known than there is between what we know and knowing nothing. In other words, on the vast spectrum from knowing nothing to knowing everything, we fall very close to the ignorance pole. Similar things might be said about aesthetic appreciation. The range of colours, sounds and smells we can perceive is limited, and thus as rich as our aesthetic appreciation may seem to us, it is grossly retarded. As for virtue, it should be clear that humans are not angels. Even the morally best humans could be so much better.

People tend to forget how much of their lives are spent tired, hungry, thirsty, in pain and being either too hot or too cold or in need of voiding their bladders and bowels. The same is true of how much time people spend bored, stressed, anxious, fearful, frustrated, irritated, sad and lonely, to name but a few examples. Also unnoticed is how bad the worst parts of a life are. They often, but not always, come later in life, but the life as a whole cannot be evaluated without considering them. Moreover, we spend a very short period of time in our prime. Most of a person's life, for those who live to old age, is spent in steady decline. Those who think that longer lives are better, all things being equal, must recognize that a lifespan of about 80 years, including periods of frailty, is terrible in comparison with a life of youthful vigour that lasts several hundred or thousand years. Our lives are much worse relative to that standard than are the lives of those who die young relative to the current standard of human longevity.

Cheery people—those who think that life is, or at least, can be good—invariably attempt to reconcile the many bad things in life with the possibility of a good life. That is to say, they offer what might be called a "secular theodicy". But, like conventional theodicies, which attempt to reconcile the vast amount of evil in the world with God's existence, the secular theodicy of optimists puts the conclusion before the evidence.

Sometimes the optimists say that the bad things in life are necessary to appreciate the good things. It is unclear whether everybody suffers from this malady. Are there not some people who would be able to appreciate the good even if there were no bad? Perhaps they are a minority. In any event, it is also not clear why those who do need to experience bad in order to appreciate the good need to experience quite so much bad. And if we were to assume that all the bad in a life is necessary in order to appreciate the good, that itself would be another very bad feature of life. It would be much better if all those bad things were not necessary.

Another optimistic response to the poor quality of human life is to argue that human life must be judged by human standards. According to this view, it is unreasonable to expect human life to be judged by unattainably higher standards. It is an implication of this view that many deficiencies and negative features of human life

shackelf@oakland.edu

that are common to all humans are excluded from consideration in determining how good a human life is. To see what an astoundingly blinkered argument this is, consider some imaginary species, which we might call *Homo infortunatus*. Members of this species have a quality of life worse than most humans. Their pain and suffering is plentiful, but life for them is not without some pleasures. In response to claims that members of this species lead poor quality lives, the optimists among them might retort that if their lives were significantly better, they simply would not be *infortunati*. That response would be unimpressive. There is a difference between (a) asking how good the lives of members of a species are and (b) asking whether a much better life is compatible with being a member of that species. Perhaps a much better life than ours would no longer be a human life. It does not follow that human life is not that much worse.

What follows from the conclusion that life is not good? It does not follow that we should all kill ourselves. There are lots of good reasons for this. For example, even if our lives are bad, they might not be bad *enough* to warrant killing ourselves. Moreover, suicide leaves bereaved people, whose lives are made worse by the death of the person who has taken his own life. Thus, in the balancing of one's own interests and those of others, one has to consider very carefully whether the quality of one's life is so bad as to warrant inflicting the trauma of one's suicide on others. This problem would be avoided if everybody took their own lives at roughly the same time, but that is not going to happen.

Nor should anybody convinced by my arguments seek to kill all people against their wishes as an act of mass (involuntary) euthanasia. There are lots of good reasons for this too, but one of them is that decisions about whether a person's life has reached an unbearably low level should, where possible, be left to the person whose life it is. A person may overestimate how good his life is. It is one thing for others to make the observation that this is the case. It is quite another to terminate that person's life.

What does follow, I think, from the conclusion that life is not good is that we should not create more of it. When we bring new people into existence, we start more lives that are not good—and we necessarily do this without the permission of those who will live those lives. We have no duty to create new people, and failing to create people can do no harm to those we fail to create. Not having children might make our own lives less good, but starting lives that are not good, merely for our own gratification, is unduly selfish.

Acknowledgment This chapter was originally published under the title "No Life Is Good" in *The Philosophers' Magazine*, Second Quarter 2011, pp. 62–66.

shackelf@oakland.edu

Antinatalism in Biological and Cultural Evolution: Fertility and Suicide

Sarah Perry

Introduction

Antinatalism is a negative view toward human fertility (and perhaps the fertility of other sentient life forms). Philanthropic antinatalism is the belief that humans should not have children for the good of the (unborn) children—that it is better, for humans, not to be born (Benatar, 2006). And if we are harmed by being born, then, as Benatar (2013) and I (Perry, 2014) have argued, the case for a moral right to die is also strengthened.

Voluntarily reducing fertility (below the number that maximizes offspring who will survive to reproduce)—especially voluntary childlessness—and voluntarily dying (suicide) are two of the most puzzling behaviors of human beings, from an evolutionary perspective. To an evolutionary psychologist, these behaviors are somewhat less puzzling; evolutionary psychologists conceive organisms (including people) not as "fitness maximizers" but as "adaptation executors" (Tooby & Cosmides, 1990). The drives to pursue pair bonds, sex, and high status are likely the most important biological adaptations that incidentally drive humans to reproduce; cultural adaptations, including fertility norms, fill in the gaps between these biological drives and the way they are instantiated in human groups. Similarly, biological adaptations such as disgust at body envelope violation, fear of heights, and fear of death protect against suicide; cultural adaptations, such as religious taboos and drug prohibition, fill in the gaps between the human organism and its social context.

One of the most surprising conclusions from examining the modern fertility transition is that people seem to be having fewer children for the good of their children—that is, they trade off the existence of more children in favor of the welfare of already born children. This has been misleadingly characterized as a

S. Perry (✉)
c/o Nine-Banded Books, P.O. Box 1862, Charleston, WV 25327, USA
e-mail: seasicksarah@gmail.com

© Springer International Publishing Switzerland 2016

T.K. Shackelford, R.D. Hansen (eds.), *The Evolution of Morality*,
Evolutionary Psychology, DOI 10.1007/978-3-319-19671-8_7

shackelf@oakland.edu

"quality–quantity" tradeoff, but it should be noted that the quality–quantity tradeoff here is not for better evolutionary quality. People do not have fewer children in hopes that those fewer children will produce more grandchildren than a larger number of children would. For example, in a large Swedish sample over many generations, children from smaller (limited) families had more education and achieved higher socioeconomic status, but had fewer children themselves (Goodman, Koupil, & Lawson, 2012). Status was passed on at the expense of evolutionary fitness. There is no mystery here. Status has been reliably associated with high fertility for so long that humans pursue it for themselves and their children for its own sake, just as they enjoy sugar for its own sake.

Parents had been asked, it seemed, for centuries to come to a miniature version of the Repugnant Conclusion (Parfit, 1986). Under the new pattern, a parent was able to form a sort of alliance with his first few children, increasing their well-being at the (unfelt) expense of the existence of their never-born siblings. When the birth of children was firmly the responsibility of God, this calculation was not even considered. But imagine an eighteenth century peasant watching his neighbor's few, well-fed children gain status through education, while his own, many children are malnourished with no prospects beyond his own. Combined with the waning cultural influence of the Catholic Church, the direct observation of Fortune's smile must have been compelling. Only those in very tightly knit religious communities, exercising a high degree of social control, and ideally rarely observing small families in practice, were immune. Only they were not compelled to compete by fertility control—at least, for a limited time and in a progressively limited manner.

There is strong evidence to think that cultural norms, rather than culture-independent biological drives, are responsible for shaping human fertility today. Testa, Cavalli, and Rosina (2014) demonstrate that in Italy, when husband and wife disagree on whether to have another child, the cultural norm prevails. DeRose, Dodoo, and Patil (2002) provide similar evidence for Ghana: couple disagreement is resolved not by husband or wife tending to "win" the dispute but by the fertility norm within the society. The partner who wishes for a number of children outside the norm tends to be overridden.

A norm may be merely a mathematical average, with elite innovators leading the way and the rest of the population gradually following suit. But norms also encompass cultural information about fertility other than the average number of children per woman. Education and wealth expectations, and the pursuit of culturally determined status in general, are some of the most salient.

The Evolution of Fertility Norms

Where cultures may only be transmitted to the genetic descendants of their people (vertical transmission, as with pathogens), fertility norms are perhaps the most important aspect of cultural evolution. Every human group ever studied by anthropologists has a population policy (Murdock, 1945); while usually not explicit,

shackelf@oakland.edu

every culture has norms and behaviors that limit population or mandate its increase. Different environments and technological packages are associated with different fertility norms. Arctic hunter populations face different pressures than medieval European farmers, and different fertility norms are successful in maintaining or even increasing a stable culture-reproducing population.

Matras (1965) identified four possible fertility strategies that a human culture might include: early marriage, uncontrolled fertility; late marriage, uncontrolled fertility; early marriage, controlled fertility; and late marriage, controlled fertility. Controlled fertility is not an artifact of modernity. Hunter-gatherers often control their fertility within marriage, whether by infanticide or by nonreproductive sexual practices or abstinence. The number of children born to the average hunter-gatherer woman ranges from 0.87 to 8.5, and in many groups is less than three (Marlowe, 2005). The higher values are consistent with uncontrolled (natural) fertility (Bongaarts, 1978), but the lower numbers indicate some form of fertility control. (Differences in nuptiality do not explain the variance.)

Hunter-gatherer populations varied substantially in population policy, depending on the demands of their environment and mode of subsistence. As some groups began to practice agriculture, however, they faced the opportunity to dramatically increase their populations and take over new lands populated by hunter-gatherers by achieving high population densities. Territory determined the success of farming cultures, and a high population was crucial to maintaining and even expanding territory. Cultures whose population policy mandated uncontrolled fertility now had a major advantage against those that did not. When an agricultural culture is expanding into a frontier populated only by hunter-gatherers, whether in the Stone Age or in the eighteenth century, only cultures establishing early, universal marriage and uncontrolled fertility as the norm are successful. Populations either increased and expanded or were kept in check by mortality.

Early marriage and uncontrolled fertility were probably the norm in all of Europe prior to the sixteenth century. This pattern exists today in sub-Saharan Africa and Afghanistan, and in isolated pockets elsewhere, though this has often been the result of a fertility transition away from an earlier, controlled norm. Improvements in agricultural productivity have mimicked the selective effects of a frontier, relaxing the need for cultural limitation of fertility, much as that experienced by the Stone Age European farmers and colonial American and Canadian farmers.

Around the sixteenth century, Europeans west of the Hajnal line began to switch to a pattern of late marriage, with a significant proportion of people never marrying, and uncontrolled fertility within the population who married. This norm was likely not adopted for the conscious purpose of limiting the population, but had the effect of keeping the population somewhat more comfortable and below the Malthusian limit (Clark, 2009). In preindustrial Japan and parts of China, however, farmers in long-settled areas kept early and universal marriage, but adopted fertility control by selective female infanticide and other means. In these populations, almost all women married and married young, but had around three children during their lives (Jones, 1990, p. 118). With industrialization and agricultural advances offering a pseudo-frontier relaxing Malthusian limits, the Japanese briefly adopted uncontrolled (or at

shackelf@oakland.edu

least much less controlled) fertility, but after World War II, they began to control their fertility once more. Most existing populations have been through multiple fertility transitions, and each transition has shaped the population.

Almost the entire world has recently undergone a single fertility transformation, one from uncontrolled fertility to controlled fertility. This transformation began in the late eighteenth century in a few small villages in France and New England and subsequently spread to every continent and almost every population in the world. Europe at this time exhibited both early and late marriage patterns, but uncontrolled fertility was the norm, a crucial part of a cultural body maintained by the Catholic Church and other institutions at the center of every community. Gradually, spurred by increasingly universal and public education, the entire world transitioned to a controlled fertility norm.

A Note on the Non-role of Child Mortality in the Fertility Transition

Since child mortality has drastically declined in the last century, many assume that child mortality had a role in people's decisions to limit fertility. This is a problematic conception: child mortality can only affect fertility when parents have a "target" parity in mind; in uncontrolled fertility regimes, no such target exists, so there is no sense in which parents might "replace" deceased children to achieve their target. The transition is from an uncontrolled fertility regime to a controlled regime; this "target parity" is the essence of the transition that must be explained.

Indeed, in sample after sample, decreasing mortality is found to have no role in decreasing fertility; it occurs at the wrong time and in the wrong place to be causal (see, e.g., Cummins, 2013; Guinane, 2010; Murphy, 2012).

Why People Used to Have Children

The nineteenth and twentieth centuries have been characterized by a massive decline in fertility, beginning in rich Western countries and spreading all over the world. It is a transformation that is still underway in a few poor countries today, especially in sub-Saharan Africa (Bongaarts, 2013).

Technological advances have, over the same period, radically decreased child mortality and increased life span. Modern parents need not have many children to ensure that one or two survive; almost all children survive to reproductive age. But Darwinian genetic interests cannot explain the modern decline in fertility (if Darwinian interests dominated, fertility should increase with increased survival, as observed in many historical elites). Rather, the fertility decline to present levels is mostly a response to the changing value of children and to the changing relationship of parents and children. The economic transformation is not spontaneous, but the product of cultural transformation through education.

shackelf@oakland.edu

The economic value of children has decreased, but this is not the most important cause of the fertility decline. The transformation of countries from predominantly agricultural to predominantly urban reduced the value of children, especially where the industrial employment of children was restricted. Each child's labor contributed positive value to a family farm or cottage industry, but in an urban setting, children began to have negative economic value. Indeed, the fertility decline correlates somewhat—though far from perfectly—with the transformation from agrarian to city life.

But the fertility decline is not merely the product of a price effect—of people having fewer children because children are more costly. Children are not normal goods (or even inferior goods, as might be surmised from low fertility among the highest income groups): they become not goods at all, but rather bundles of claims on their parents. This transformation is a culturally controlled change in direction of the flow of resources. Before the fertility decline, resources flowed from children to parents (and even up to grandparents and kin); after the transformation, resources flowed from parents to children. Caldwell (1980) argues that the vector of this cultural transformation has been mass education. He characterizes it as the replacement of "family morality," in which children are expected to "work hard, demand little, and respect the authority of the old," with "community morality," in which children are dependent on their parents to become future productive citizens (perhaps even upwardly mobile) for the good of the country.

Caldwell identifies five mechanisms by which education reduces fertility by reshaping the economic relationship of parents and children. First, education reduces the ability of a child to work inside and outside the home—not just because school and studying take up time but also because the child's student status makes others reluctant to enforce traditional duties. Second, education increases the expense of raising a child, again not just because school is expensive but because education increases a child's demands on his parents for nonschool expenses in a manner Caldwell describes as unprecedented. Third, education increases the dependency of children, reframing a formerly hardworking, productive child as primarily a future producer and citizen. Fourth, schooling speeds up cultural change and creates new cultures. Finally, in the developing world, education specifically transmits the values of the Western middle class, which is contemptuous of traditional "family morality" as described above.

In each country, before the demographic transition, children were essentially the property of their parents. Their labor could be used for the parents' good, and they were accustomed to strict and austere treatment. Parents had claims not only to their children's labor in childhood, but even to their wealth in adulthood. To put it crudely, marrying a wife meant buying a slave factory, and children were valuable slaves.

After the transition, mediated by mass education, children were transubstantiated into persons. Their individual status increased, and parents no longer had a culturally recognized claim on their labor. Children's culturally supported entitlements increased, including not only food and clothing but also study and play time. Their relationship with their parents became more egalitarian and friendly, their treatment less strict.

shackelf@oakland.edu

But children do not exactly own themselves in the present situation: the government has claims on their future earnings, through taxation and other mandatory payments (and, increasingly, education loans). In essence, mass education is a communist transformation: individually owned "goods" (children) are brought under national ownership, and returns from children flow to the country as a whole (through tax-based entitlement programs), rather than individually to their previous "owners." When farms are communally owned, production suffers and famine results; when children are communally "owned," fertility decline results. Social Security programs likely reflect this: the government provides (often poor quality) assistance to old people in place of their children, while undermining their direct claims on their children for assistance in old age.

There is another, related shift in the direction of resource flow during this time: resources (including labor) stop flowing from wives to husbands and instead flow from husbands to wives, as a result of Western-style female liberation. This trend is also a result of education and amplifies the trend toward low fertility. Since the emancipation of women frequently follows the child–parent economic transformation, it does not seem to be the first cause; Japan's fertility decline occurred in the postwar 1940s, tracking the forced implementation of Western-style mass education, but women's opportunities for education, professional employment, and political participation continue to be limited and were much more so in the 1940s, despite American-imposed female suffrage. Few would describe Japan in the 1940s as a hotbed of feminism and licentiousness, yet its fertility declined steeply and has not recovered since.

It does not seem that female emancipation was the primary cause of the fertility decline, although Caldwell details the many ways in which it amplifies the existing trend once established. It has long been noted by charitable organizations in poor countries that when resources are distributed directly to women, they are more likely to be spent on children's needs, and when distributed to men, more likely to be spent on the men's status and drug needs. Education and control of finances by women embrace and amplify the new flow of resources from parents to children, rather than children to parents. Educated children are expensive and demanding, and an educated wife makes them more so. Higher education and labor force participation by women limits the time available for child bearing and rearing, especially during the more fertile periods of women's lives.

Caldwell reports that the transition in Ghana was underway in the 1960s, and in many cases, families had both children who had been to school and children who had not. Children who had been to school were treated drastically differently from their "illiterate siblings," though they were often oblivious to this fact. That the transformation could be observed at this level—the treatment of children within the same family—suggests that changes in the status of children (expected to play, to devote time to studies, to be dependent on their parents) precede and underlie changes in gender roles.

Parental control of children's sexuality and marriage does not last long once children have been transformed into persons, and with it goes the right to collect bride price, formerly a compensation for the burden of raising a female child.

shackelf@oakland.edu

Even in dowry societies, dowry is increasingly supplemented or replaced with education, Caldwell notes, as a wealthy but uneducated woman is not seen as marriageable by Westernized elites. But this is only one aspect of the fertility transformation, rather than the prime driver; in a few countries, parental control over children's marriage survived long after the fertility decline.

Industrialization negatively affected women's productivity earlier than men's productivity, usurping traditional female work from spinning and weaving to food production. The declining economic value of both women and children necessitated that they be granted symbolic value instead. The "cult of motherhood" beginning in the 1820s in England was a response to this—granting motherhood special status as a full-time occupation and as a fulfilling life's work (Baumeister, 1991). Similarly, as the economic value of children fell, their sacred value increased. Both of these value transformations are not spontaneously occurring but culturally transmitted, and the vector for their transmission is mass Western-style education. Literature for the masses, from pamphlets of the 1820s to television, also plays a major role.

For many decades prior to the 1970s, the value of an adult (in terms of his productivity and real wages) rose; but the economic value of even an adult person has fallen in recent decades, as real wages attest. Fertility trends do not track the economic value of a human, as they might be expected to do if parents could realize a portion of the value of their offspring. Fertility trends can only respond to that share of the value of a person that a "producer of children" can recover—and the memetic transformation occasioned by mass education has essentially eliminated this share. Governments, meanwhile, claim an ever-larger share of their citizens' resources. And accessing parental money by catering to (and creating, if necessary) the "needs" of children remains a profitable business plan. The producers of children have not benefitted from their children's adult productivity in a long time, just as farm workers in China during the Great Leap Forward did not benefit from their labor.

So why did people used to have children? It's hard for us even to imagine, but children used to be valuable—they used to be much more like slaves or farm animals, which are both very valuable. They were also treated much more like slaves, with patriarchs (at least) maintaining distance from children, as Caldwell notes. Consider the history of the room known as the "study," compared to the lowly and shameful "man cave," for a sense of the old style of family relations. Making a new "person"—on which the state has claims, but you do not, and toward whom you have (class-dependent) obligations—is a much less economically attractive proposition than making a new "slave."

The somewhat uncomfortable claim here is that children used to be valuable, in the way that slaves and farm animals are valuable. A line of evidence against this is that children in some hunter-gatherer and farming societies did not, on net, contribute positive economic value to their families in terms of material production. Bergstrom (1996) relates a few examples of studies of hunter-gatherers and peasant farmers; Peruvian and Paraguayan hunter-gatherer children consumed more food than they caught up to age 18 years, and the same was true of peasant agriculturalists in contemporary India and Egypt. The rate of return on the "investment" in children, measured by their providing for parents' retirement, was only 1 %. In some societies,

shackelf@oakland.edu

measured by material consumption and production, children appear to have been a very poor investment indeed. Of course, this must be compared to other investments available in those contexts (extremely limited); and it is important to think on the margins, rather than in terms of absolutes.

However, not all studies agree that children are not economically valuable—methodology strongly influences the result, and populations vary in terms of their children's helpfulness and self-sufficiency. In many populations studied, children make significant and even net positive economic contributions, and the upward wealth flow is measurable even in merely material terms. Kramer (2005) summarizes many studies that have investigated whether children "help" and reports that Maya subsistence agriculturalist children produced more than half of their consumption by the age of 7 years for boys and 6 years for girls and produced the equivalent of their consumption at 16 years for boys and 15 years for girls. Though children are not very productive compared to adults, they are cheap—they have a very low opportunity cost for work compared to adults and are expected to work long hours even with low productivity. Kramer reports that agriculturalist children spend many more hours per day working than hunter-gatherer children and pastoralist children most of all. Hunter-gatherer children frequently become self-sufficient at a very young age. Farm children cannot become self-sufficient so early and therefore need more from their parents, but their parents demand more from them in return.

But the impact of children in terms of material production compared to consumption, and on net wealth, is not the main driver of fertility; children were valuable in other ways, and mass education interfered with all of them, not just their economic contribution. To return to the central analogy, slaves are valuable for many reasons besides their ability to produce more than they consume: they may help with childcare, provide companionship, and serve as status goods (from the point of view of peers). The type of companionship slaves provide is relevant: they are low-status beings, and with their servile behavior, they provide the owner with constant reminders that he is powerful and high status. A slave of this type's mere presence represents a type of consumption on the part of the owner, similar to the consumption of entertainment.

The practice of apprenticeship and child servitude suggests that many children even in complex societies contributed positive economic value at a young age. Much of the value that they contribute, though, is social: they make parents (or other adults) feel both needed and comparatively high in status. Submissive, servile behavior, instilled by harsh treatment and often violence, likely made them more pleasing for parents to be around. Having low-status underlings around seems to be a common human desire, expressed in a celebrity's "entourage" and, especially, in pet ownership. This human trait may even be relevant to the formation of complex hierarchies. In a sense, children used to provide a social service; education deprives them of most of their ability or willingness to engage in these behaviors.

What do children help with? They are primarily useful for the work of having a large family. Among the Maya, Kramer (2005, p. 231) reports:

> If children produced nothing, Maya parents would have to work 2.5 times as hard as they do to maintain their children's consumption between the 20th and 33rd years of the family life cycle. Were it not for the economic contributions of children, parents in their fourth and

shackelf@oakland.edu

fifth decades would have to increase their work effort by 150 %, each parent working more than 16.5 hours a day (real-time hours).

Children's work makes it possible for parents to raise large families. This may be even more true among agriculturalists and pastoralists than hunter-gatherers; among a group of Pacific Island agriculturalists, the Ifaluk, Turke (1988) found that having a daughter (or, better, two in a row) increased the completed fertility of women compared to those who had sons first; this was not the case among !Kung foragers. The nature of available work and the gender division of labor account for the difference; Kramer (2005) finds that the percentage of childcare provided to infants by their sisters ranged from 10 to 33 % in the available studies. In no case, however, did care from fathers account for a higher percentage of an infant's time than care from sisters.

But there is another way in which children used to contribute: they gave a parent his status as a free adult, and marriage and children were the only path to free adulthood. Boswell (1998) notes that in Greek, Latin, Arabic, Syriac, and many medieval languages, terms for "child," "boy," or "girl" were frequently used to mean "slave" or "servant." He reminds us that only a few hundred years ago, only a small proportion of the population married and raised children; the rest remained under someone else's control, often as servants. Similarly, in the Nakaya language, a "child" is someone who has not yet had children; one does not obtain adult status until having children of one's own. Having children was formerly the only path to achieving adult status; education changed all that, providing a new means by which to measure status and changing the status relationship between parents and children. In summary, children used to be hardworking and helpful, especially at the work of raising a large family, self-sufficient at an early age, submissive to adults, and the only path to adult status.

Education, specifically Western education promoting democratic values, interferes with children's work and their parents' expectations for their work. It makes them more dependent on their parents and makes them less likely to be servile and submissive to parents. And education itself provides an alternate means of achieving adult status other than having children. In the presence of these conditions, the demand for children is apparently low. People increasingly would rather spare their possible future children the burdens of existence, in order to maximize well-being and status for their existing children and, perhaps, for themselves.

Alone among animals, humans have brought reproduction (the most important evolutionary act) under conscious control. Conscious control of reproduction has thus become a locus for selection, with new niches arising for biological and cultural adaptations that promote reproduction against the dangerous innovation of human consciousness. Fertility is not the only evolutionarily crucial domain that has come under conscious control, however. The peculiarities of human self-aware, future-oriented consciousness also give humans conscious control over the duration of their lives, in the form of suicide. To the extent that suicide has been a selective force in human history, just as with fertility, biological and cultural innovations have had room to arise to mitigate the damage of conscious self-awareness. The next section outlines some evolutionary and cultural evolutionary perspectives on suicide: while it may be adaptive (fitness maximizing) in some circumstances, it is likely

shackelf@oakland.edu

maladaptive in others, so that even if suicide is a rational, desirable act in many circumstances, biological and cultural innovations that prevent suicide would be successful at reproducing themselves. While cultural norms that promote fertility have been largely eroded, those that prevent suicide are as strong as ever.

Is Suicide Adaptive?

Is suicide an adaptive phenomenon? Why do our brains, themselves the product of millions of years of evolution, occasionally allow us to kill ourselves and thereby destroy any hope of future reproduction? de Catanzaro has been researching the evolutionary biology of human suicide since the early 1980s (see, e.g., Maris, 1992). He outlines both adaptive and nonadaptive possibilities of the behavior of suicide. In considering the possibilities, keep in mind that suicide may not have a single explanation; some suicides (or suicide attempts) may be better explained as adaptive behavior, others as maladaptive. For instance, in terms of being the product of adaptation, the suicides of the elderly might be in an entirely different category from those of adolescents (Confer et al., 2010).

Here are the possibilities, broadly. First, suicide might not be an adaptation. It might be a logical decision made rationally by the individual and not specifically influenced by inherited traits. The positive relationship between suicide and IQ on the national level, as well as the decreased risk of suicide in mentally retarded individuals, both make this hypothesis more likely; if suicide requires a minimum IQ in order to occur, then it must be a relatively new phenomenon in the development of human beings, with not much time for adaptations to occur. While nonhuman animals sometimes engage in behavior that is lethal to themselves for kin-altruistic reasons, there is no true analogue to human suicide among other animals. If Everett (2010) is correct that suicide is completely absent among the Piraha people, those humans least engaged in abstract thought, then this is more evidence that the level of abstraction required to commit suicide has only recently been reached by humans.

However, since suicide seems to account for a significant proportion of deaths in virtually every human group ever studied, it is likely that populations have had some time to develop adaptations to this eventuality (de Catanzaro, 1991). But there is another possibility, according to de Catanzaro: the adaptations that lead to suicide in the modern world did not lead to suicide in ancestral environments. That is, there is a mismatch between human nature and the modern environments in which we find ourselves, and this mismatch is the cause of suicide. (Note that modern technology, including painless methods for suicide, are part of the "mismatch" that might be said to "cause" suicide; the absence of painless or reliable methods of suicide in past societies, in this view, would have caused people to continue living even though they did not want to.)

The most intriguing possibility is that suicide is itself an adaptive behavior, under certain circumstances. How could this be so, given that suicide ends one's survival

shackelf@oakland.edu

and destroys any future chance at reproduction, the two most crucial factors for selection? The answer lies in kin selection. Genes are not carried by the individual only, but shared with relatives; offspring are not the only chance for reproductive success. Parents, siblings, nieces, and nephews also carry one's genes. And future children are not the only path to evolutionary success; investment in previously born children (and their reproductive success) also advances genetic interests. One gets "inclusive fitness"—an increase in the chance of one's genes being passed on—from promoting the survival and reproduction of close kin as well as by promoting one's own survival and reproduction. de Catanzaro (1995) proposes a mathematical model of "adaptive suicide" in which individuals monitor their likelihood of having future surviving offspring, plus the ability to contribute to the survival and reproduction of existing relatives in proportion to their relatedness.

Under certain conditions, one's expected contribution to one's own genetic fitness (likelihood of reproduction, likelihood of the survival of one's future offspring to reproduce, effectiveness at materially supporting one's offspring and other relatives) may fall to virtually nothing. However, as long as one survives under these circumstances, an individual not only contributes nothing to his own genetic fitness but also drains the resources of his genetic relatives. His internal meter of his inclusive fitness would read a negative value, meaning that his continued survival is contrary to his genetic interests. Therefore, suicide, in this limited situation, must be said to be adaptive.

This adaptation would require that humans have a kind of "inclusive fitness monitor," noticing factors such as future fertility, ability to contribute, and burdensomeness on close kin. It would require that the brain has a mechanism for causing suicide (or mechanisms for inhibiting suicide that it could cease to engage), and this suicide mechanism would have to be triggered by a negative reading on the inclusive fitness meter. Also, for this adaptation to have come into existence, situations in which people were a significant burden on the genetic interests of their kin must have been so common in human history as to be a selective force.

Unlike fertility, there has been no modern, dramatic "suicide transition"—suicide is still relatively rare, accounting for around 1.4 % of deaths worldwide, according to the World Health Organization (2014). But unlike the case of fertility, cultural practices have accelerated to stop people from committing suicide. Public health initiatives such as drug prohibitions and emissions reduction in vehicles (so that carbon monoxide suicide is less effective) make it more difficult for people to commit suicide. Those who attempt suicide are often halted with medical intervention without their consent; billions are spent every year in emergency rooms hospitalizing those who tried to end their lives and failed.

An observation an economist might make from observing suicides is that methods of suicide are not good substitutes for each other; or, perhaps stretching the term a bit, the "elasticity" for method is low. When a popular method of suicide is made illegal or more difficult, the overall suicide rate often goes down; people do not simply substitute a different method of suicide. After Australia tightened motor vehicle exhaust restrictions, making suicide by carbon monoxide poisoning more

shackelf@oakland.edu

difficult, the regional suicide rate decreased. Suicide attempts using this method remained popular but became less lethal, resulting in fewer suicides. More impressively, a few studies of suicide barriers on bridges have found that installing a suicide barrier on a bridge does not increase suicides from nearby bridges. Gun ownership increases the risk of suicide; the availability of a method that is fairly reliably lethal confers such a reduction in the cost of suicide that merely owning a gun makes one more likely to commit suicide. One of the most universal findings about suicide is that men successfully commit suicide about four times more often than women; but where methods preferred by women are available, such as the lethal poisons that may be ingested by mouth that are available in China and India, the female suicide rate sometimes exceeds that of men. People do not seem to freely substitute one method for another.

Therefore, cultural control of suicide by controlling methods that are available has been very successful. Almost every culture's view about suicide is purely negative, whereas fertility is at least viewed ambiguously. It is important to see fertility as at least as morally important as suicide; childbearing is the act that causes a human to have a life that may be burdensome, and the burdensomeness of the life may in many cases make suicide the only desirable response.

Conclusion

The modern low-fertility pattern has allowed humans to spread the fruits of economic development among fewer people, resulting in a standard of living that pre-1800 populations never dreamed of. Starvation, disease, and genocide, the dangers of high population density, have been averted on a massive scale. It is difficult to imagine the misery that would have accompanied industrialization without controlled fertility. This transition has come with a drastic change in the relationship between parents and children, with children assuming the status of "persons" instead of slaves or property. People increasingly choose to have fewer children, for the good of those same children. Voluntary childlessness may be seen as an extreme form of the recent fertility transition: choosing to have zero children for the good of those children never born.

This fertility transition must be regarded as "maladaptive" from the gene's eye view. Similarly, suicide is usually (though likely not always) maladaptive from this perspective. While reproduction far below biological potential has become common, suicide remains relatively rare. Both the biological and cultural adaptations preventing suicide appear to be stronger and more resilient to technological change than those preventing subfertility. Entirely new cultural adaptations, such as drug prohibition and advanced emergency room treatment, have kept up with technological and social changes that might otherwise make suicide easier and more desirable. No such adaptations have arisen to reliably raise fertility, except in isolated pockets of insular religious communities.

shackelf@oakland.edu

References

Baumeister, R. (1991). *Meanings of life*. New York: Guildford Press.

Benatar, D. (2006). *Better never to have been: The harm of coming into existence*. Oxford, UK: Oxford University Press.

Benatar, D. (2013). Suicide: A qualified defence. In J. S. Taylor (Ed.), *The ethics and metaphysics of death: New essays*. New York: Oxford University Press.

Bergstrom, T. (1996). Economics in a family way. *Journal of Economic Literature, 34*, 1903–1934.

Bongaarts, J. (2013). *How exceptional is the pattern of fertility decline in sub-Saharan Africa?* (United Nations Population Division, Expert Paper 2013/4).

Bongaarts, J. (1978). A framework for analyzing the proximate determinants of fertility. *Population and Development Review, 4*(1), 105–132.

Boswell, J. (1998). *The kindness of strangers: The abandonment of children in Western Europe*. Chicago: The Chicago University Press.

Caldwell, J. C. (1980). Mass education as a determinant of the timing of fertility decline. *Population and Development Review, 6*, 225–255.

Clark, G. (2009). *A farewell to Alms*. Princeton, NJ: Princeton University Press.

Confer, J. C., Easton, J. A., Fleischman, D. S., Goetz, C. D., Lewis, D. M., Perilloux, C., et al. (2010). Evolutionary psychology: Controversies, questions, prospects, and limitations. *American Psychologist, 65*(2), 110–126.

Cummins, N. (2013). Marital fertility and wealth during the fertility transition: Rural France, 1750–1850. *The Economic History Review, 66*(2), 449–476.

de Catanzaro, D. (1991). Evolutionary limits to self-preservation. *Ethology and Sociobiology, 12*(1), 13–28.

de Catanzaro, D. (1995). Reproductive status, family interactions, and suicidal ideation: Surveys of the general public and high-risk groups. *Ethology and Sociobiology, 16*(5), 385–394.

DeRose, L. F., Dodoo, N.-A., & Patil, V. (2002). Fertility desires and perceptions of power in reproductive conflict in Ghana. *Gender & Society, 16*, 53–73.

Everett, D. (2010). The Pirahã: People who define happiness without god. *Freethought Today, 27*(3). https://ffrf.org/publications/freethought-today/item/13492-the-pirahae-people-who-define-happiness-without-god.

Goodman, A., Koupil, I., & Lawson, D. W. (2012). Low fertility increases descendant socioeconomic position but reduces long-term fitness in a modern post-industrial society. *Proceedings of Biological Sciences/The Royal Society, 279*(1746), 4342–4351.

Guinane, T. (2010, October). *The historical fertility transition: A guide for economists* (Center Discussion Paper 990). Economic Growth Center, Yale University.

Jones, H. R. (1990). *Population geography*. New York: Guilford Press.

Kramer, K. (2005). Children's help and the pace of reproduction: Cooperative breeding in humans. *Evolutionary Anthropology, 14*(6), 224–237.

Maris, R. W. (1992). *Assessment and prediction of suicide* (pp. 612 et seq.). New York, NY: The Guilford Press.

Marlowe, F. (2005). Hunter-gatherers in human evolution. *Evolutionary Anthropology, 14*, 54–67.

Matras, J. (1965). The social strategy of family formation: Some variations in time and space. *Demography, 2*, 349–362.

Milot, E., Mayer, F. M., Nussey, D. H., Boisvert, M., Pelletier, F., & Reale, D. (2011). Evidence for evolution in response to natural selection in a contemporary human population. *Proceedings of the National Academy of Sciences of the United States of America (PNAS), 108*(41), 17040–17045.

Murdock, G. (1945). The common denominator of cultures. In R. Linton (Ed.), *The science of man in the world crisis* (pp. 123–142). New York: Columbia University Press.

shackelf@oakland.edu

Murphy, T. (2012, September). *Old habits die hard (sometimes): Can département heterogeneity tell us something about the French fertility decline?* [No 364, Working Papers from IGIER (Innocenzo Gasparini Institute for Economic Research)]. Bocconi University.

Parfit, D. (1986). *Reasons and persons*. Oxford, UK: Oxford University Press.

Perry, S. (2014). *Every cradle is a grave: Rethinking the ethics of birth and suicide*. Nine Banded Books.

Testa, M. R., Cavalli, L., & Rosina, A. (2014). The effect of couple disagreement about child-timing intentions: A parity-specific approach. *Population and Development Review, 40*(1), 31–53.

Tooby, J., & Cosmides, L. (1990). The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology and Sociobiology, 11*, 375–424.

Turke, P. (1988). Helpers at the nest: Childcare networks on Ifaluk. In L. L. Betzig, M. B. Mulder, & P. Turke (Eds.), *Human reproductive behaviour: A Darwinian perspective*. Cambridge, UK: Cambridge University Press.

World Health Organization. (2014). *Preventing suicide: A global imperative*. Luxembourg: WHO Press.

shackelf@oakland.edu

Part III
Animal Behavior and Comparative Psychology

shackelf@oakland.edu

A Comparative Perspective on the Evolution of Moral Behavior

Katie Hall and Sarah F. Brosnan

> *The following proposition seems to me in a high degree probable – namely that any animal whatever, endowed with well-marked social instincts, the parental and filial affections being here included, would inevitably acquire a moral sense or conscience, as soon as its intellectual powers became as well developed, or nearly as well developed, as in man. For firstly, the social instincts lead an animal to take pleasure in the society of its fellows, to feel a certain amount of sympathy for them, and to perform various services for them.*
>
> —Darwin (1871)

Introduction

One of the characteristics often argued to set humans apart from the rest of the animal kingdom is our advanced and complex system of morals; upon closer inspection, however, precursors for moral behavior exist in many species. Oddly, despite being less popular, this latter view of continuity across species was the dominant one until

K. Hall (✉)
Keeling Center for Comparative Medicine and Research, UT MD Anderson Cancer Center, 650 Cool Water Drive, Bastrop, TX 78602, USA
e-mail: katiehallphd@gmail.com

S.F. Brosnan
Keeling Center for Comparative Medicine and Research, UT MD Anderson Cancer Center, 650 Cool Water Drive, Bastrop, TX 78602, USA

Departments of Psychology and Philosophy, Neuroscience Institute, Georgia State University, Atlanta, GA, USA

© Springer International Publishing Switzerland 2016

T.K. Shackelford, R.D. Hansen (eds.), *The Evolution of Morality*, Evolutionary Psychology, DOI 10.1007/978-3-319-19671-8_8

shackelf@oakland.edu

recently. There is a long-standing precedent for, and much evidence to support, the concepts of biological, emotional, and cognitive continuity across nonhuman animals and humans. As the opening quote indicates, Charles Darwin saw no reason to consider that moral behavior is any different and joined earlier thinkers such as Adam Smith in the view that the roots of human morality are shared across species (Smith, 1817). Therefore, it is curious that more modern thinkers have seen fit to separate out humans (Bonnie & de Waal, 2004; Brosnan, 2014b; Flack & de Waal, 2000; Haidt, 2003).

One likely reason for this discrepancy is different understandings in different disciplines of what moral behavior is, how it functions, and where it came from. For instance, anyone who denies a biological foundation for human behavior will dismiss this proposition out of hand, a view that is unfortunately common—although rapidly diminishing—in many humanities and social science departments. More commonly, however, is the implicit (or explicit) assumption that morality is something special that is too complex to be explained biologically or too complex to exist in the absence of the human intellect (e.g., that it necessitates language). However, morality need not be a higher-order phenomenon that requires language. The function of moral behavior is to reduce tension and uphold social regularity in groups of organisms (Flack & de Waal, 2000). As group living is widespread throughout the animal kingdom, we would anticipate that some aspects of what we call moral behavior are present in other species.

Morality and moral behavior are more than just a mechanism for allowing species to successfully live in social groups. In humans, morality has extended to individuals outside of the social group and, perhaps more importantly, has gone beyond mechanisms for smooth group functioning and has developed into a set of prescriptions for behavior. These aspects of morality and, in particular, the prescriptive component are more difficult to study in other species and, given the lack of language in other species, will be much less advanced, if they exist at all. However, this lack of "advanced" morality in species other than humans does not mean that the roots of moral behavior are not in these species, and we can learn much about the evolution of our own behavior by studying these roots.

Here we focus on four principal components of morality that have most frequently been explored in nonhuman species. These include conflict resolution, reciprocity, reactions to inequity, and empathy (Brosnan, 2006, in press; de Waal, 2006a; Flack & de Waal, 2000). Specifically, we argue that the precursors to moral behavior exist in rudimentary form in other species (here we focus on nonhuman primates, our closest living phylogenetic relatives) and that some of these building blocks have evolved through natural selection into the more complex behavior we see in humans. By studying these behaviors in other species, we might gain insight into the selective forces that produced our own moral system in order to better understand why we make the decisions that we do, the function of moral behavior, and how the decisions we make are constrained by evolutionary history.

shackelf@oakland.edu

Why Study Other Species?

There are a number of benefits to studying behavior in species other than humans. A primary practical advantage is that we have much greater knowledge of and control over subjects' histories (Brosnan, 2011a). In captive studies, we usually know the subjects' lifelong social history, sometimes in remarkable detail, allowing for deeper examination into the effects of social relationships on moral behavior (Brosnan, Newton-Fisher, & Van Vugt, 2009; Brosnan, Schiff, & de Waal, 2005). This is much more difficult in humans, where studies often take place either anonymously or with unfamiliar people in a laboratory setting rather than over an extended period with known individuals (Fessler, 2009; Trivers, 2006). Additionally, even when studies are not anonymous, humans may be influenced by the social or environmental context, such as one's reputation or group identity, in more substantial ways than we see in other species (Hagan & Hammerstein, 2006). Furthermore, human moral behavior can vary culturally (Eisenberg, 1992; Eisenberg & Mussen, 1989; Tomasello & Vaish, 2013; Trommsdorff, Friedlmeier, & Mayer, 2007), making it difficult to tease apart the influences and interactions of the social context on moral behavior. Again, although nonhuman primates also show cultural variation in behavior (Fragaszy & Perry, 2008; Laland & Galef, 2009; Whiten et al., 1999; Wrangham, McGrew, de Waal, & Heltne, 1994), the differences tend not to be as great as in humans, allowing experimenters to explore the biological foundations of behavior without the confound of human culture.

The second reason for studying other species is that using a comparative approach can give us insight into the biological foundations of human behavior. It can reveal how certain selective pressures may have shaped a behavior, from its initial building blocks to its current expression in extant species. Natural selection slowly shapes small changes in existing structures over many generations; therefore, such precursor behaviors can typically be found, even if something manifests differently or is used for a different purpose (Brosnan, 2014b).

This evolutionary approach to behavior was best described by Tinbergen, who famously outlined the four key questions that are essential for studying any behavior: what is the evolutionary history of a behavior, how is behavior functionally adaptive, what are the underlying mechanisms motivating the behavior, and how do these behaviors develop in the individual (Tinbergen, 1963)? These questions highlight perhaps the key distinction that needs to be made, between the ultimate function of a trait and the proximate mechanism(s) that underlies it. The ultimate function of a trait, what the trait does for those who display it, relates to how the trait evolved and under what circumstances. The proximate mechanism of a trait is how it develops and is displayed within an individual, for example, due to ontogenetic development, genetics, or hormone pathways within the organism. This distinction is important for many reasons, but perhaps the most important is that confusing ultimate and proximate levels of analysis leads to problems in interpretation.

The first problem is that while an agent's behavior may function to produce a certain result, such as altruism or deception, the agent need not recognize or

shackelf@oakland.edu

understand that his or her action was altruistic or deceptive (e.g., the mechanism need not be altruistic). That is, a subject can *act* altruistically or deceptively without an altruistic or deceptive *motivation*. A good example of this distinction between the function and motivation of the behavior is that some deceptive alarm calls (function) may be due to increased stress hormones produced by fear or stress (mechanism) rather than motivated by the goal of deception (Bercovitch, Hauser, & Jones, 1995; Blumstein, Patton, & Saltzman, 2006; Boinski, Gross, & Davis, 1999). In terms of evaluating behaviors as moral, we may observe an outcome that appears to be "moral," despite the fact that the organism did not intend to benefit another individual. If an outcome functions to benefit others (though the actor need not realize it as such), the behaviors that contribute to that outcome can be acted upon by natural selection over generations to shape behaviors that help others. It is possible that, over time, mechanisms that become other-regarding may evolve, but this need not be the case.

A second and related issue is that similar outcomes may result from different mechanisms, and the same underlying mechanism may lead to different outcomes. A failure to recognize this may lead to, on the one hand, an assumption of similarity that is not warranted and, on the other, an assumption of differences where none exist. Behaviors that evolve the same outcome with different mechanisms highlight that the outcome was so important that natural selection found the outcome several times. When similarities in mechanism lead to differences in outcome, however, we can better understand how local selective pressures may have influenced behaviors to evolve differently.

This leads us to a third important distinction, in that similarities between species can result through one of the two different processes, homology and convergence (Brosnan, 2014a, 2014b; Brosnan, Newton-Fisher & Van Vugt, 2009). When a trait is similar between closely related species, it is likely that the similarity is due to shared ancestry (homology). For example, homology can explain why most bird species can fly; they all descended from a common ancestor with that trait. Another possibility is that a similarity between species is due to similar environmental pressures faced by two species that do not share a common ancestor with the trait in question (convergence or homoplasy). For example, the fact that birds, bats, and insects can fly is due to convergent evolution; these distantly related lineages faced similar environmental pressures and opportunities to fill a certain ecological niche and all did so by evolving flight. One important caveat is that closely related species may evolve similar traits due to convergence, and distantly related species may share traits due to homology, so although it is reasonable to hypothesize that closely related species, such as those in the same taxon, share traits through homology, this is a hypothesis that must be tested (Brosnan, 2014a; Talbot, Freeman, Williams, & Brosnan, 2011). Finally, as mentioned above, two closely related species may evolve in different directions due to different ecological pressures, a process known as divergence; to continue the above example, penguins use their wings to "fly" through the water, not the air. This process may result in similar mechanisms that lead to different outcomes.

One final caveat is in order; evolution is not biological or genetic determinism (e.g., how selection or our genes preordain that we behave). Instead, the influence

shackelf@oakland.edu

of evolutionary history on the expression of a trait may best be described by the metaphor of the tangled wing (Konner, 1982); evolution constrains the direction that a behavior can take much in the way a tangled wing may have limited options for flight, but does not determine exactly which direction it must take. Therefore, by arguing that a trait evolved, we are not saying that it is invariant or that it is only expressed in one way. Similarly, the discussion of an evolutionary/comparative approach to the study of morality, i.e., a description of what exists in terms of biological precursors and underlying mechanisms, should not be confused with a normative discussion of what should be, i.e., the moral value of a behavior or the ethicist's goal of improving a behavior (Brosnan, 2014b). We discuss what animals (including humans) do, not what they should do.

Conflict Resolution

Many species in addition to humans live in social groups, but although there are benefits to social living, such as defense against predators or other groups of conspecifics, group living also has costs, such as competition for resources. Therefore, for social groups to function successfully, there must be behaviors that serve to reduce tension and avoid conflict to maintain peaceful interactions. Perhaps the simplest means of maintaining relationships and social expectations is through a dominance hierarchy (Bernstein, 1981; de Waal, 1996; Flack & de Waal, 2000; Mendoza & Barchas, 1983; Preuschoft & van Schaik, 2000). However, dominance hierarchies do not solve all problems, and fights still occur. When fights occur, there are several behaviors that serve to minimize the disruption to group harmony.

First, other individuals may intervene in fights. In some cases, this is self-serving, such as when individuals intervene on behalf of the individual already winning the fight (de Waal, 1984; Machida, 2006; Watanabe, 2001). In other cases, however, it is less clear that the behavior is self-serving. For instance, individuals may intervene impartially, without taking sides (de Waal, 1982; de Waal & Luttrell, 1988; Goodall, 1986; Sicotte, 1995; Silk 1992), or even on behalf of the loser (de Waal, 1978; Watanabe, 2001). The former is often referred to as policing (Flack, de Waal, & Krakauer, 2005; Flack, Girvan, de Waal, & Krakauer, 2006) and serves to maintain group harmony. Loser support may serve the same role, although it also may result from individuals intervening on behalf of their kin or supporters. Note however that although individual motivation to intervene may stem from a desire to end the fighting (an aversive stimulus), punish a competitor (by joining in the fight against him or her) or help an ally, or help kin, the interventions may still *function* to reduce group tension and ultimately restore peace (Flack & de Waal, 2000). This is an example of the importance of carefully disentangling a behavior's function from its underlying mechanism (in this case, the motivation for the intervention).

Another form of conflict resolution that may support group harmony is post-conflict reconciliation or a friendly reunion of former opponents shortly after a disagreement (de Waal, 1989a; de Waal & van Roosmalen, 1979; Flack & de Waal, 2000). Reconciliation has been hypothesized to not only mend the relationship of

shackelf@oakland.edu

the quarreling individuals but also to reduce the stressful effects of fighting on the group by reducing tension between the combatants (Aureli, 1997; Aureli & van Schaik, 1991; de Waal & Aureli, 1996; de Waal & van Roosmalen, 1979; Kappeler & Schaik 1992). Sometimes a mediator is involved in reconciliation, by, for instance, alternately grooming the previous combatants until they are brought together to groom one another (de Waal & van Roosmalen, 1979; Flack & de Waal, 2000). Some have argued that mediation demonstrates that uninvolved third parties can be concerned with repairing relationships that are not their own and shows concern for the community in that all members benefit when tension is reduced (Flack & de Waal, 2000). However, as such interactions have rarely been observed, more research will need to be done to determine what motivates this mediation behavior, as well as its role in maintaining group harmony.

A related behavior is consolation, in which individuals who were not involved in the altercation provide post-conflict support to the loser (Clay & de Waal, 2013; de Waal, 1982; de Waal & Aureli, 1996; de Waal & van Roosmalen, 1979). This behavior has been of particular interest to researchers, as unlike reconciliation, which is widespread across the animal kingdom, consolation seems to be relatively limited (e.g., within the primates to the apes; de Waal & Aureli, 1996). This, and the apparent lack of a personal reason for the consoler to intervene, has led some researchers to argue that consolation is an example of empathetic concern. However, others find that consolation also provides benefits to the consoler. Specifically, acting as a consoler reduces the stress of that individual, which indicates that consolers may be motivated to do so to lower their own stress and anxiety following a group conflict, rather than by empathetic concern (Koski & Sterck, 2007). Whether this is the case, the behavior still serves to help the victim as well and, therefore, may be one of the behaviors that natural selection acted on to promote group harmony.

Reciprocity

Reciprocity is often considered one of the pillars of moral behavior (de Waal, 2006a). First described by Trivers (1971), reciprocity is the idea that the short-term cost of acting to benefit another can be balanced when the recipient of the good deed returns the favor to the agent at some point in the future. In principle, this can occur through direct reciprocity, in which two individuals swap favors, or generalized reciprocity, in which one individual "pays it forward" to another, with a net benefit to all participants (see Rutte & Taborsky, 2007, 2008, for examples in rodents and Claidière et al., 2015, for a possible example in primates). Given the rarity of evidence for generalized reciprocity in other animals, and the fact that in direct comparisons, direct reciprocity is much stronger than generalized reciprocity (Rutte & Taborsky, 2007), we hereafter focus exclusively on direct reciprocity.

Although reciprocal altruism is often assumed to have high cognitive demands, limiting its distribution among species besides humans (Stevens & Hauser, 2004), others have argued that the cognitive demands are much lower and that reciprocity is therefore widespread (e.g., Brosnan, Salwiczek, & Bshary, 2010; Schino &

shackelf@oakland.edu

Aureli, 2010). Brosnan and de Waal (2002) suggest three levels of reciprocity, increasing in their cognitive demands. The most basic is symmetry-based reciprocity, which is noncontingent and requires no comprehension, and results from the naturally parallel distribution of favors when two individuals spend a lot of time together. These individuals are more likely to groom or share food with each other than with others with whom they spend less time, and the result is a correlated distribution. Next is attitudinal reciprocity, which is contingent and partner-specific but does not require comprehension. In this case, an act by individual A toward B will induce a positive feeling in B toward A (e.g., a "warm glow"; Andreoni 1990; de Waal, 2000; de Waal, Leimgruber, & Greenberg, 2008; Harbaugh 1998; Hemelrijk, 1994; Nunes & Schokkaert, 2003), which then makes it more likely that B will behave in a way to benefit A. Note that even in this contingent case, B does not plan to benefit A, but the change in affect makes B more likely to benefit A. However, this only works if the positive feeling is directed specifically at A (e.g., de Waal, 1989b). The most cognitively complex level is calculated reciprocity, in which individuals keep track of favors exchanged and act accordingly. There are few examples of calculated reciprocity in the animal kingdom, and even among humans most reciprocal interactions may not be calculated, so we assume in our consideration of reciprocity that we are discussing one of the other two mechanisms unless we state otherwise.

Among our closest relatives, the nonhuman primates, ample evidence exists for behavior that is reciprocally patterned. The best examples come from the reciprocal exchange of grooming for grooming (Barrett, Henzi, Weingrill, Lycett, & Hill, 1999; Gomes, Mundry, & Boesch, 2009; Manson, Navarrete, Silk, & Perry, 2004; Schino & Aureli, 2008), for agonistic support (Hemelrijk, 1994; Seyfarth & Cheney, 1984), for mating opportunities (de Waal, 1982), and for food (Gilby, 2006; Jaeggi, De Groot, Stevens, & Van Schaik, 2013); the exchange of meat (Gomes & Boesch, 2009) or high-value plant food (Crick, Suchak, Eppley, Campbell, & de Waal, 2013; Kuroda, 1984; Pruetz & Lindshield, 2012) for mating opportunities (but see Gilby, Emery Thompson, Ruane, & Wrangham, 2010); food sharing (de Waal, 1989b, 1997, 2000); and coalitionary support and interventions (de Waal, 1987; de Waal & Luttrell, 1988; Noë, 1986, Packer, 1977; Smuts, 1985).

Contrary to this, however, evidence for contingent reciprocal exchange in a laboratory setting is remarkably weak, which is likely one of the main reasons that the existence of reciprocity is hotly disputed. For instance, two recent studies of chimpanzees showed that subjects were no more likely to share food with a group mate when there was an opportunity for reciprocity than when there was not (Brosnan, Silk, et al., 2009; Melis, Hare, & Tomasello, 2008). Such results seem contrary to the idea that reciprocity is a foundation that is shared with humans and other species. However, several factors indicate that reciprocity is not a behavior that should be expected in all circumstances, and it may be particularly difficult to produce in experimental laboratory settings. First, although the time interval for reciprocation in a natural setting may be hours, weeks, or months, in the laboratory studies take place over much shorter time periods, typically less than 30 min, which may not allow sufficient time for reciprocation to emerge (Brosnan, 2014a). Moreover, experimental subjects rarely have the opportunity to choose their partners, which are chosen for them by the experimenter, whereas in the wild, subjects choose their

shackelf@oakland.edu

own partners (Brosnan, 2014a; Brosnan, Freeman, & de Waal, 2006). Partnerships in laboratory studies are typically chosen based on whether individuals are compatible (e.g., are willing to separate with one another), which may bias results. For instance, subjects who have lived together for longer are more likely to separate together than those who have lived together for shorter periods (Brosnan et al., 2015). Moreover, food sharing studies involving the entire group show that relationship quality influences whether reciprocal food sharing occurs (Silk, Brosnan, Henrich, Lambeth, & Schapiro, 2013).

These results bring up a point that we think is worth a digression: If experimental laboratory studies do not match results seen in the field, why do them? There are several reasons for conducting laboratory studies. First, even the most careful observational studies can only provide evidence of correlation, not causation. To determine causation, controlled experiments that manipulate a single variable at a time are required. Second, laboratory studies allow us to explore behaviors that may be difficult to observe in a natural environment. This may relate to social behaviors (e.g., by exploring the interactions among individuals who would not typically interact in a field setting) or cognition (e.g., deception, metacognition). Most importantly, however, is that field and laboratory studies should complement one another (Janson & Brosnan, 2013). Ideally, laboratory studies will be based on observations of animals' natural behaviors, and findings from the lab will be explored in the field. This may involve complementary studies or observations in the lab and experiments in the field (e.g., theory of mind; Cheney & Seyfarth, 1990). This will lead to a better understanding of both the function and the mechanism of a behavior.

Responses to Inequity

Reciprocity underlies many social interactions such as cooperation, coordination, and collaboration; but in cooperative interactions, unlike reciprocal interactions, both partners have the opportunity to receive immediate benefits (Brosnan, 2014a). In primates, this has most commonly been studied in a cooperative bar-pull task, in which subjects must cooperate to pull in a tray with rewards on it (for a review, see Brosnan, 2011b). Rewards can be distributed in a variety of ways to test the conditions under which partners will work together. These studies have been done most frequently with capuchin monkeys, chimpanzees, and bonobos, all of whom easily learn to complete the task. Moreover, evidence suggests that subjects are sensitive to both their own and their partners' outcomes; in general, they have better success at the task when the partner is more effective at the task (Melis, Hare, & Tomasello, 2006a) and is tolerant and shares the food (Brosnan, Freeman & de Waal, 2006; de Waal & Berger, 2000; Hare, Melis, Woods, Hastings, & Wrangham, 2007; Melis, Hare, & Tomasello, 2006b). Brosnan and Houser et al. (2010) tested capuchins' willingness to pull the apparatus when doing so led to their partner getting more rewards than they did. Subjects were willing to pull when their partner got only a little more than they did, demonstrating their strong tendency to help each other, but

shackelf@oakland.edu

did not do so when it would have meant that their partner got substantially more than they did. Results such as these suggest that the distribution of rewards, or equity, is important to these monkeys.

In fact, responding to inequity has been proposed to underlie the evolution of cooperative behavior (Brosnan, 2006; Fehr & Schmidt, 1999). Once individuals can recognize and judge their own outcomes relative to others', they can change their behavior to improve their outcomes (Brosnan, 2011a), for example, by choosing a cooperative partner who does not take more than his or her fair share. It is relatively simple to explain how an ability to recognize when one was treated inequitably evolved; if subjects recognized when they got less and responded by finding a new partner, they would benefit and the behavior would be selected for. We first consider evidence that primates respond to inequity, or getting less than a social partner. However, does this have anything to do with moral behavior? When considering human behavior, we would see this as self-interested but not fair, which is a critical element for a behavior to be moral (Brosnan, 2014a, 2014b; Brosnan & de Waal, 2014). Following this, then, we will consider whether individuals in other species react to receiving more than a partner and what this tells us about the evolution of fairness and morality.

Several species respond negatively when a social partner is treated better for completing the same task, refusing food rewards or refusing to participate in the task (capuchins, Brosnan & de Waal, 2003; Fletcher, 2008; van Wolkenten, Brosnan, & de Waal, 2007; chimpanzees, Brosnan, Schiff & de Waal, 2005; Brosnan, Talbot, Ahlgren, Lambeth, & Schapiro, 2010; macaques, Hopper, Lambeth, Schapiro, Bernacky, & Brosnan, 2013; Massen, Van Den Berg, Spruijt, & Sterck, 2010; Massen, Van den Berg, Spruijt, & Sterck, 2012; bonobos, Bräuer, Call, & Tomasello, 2009; corvids, Wascher & Bugnyar, 2013; dogs, Horowitz, 2012; Range, Horn, Virányi, & Huber, 2009; Range, Leitner, & Virányi, 2012). In these studies, the typical procedure involves having a subject and a partner complete a species-appropriate task (e.g., returning a token to the experimenter), after which they each receive a reward. We compare subjects' reactions when they are given a less preferred reward than their partner (inequity) to their reactions when the partner gets the same, less preferred reward (equity). In another condition, the subject and partner are initially offered the preferred reward, but after completing the task, they receive only the less preferred one; we compare their reactions in this scenario to their reactions in the prior situations to distinguish reactions based on social comparison from those based on individual expectations. Although in both cases we are measuring the subjects' responses to violations of their expectations, indicating that they are underpinned by a similar cognitive mechanism, in the case of inequity, the referent for the expectation is what their partner got, whereas in the case of individual expectation, also known as contrast (Reynolds, 1961), the referent for the expectation is what they were initially offered. It is critical to disambiguate social versus individual expectations to be sure that one is measuring a response to inequity.

One important point to consider is what refusals mean. Refusing to accept one's own reward when the partner gets a better reward increases both relative and absolute inequity (your partner still gets the preferred food, but now you get nothing) (Henrich, 2004). This response is seen in humans as well, who routinely refuse

shackelf@oakland.edu

outcomes when they are unequal, despite the fact that this refusal does not alter the proposer's outcome (Lamichhane, Adhikari, Brosnan, & Dhamala, 2014; Yamagishi et al., 2009). However, in a naturally occurring situation of cooperation (e.g., outside of the lab), refusing to participate would result in neither individual receiving any reward, which would result in relative equity (although both individuals would be absolutely less well off). Given that the currency of natural selection is relative outcomes, relative equity is more important than absolute gains. Thus, these refusals are likely an artifact of a partner choice mechanism that, in a natural setting, would result in the subject attempting to increase their outcomes by finding another partner.

Studying other species provides two avenues for understanding the evolution of inequity responses in humans. First, we can explore the contexts and conditions that lead to the response, and second, we can consider what factors co-occur with the response to inequity, and thereby gain evidence for the function of the behavior. Considering the former, a key finding is that the inequity response is highly variable (reviewed in Brosnan, 2013). Using chimpanzees as an example, some studies have found evidence for a response to inequity (Brosnan et al., 2005, 2015; Brosnan, Talbot, et al., 2010; Hopper, Lambeth, Schapiro, & Brosnan, 2014b) and others have not (Bräuer, Call & Tomasello, 2006, 2009; Hopper, Lambeth, Schapiro, Bernacky & Brosnan, 2013), but this cannot be entirely explained by differences between groups, as in no study has every subject responded. Considering the factors that influence responses, the presence of a task, such as exchanging a token, is important. Although not every study involving a task finds a response to inequity (e.g., Bräuer et al, 2009; Silberberg, Crescimbene, Addessi, Anderson, & Visalberghi, 2009), no study without a task has demonstrated an inequity reaction (capuchins, Dindo & de Waal, 2006; Dubreuil, Gentile, & Visalberghi, 2006; Roma, Silberberg, Ruggiero, & Suomi, 2006; chimpanzees, Bräuer et al., 2006), and in studies that include conditions with and without a task, inequity responses only occur when a task is present (Brosnan et al., 2015; Brosnan, Talbot, et al., 2010; van Wolkenten, Brosnan & de Waal, 2007). It is uncertain why this is so. One possibility is that subjects are used to being handed food for "free" by experimenters and so do not respond. A more intriguing hypothesis is that if the response to inequity evolved in conjunction with cooperation, subjects only react in contexts that are perceived as cooperative; in this explanation, being handed food is not seen as cooperative, whereas working to obtain it is. We consider the function of the response to inequity in more detail below.

Other factors also affect responses, but their importance to the inequity response is less clear. For instance, dominance plays a role, with dominant subjects responding more strongly to inequity than their subordinate partners (Bräuer et al., 2006; Brosnan, Talbot, et al., 2010). However, this has not been seen in all studies, making it unclear whether this is restricted to some species (e.g., chimpanzees) or is part of a suite of factors that work together to influence responses. Likewise, relationship quality is an important influence in some studies related to cooperation and food sharing (Melis et al., 2006a, 2006b; Silk et al., 2013), but it is not clear the degree to which it influences responses to inequity. Two studies have found evidence that the length of time subjects have lived together affects responses to inequity (Brosnan et al., 2005; Hopper et al., 2014), but another study that controlled for

shackelf@oakland.edu

both relationship and personality found that personality but not relationship quality or duration was a significant predictor of responses (Brosnan et al., 2015). Finally, although age has not been a significant predictor of adult subjects' responses in any study that included it, the one study to test juveniles found that the response was not present at 1 year of age, but appeared by 27 months (Hopper et al., 2013). The take-home message from these studies is that, like humans, other species' responses are influenced by their environment and other factors. Future research is needed to determine which factors are important in which circumstances and how they interact.

At this point in time, we have data on 10 species' responses to inequity, not including humans. This allows us to use the comparative approach to consider the function of the inequity response. The distribution of data demonstrates that the response is not a homology in the primates (Talbot, Freeman, Williams & Brosnan, 2011). Given this, we can investigate which factors co-occur with the response to inequity, which provides evidence that it evolved convergently. Current evidence indicates that the response is not an artifact of brain size (large-brained orangutans do not respond to inequity; Brosnan et al., 2011; Bräuer et al., 2006, 2009) or living in a social group (highly gregarious squirrel monkeys do not respond to inequity; Freeman et al., 2013; Talbot et al., 2011). The one factor that does co-occur with responses to inequity is the frequency with which subjects cooperate with non-kin (Price & Brosnan, 2012; see in particular Table 2, which outlines studies to date). Species that do so routinely, engaging in cooperative hunting, coalitions and alliances, and/or food sharing, including chimpanzees, bonobos, macaques (rhesus and long tailed), and capuchin monkeys, show evidence of responding to inequity. Species that do not do so as routinely, including orangutans and squirrel monkeys, do not respond to inequity. Although we cannot know from this correlation which came first, the comparative evidence supports earlier hypotheses of the importance of inequity responses for cooperation.

In fact, one point of support comes from an exception that "proves the rule." The cooperative breeding hypothesis proposes that cooperative breeders, species in which both parents and, often, adult offspring are required to rear the young, have a unique set of behaviors due to the constraints of this highly interdependent system (Hrdy, 2009; Burkhart & van Schaik, 2010). In particular, they are argued to be exceptionally prosocial due to the importance of their partners. The cooperative breeding hypothesis also suggests that for these species, the cost of finding a new partner is sufficiently high that subjects should accept less costly inequity to a greater degree than do other species. Thus far, three species that are either cooperative breeders or show biparental care (e.g., both parents are obligate care providers for offspring, but help from their adult offspring is not essential), marmosets, tamarins, and owl monkeys, do not react negatively to experimentally induced inequity (Freeman et al., 2013; Neiworth, Johnson, Whillock, Greenberg, & Brown, 2009). It may be that for these species, the costs of finding a new partner are sufficiently high that they are more tolerant of inequity. It would be interesting to investigate whether these species are more sensitive to inequity during partner formation than they are after partnerships are established and offspring are present (Brosnan, 2013). Additionally, although punishment is rare in primates (Jensen, 2010; Raihani, Thornton, & Bshary,

shackelf@oakland.edu

2012), it may be that species with limited partner choice are more likely to engage in partner control or punishment (Raihani & McAuliffe, 2012).

What, however, of subjects who get *more* than a partner? All of the above addresses how subjects respond when they get less than a partner. This response is the likely first step in the evolution of inequity responses (Brosnan & de Waal, 2014). However, as discussed above, when considering humans, we expect more than this; to qualify as a true sense of fairness, subjects must also respond negatively to inequity that favors them (Brosnan & de Waal, 2014). Thus in order to more fully understand the evolution of fairness, it is necessary to explore whether other species also are sensitive to being advantaged. Although there is little evidence for this response outside of humans, there is evidence in chimpanzees. Chimpanzees are more likely to refuse a preferred reward if their partner gets a less preferred one as compared to when their partner also gets the more preferred reward. This response is less strong than their response when they receive less than the partner, and from this study we do not know the mechanism for the response, but nonetheless it shows that they notice when they get more than a partner (Brosnan, Talbot, et al., 2010). Moreover, in some cases chimpanzees change their preferences from distributions that favor themselves to those that are equally split (e.g., in an ultimatum game: Proctor, Williamson, Latzman, de Waal, & Brosnan, 2014, but see also Jensen, Call, & Tomasello, 2007). Notably, however, this occurs only in situations in which their partner has recourse to block both subjects from receiving any reward. Thus, it has recently been hypothesized that this response evolved to help subjects maintain beneficial social interactions by keeping inequity from becoming too great; in these cases, it is worth the short-term cost to a subject to rectify inequity to maintain the long-term benefits of a cooperative relationship (Brosnan & de Waal, 2014).

This raises an interesting point regarding moral behavior; most people consider moral behavior to be good behavior that benefits others rather than the self. However, no trait, including moral behavior, will be selected unless it provides some benefit. In this case, fairness is a moral behavior that we celebrate, yet it evolved for a purpose that benefits the self: to maintain beneficial relationships. Understanding why moral behaviors evolved and what purpose they served may help us to increase behavior that is moral by setting up society in ways that promote its occurrence.

Empathy

Group living animals are aware of their partners to some degree (e.g., to their presence, actions, emotions, etc.). Whether this extends to the level of empathy is another question. In part, this depends on the definition of empathy; many scientists treat empathy as a continuum ranging from basic emotional contagion, which does not involve awareness on the part of the animal, to so-called cognitive empathy, in which the individual is aware of their interest in the other's well-being (Preston & de Waal, 2002). Anecdotes abound of actions by animals that may indicate empathetic concern (Bekoff, 2000; de Waal, 2006b), but as discussed above, without

shackelf@oakland.edu

controlled experiments, it is not possible to identify the underlying mechanisms for these behaviors.

Beyond anecdotal examples of animals helping one another, there are other common behaviors that may be indicative of empathetic concern. For instance, when individuals interact with others who are younger or disabled, they tend to adjust their own behavior to accommodate the limits of the other (de Waal, 1996; de Waal, Uno, Luttrell, Meisner, & Jeanotte, 1996; Fedigan & Fedigan, 1977; Hayaki, 1985). This learned adjustment might reflect empathy as it demonstrates that individuals may understand the needs of others (Flack & de Waal, 2000). Other behaviors that indicate concern for others are succorant behavior (Scott, 1971) and consolation (Clay & de Waal, 2013; de Waal & van Roosmalen, 1979). As we mentioned above, it is possible that consolation is motivated by reducing an aversive stimulus and calming the consoler rather than the victim (Koski & Sterck, 2007), but the behavior may nonetheless function to calm the victim (Brosnan, 2014b).

Experimental evidence of empathy in nonhuman species is rather sparse. From the beginning, studies attempting to find evidence of empathy have focused on whether animals take action to alleviate distress or pain in others. The first such evidence emerged in rats and monkeys in the 1950s and 1960s (Church, 1959; Wechkin, Masserman, & Terris, 1964), but has been challenged on the grounds that the subjects helped their partners to stop the signals (e.g., calls) given by the distressed animals, which are aversive. By this interpretation, the *motivation* of the animals was not based on empathetic concern, although the functional outcome of the behavior nonetheless served to help their partners. More recent evidence shows that mice respond more strongly to pain when they are housed with another mouse also experiencing pain (Langford et al., 2006), but that this response depends on the mouse's genetics, suggesting selection for this behavior (Chen, Panksepp & Lahvis, 2009). An elegant series of studies in rats demonstrates that rats will free co-housed social partners, even releasing them prior to opening another container with chocolate (a favored food that they then must share; Bartal, Decety, & Mason, 2011). Moreover, this behavior is based on social experience, with rats only freeing those of strains with which they have previous familiarity (intriguingly, this is independent of the subject's own strain), showing that this is not a reflexive response but has a social component (Bartal, Rodgers, Bernardez Sarria, Decety, & Mason, 2014).

Thus, like humans (Jackson, Rainville, & Decety, 2006), other species show evidence of behaviors that may be based on empathetic concern. Whether these behaviors in other species are due to a form of empathy, and if they are, whether it is homologous with the empathy seen in humans, is still a matter of debate, but empathy does provide a plausible mechanism for the behaviors seen above (de Waal, 2008; Yamamoto & Takimoto, 2012), and many have argued that the null assumption in cases of shared behavior across species should be homology (i.e., we should consider related outcomes in related species to be the result of shared mechanisms until it is shown otherwise; de Waal, 2006b). Even if it is not developed to the degree seen in humans, or even if the underlying mechanisms turn out to be different, the presence of behavior that functions to relieve others' distress, or empathy, may provide us with insight into the pathway by which human empathy evolved.

shackelf@oakland.edu

Conclusion: The Importance of Comparative Research for Understanding Human Morality

What can studying morality in other species tell us about humans? By seeking evidence of precursors to moral behavior in our closest living phylogenetic relatives, we can learn more about how and why these traits evolved, how they are constrained, and which factors of social life are most important to maintain. Data from a wide variety of species are upending long-held beliefs about the degree to which morality sets humans apart from other animals. Although human moral systems are undoubtedly more complex than in animals, they share commonalities in function and, possibly, outcome, which allow us to explore the roots of our own moral behavior. Understanding these roots may be critical if we are to understand why humans make the decisions that we do, or to understand situations in which it is difficult to produce moral behavior (and, hopefully, change the context to elicit more preferred outcomes). By continuing to study and compare the precursors of moral behavior in other species, we learn more not only about these other animals but also about ourselves.

References

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal, 100*, 464–477.

Aureli, F. (1997). Post-conflict anxiety in nonhuman primates: The mediating role of emotion in conflict resolution. *Aggressive Behavior, 23*(5), 315–328.

Aureli, F., & van Schaik, C. P. (1991). Post-conflict behaviour in long-tailed macaques (*Macaca fascicularis*). *Ethology, 89*(2), 101–114.

Barrett, L., Henzi, S. P., Weingrill, T., Lycett, J. E., & Hill, R. A. (1999). Market forces predict grooming reciprocity in female baboons. *Proceedings of the Royal Society B: Biological Sciences, 266*(1420), 665–670.

Bartal, I. B. A., Decety, J., & Mason, P. (2011). Empathy and pro-social behavior in rats. *Science, 334*(6061), 1427–1430.

Bartal, I. B. A., Rodgers, D. A., Bernardez Sarria, M. S., Decety, J., & Mason, P. (2014). Pro-social behavior in rats is modulated by social experience. *eLife, 3*, e01385. doi:10.7554/eLife.01385.

Bekoff, M. (Ed.). (2000). *The smile of a dolphin: Remarkable accounts of animal emotions*. New York: Random House/Discovery Books.

Bercovitch, F. B., Hauser, M. D., & Jones, J. H. (1995). The endocrine stress response and alarm vocalizations in rhesus macaques. *Animal Behaviour, 49*, 1703–1706.

Bernstein, I. S. (1981). Dominance: The baby and the bathwater. *Behavioural and Brain Sciences, 4*, 419–458.

Blumstein, D. T., Patton, M. L., & Saltzman, W. (2006). Faecal glucocorticoid metabolites and alarm calling in free-living yellow-bellied marmots. *Biology Letters, 2*, 29–32.

Boinski, S., Gross, T. S., & Davis, J. K. (1999). Terrestrial predator alarm vocalizations are a valid monitor of stress in captive brown capuchins (*Cebus apella*). *Zoo Biology, 18*, 295–312.

Bonnie, K. E., & de Waal, F. B. (2004). Primate social reciprocity and the origin of gratitude. In R. A. Emmons & M. E. McCullough (Eds.), *The psychology of gratitude*. Oxford: Oxford University Press.

Bräuer, J., Call, J., & Tomasello, M. (2006). Are apes really inequity averse? *Proceedings of the Royal Society B: Biological Sciences, 273*(1605), 3123–3128.

shackelf@oakland.edu

Bräuer, J., Call, J., & Tomasello, M. (2009). Are apes inequity averse? New data on the token-exchange paradigm. *American Journal of Primatology, 71*(2), 175–181.
Brosnan, S. F. (2006). Nonhuman species' reactions to inequity and their implications for fairness. *Journal of Social Justice, 19*, 153–185.
Brosnan, S. F. (2008). The evolution of inequity. In D. Houser & K. McCabe (Eds.), Neuroeconomics (pp. 99–124). Bingley, UK: Emerald.
Brosnan, S. F. (2011a). An evolutionary perspective on morality. *Journal of Economic Behavior & Organization, 77*(1), 23–30. doi:10.1016/j.jebo.2010.04.008.
Brosnan, S. F. (2011b). A hypothesis of the co-evolution of cooperation and responses to inequity. *Frontiers in Neuroscience, 5*, 43.
Brosnan, S. F. (2013). Justice and fairness related behaviors in non-human primates. *Proceedings of the National Academy of Sciences, 110*, 10416–10423. doi:10.1073/pnas.1301194110.
Brosnan, S. F. (2014a). Why an evolutionary perspective is critical to understanding moral behavior in humans. In M. Bergmann & P. P. Kain (Eds.), *Challenges to moral and religious belief: Disagreement and evolution*. Oxford: Oxford University Press.
Brosnan, S. F. (2014b). Precursors of morality—Evidence for moral behaviors in non-human primates. In M. Christen, C. van Schaik, J. Fischer, M. Huppenbauer, & C. Tanner (Eds.), *Empirically informed ethics: Morality between facts and norms* (pp. 85–98). New York: Springer International Publishing.
Brosnan, S. F., & de Waal, F. B. M. (2002). Variations on tit-for-tat: Proximate mechanisms of cooperation and reciprocity. *Human Nature, 13*(1), 129–152.
Brosnan, S. F., & de Waal, F. B. M. (2003). Monkeys reject unequal pay. *Nature, 425*, 297–299.
Brosnan, S. F., & de Waal, F. B. M. (2014). Evolution of responses to (un) fairness. *Science, 346*(6207), 1251776.
Brosnan, S.F., Flemming, T., Talbot, C.F., Mayo, L. & Stoinski, T. (2011). Orangutans (Pongo pygmaeus) do not form expectations based on their partner's outcomes. *Folia Primatologica, 82*(1), 56–70.
Brosnan, S. F., Freeman, C., & de Waal, F. (2006). Partner's behavior, not reward distribution, determines success in an unequal cooperative task in capuchin monkeys. *American Journal of Primatology, 68*(7), 713–724.
Brosnan, S. F., Hopper, L. M., Richey, S., Freeman, H. D., Talbot, C. F., Gosling, S. D., et al. (2015). Personality influences responses to inequity and contrast in chimpanzees. *Animal Behavior, 101*, 75–87. doi:10.1016/j.anbehav.2014.12.019.
Brosnan, S. F., Houser, D., Leimgruber, K., Xiao, E., Chen, T., & de Waal, F. B. M. (2010). Competing demands of prosociality and equity in monkeys. *Evolution and Human Behavior, 31*, 279–288.
Brosnan, S. F., Newton-Fisher, N. E., & Van Vugt, M. (2009). A melding of the minds: When primatology meets personality and social psychology. *Personality and Social Psychology Review, 13*(2), 129–147.
Brosnan, S. F., Salwiczek, L., & Bshary, R. (2010). The interplay of cognition and cooperation. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 365*(1553), 2699–2710. doi:10.1098/rstb.2010.0154.
Brosnan, S. F., Schiff, H. C., & de Waal, F. B. M. (2005). Tolerance for inequity may increase with social closeness in chimpanzees. *Proceedings of the Royal Society B: Biological Sciences, 272*(1560), 253–258.
Brosnan, S. F., Silk, J. B., Henrich, J., Mareno, M. C., Lambeth, S. P., & Schapiro, S. J. (2009). Chimpanzees (*Pan troglodytes*) do not develop contingent reciprocity in an experimental task. *Animal Cognition, 12*(4), 587–597. doi:10.1007/s10071-009-0218-z.
Brosnan, S. F., Talbot, C., Ahlgren, M., Lambeth, S. P., & Schapiro, S. J. (2010). Mechanisms underlying responses to inequitable outcomes in chimpanzees, *Pan troglodytes. Animal Behaviour, 79*, 1229–1327. doi:10.1016/j.anbehav.2010.02.019.
Burkhart, J.M. & van Schaik, C.P. (2010). Cognitive consequences of cooperative breeding in primates? *Animal Cognition, 13*(1), 1–19.

shackelf@oakland.edu

Chen, Q., Panksepp, J. B., & Lahvis, G. P. (2009). Empathy is moderated by genetic background in mice. *PLoS One, 4*(2), e4387.
Cheney, D. L., & Seyfarth, R. M. (1990). *How monkeys see the world: Inside the mind of another species*. Chicago: University of Chicago Press.
Church, R. M. (1959). Emotional reactions of rats to the pain of others. *Journal of Comparative and Physiological Psychology, 52*(2), 132.
Claidière, N., Whiten, A., Mareno, M. C., Messer, E. J. E., Brosnan, S. F., Hopper, L. M., et al. (2015). Selective and contagious prosocial resource donation in capuchin monkeys, chimpanzees and humans. *Scientific Reports, 5*, 7631. doi:10.1038/srep07631.
Clay, Z., & de Waal, F. B. (2013). Bonobos respond to distress in others: Consolation across the age spectrum. *PLoS One, 8*(1), e55206.
Crick, J., Suchak, M., Eppley, T. M., Campbell, M. W., & de Waal, F. B. (2013). The roles of food quality and sex in chimpanzee sharing behavior (*Pan troglodytes*). *Behaviour, 150*(11), 1203–1224.
Darwin, C. (1871). *The descent of man and selection in relation to sex*. New York: D. Appleton & Company.
de Waal, F. B. M. (1989a). *Peacemaking among primates*. Cambridge, MA: Harvard University Press.
de Waal, F. B. M., & Aureli, F. (1996). Consolation, reconciliation, and a possible cognitive difference between macaques and chimpanzees. In A. E. Russon, K. A. Bard, & S. T. Parker (Eds.), *Reaching into thought: The minds of the great apes* (pp. 80–110). Cambridge, UK: Cambridge University Press.
de Waal, F. B., Leimgruber, K., & Greenberg, A. R. (2008). Giving is self-rewarding for monkeys. *Proceedings of the National Academy of Sciences, 105*(36), 13685–13689.
de Waal, F. B. M. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology, 59*, 279–300.
de Waal, F. B. M., & van Roosmalen, A. (1979). Reconciliation and consolation among chimpanzees. *Behavioral Ecology and Sociobiology, 5*(1), 55–66.
de Waal, F., Uno, H., Luttrell, L. M., Meisner, L. F., & Jeanotte, L. A. (1996). Behavioral retardation in a macaque with autosomal trisomy and aging mother. *American Journal on Mental Retardation, 100*(4), 378–390.
de Waal, F. B. M. (1978). Exploitative and familiarity-dependent support strategies in a colony of semi-free living chimpanzees. *Behaviour, 66*, 268–312.
de Waal, F. B. M. (1982). *Chimpanzee politics: Sex and power among apes*. London: Jonathan Cape.
de Waal, F. B. M. (1984). Sex differences in the formation of coalitions among chimpanzees. *Ethology and Sociobiology, 5*(4), 239–255.
de Waal, F. B. M. (1987). Tension regulation and nonreproductive functions of sex in captive bonobos (*Pan paniscus*). *National Geographic Research, 3*(3), 318–335.
de Waal, F. B. M. (1989b). Food sharing and reciprocal obligations among chimpanzees. *Journal of Human Evolution, 18*(5), 433–459.
de Waal, F. B. M. (1996). *Good natured: The origins of right and wrong in primates and other animals*. Cambridge, MA: Harvard University Press.
de Waal, F. B. M. (1997). The chimpanzee's service economy: Food for grooming. *Evolution and Human Behavior, 18*(6), 375–386.
de Waal, F. B. M. (2000). Attitudinal reciprocity in food sharing among brown capuchin monkeys. *Animal Behaviour, 60*(2), 253–261.
de Waal, F. B. M. (2006a). The animal roots of human morality. *New Scientist, 192*(2573), 60–61.
de Waal, F. B. M. (2006b). *Primates and philosophers*. Princeton, NJ: Princeton University Press.
de Waal, F. B. M., & Berger, M. L. (2000). Payment for labour in monkeys. *Nature, 404*(6778), 563.
de Waal, F. B. M., & Luttrell, L. M. (1988). Mechanisms of social reciprocity in three primate species: Symmetrical relationship characteristics or cognition? *Ethology and Sociobiology, 9*(2), 101–118.
Dindo, M., & de Waal, F. (2006). Partner effects on food consumption in brown capuchin monkeys. *American Journal of Primatology, 69*(4), 448–456.

shackelf@oakland.edu

Dubreuil, D., Gentile, M. S., & Visalberghi, E. (2006). Are capuchin monkeys (*Cebus apella*) inequity averse? *Proceedings of the Royal Society B: Biological Sciences, 273*(1591), 1223–1228.
Eisenberg, N. (1992). *The caring child* (Vol. 28). Cambridge, MA: Harvard University Press.
Eisenberg, N., & Mussen, P. H. (Eds.). (1989). *The roots of prosocial behavior in children*. Cambridge, UK: Cambridge University Press.
Fedigan, L. M., & Fedigan, L. (1977). The social development of a handicapped infant in a free-living troop of Japanese monkeys. In S. Chevalier-Skolnikoff & R. E. Poirier (Eds.), *Primate bio-social development: Biological, social and ecological determinants* (pp. 205–222). New York: Garland.
Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics, 114*, 817–868.
Fessler, D. M. (2009). Return of the lost letter: Experimental framing does not enhance altruism in an everyday context. *Journal of Economic Behavior & Organization, 71*(2), 575–578.
Flack, J. C., Girvan, M., de Waal, F. B., & Krakauer, D. C. (2006). Policing stabilizes construction of social niches in primates. *Nature, 439*(7075), 426–429.
Flack, J. C., de Waal, F. B. M., & Krakauer, D. C. (2005). Social structure, robustness, and policing cost in a cognitively sophisticated species. *The American Naturalist, 165*(5), E126–E139.
Flack, J. C., & de Waal, F. B. M. (2000). 'Any animal whatever'. Darwinian building blocks of morality in monkeys and apes. *Journal of Consciousness Studies, 7*(1–2), 1–29.
Fletcher, G. E. (2008). Attending to the outcome of others: Disadvantageous inequity aversion in male capuchin monkeys (*Cebus apella*). *American Journal of Primatology, 70*(9), 901–905.
Fragaszy, D. M., & Perry, S. (Eds.). (2008). *The biology of traditions: Models and evidence*. Cambridge, UK: Cambridge University Press.
Freeman, H., Sullivan, J., Hopper, L., Talbot, C., Holmes, A., Schultz-Darken, N., et al. (2013). Different responses to reward comparisons by three primate species. *PLoS One, 8*, e76297. doi:10.1371/journal.pone.0076297.
Gilby, I. C. (2006). Meat sharing among the Gombe chimpanzees: Harassment and reciprocal exchange. *Animal Behaviour, 71*(4), 953–963.
Gilby, I. C., Emery Thompson, M., Ruane, J. D., & Wrangham, R. (2010). No evidence of short-term exchange of meat for sex among chimpanzees. *Journal of Human Evolution, 59*(1), 44–53.
Gomes, C. M., & Boesch, C. (2009). Wild chimpanzees exchange meat for sex on a long-term basis. *PLoS One, 4*(4), e5116.
Gomes, C. M., Mundry, R., & Boesch, C. (2009). Long-term reciprocation of grooming in wild West African chimpanzees. *Proceedings of the Royal Society B: Biological Sciences, 276*(1657), 699–706.
Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behavior*. Cambridge, MA: Belknap Press.
Hagan, E.H. & Hammerstein, P. (2006). Game theory and human evolution: A critique of some recent interpretations of experimental games. *Theoretical Population Biology, 69*(3), 339–348.
Haidt, J. (2003). The moral emotions. *Handbook of affective sciences* (pp. 852–870). Oxford: Oxford University Press.
Harbaugh, W. T. (1998). What do donations buy? A model of philanthropy based on prestige and warm glow. *Journal of Public Economics, 67*(2), 269–284.
Hare, B., Melis, A. P., Woods, V., Hastings, S., & Wrangham, R. (2007). Tolerance allows bonobos to outperform chimpanzees on a cooperative task. *Current Biology, 17*(7), 619–623.
Hayaki, H. (1985). Social play of juvenile and adolescent chimpanzees in the Mahale Mountains National Park, Tanzania. *Primates, 26*(4), 343–360.
Hemelrijk, C. K. (1994). Support for being groomed in long-tailed macaques, *Macaca fascicularis. Animal Behaviour, 48*(2), 479–481.
Henrich, J. (2004). Inequity aversion in capuchins? *Nature*. 428, 139. doi:10.1038/428139a.
Hopper, L. M., Lambeth, S. P., Schapiro, S. J., Bernacky, B. J., & Brosnan, S. F. (2013). The ontogeny of social comparisons by rhesus macaques (*Macaca mulatta*). *Journal of Primatology, 2*, 109. doi:10.4172/2167-6801.1000109.

shackelf@oakland.edu

Hopper, L. M., Lambeth, S. P., Schapiro, S. J., & Brosnan, S. F. (2014). Social comparison mediates chimpanzees' responses to loss, not frustration. *Animal Cognition, 17*, 1303–1311. doi:10.1007/s10071-014-0765-9.

Horowitz, A. (2012). Fair is fine, but more is better: Limits to inequity aversion in the domestic dog. *Social Justice Research, 25*(2), 195–212.

Hrdy, S. B. (2009). Meet the alloparents: Shared child care may be the secret of human evolutionary success. *Natural History, 118*(3), 24–29.

Jackson, P. L., Rainville, P., & Decety, J. (2006). To what extent do we share the pain of others? Insight from the neural bases of pain empathy. *Pain, 125*(1), 5–9.

Jaeggi, A. V., De Groot, E., Stevens, J. M., & Van Schaik, C. P. (2013). Mechanisms of reciprocity in primates: Testing for short-term contingency of grooming and food sharing in bonobos and chimpanzees. *Evolution and Human Behavior, 34*(2), 69–77.

Janson, C. H., & Brosnan, S. F. (2013). Experiments in primatology: From the lab to the field and back again. In M. Blair, E. Sterling, & N. Bynum (Eds.), *Primate ecology and conservation: A handbook of techniques*. Oxford: Oxford University Press.

Jensen, K. (2010). Punishment and spite, the dark side of cooperation. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 365*(1553), 2635–2650. doi:10.1098/rstb.2010.0146.

Jensen, K., Call, J., & Tomasello, M. (2007). Chimpanzees are rational maximizers in an ultimatum game. *Science, 318*(5847), 107–109.

Kappeler, P. M., & Schaik, C. P. (1992). Methodological and evolutionary aspects of reconciliation among primates. *Ethology, 92*(1), 51–69.

Konner, M. (1982). *The tangled wing: Biological constraints on the human spirit*. New York: Henry Holt.

Koski, S. E., & Sterck, E. H. (2007). Triadic postconflict affiliation in captive chimpanzees: Does consolation console? *Animal Behaviour, 73*(1), 133–142.

Kuroda, S. (1984). Interaction over food among pygmy chimpanzees. In R. L. Susman (Ed.), *The pygmy chimpanzee* (pp. 301–324). Springer US.

Laland, K. N., & Galef, B. G. (Eds.). (2009). *The question of animal culture*. Cambridge, MA: Harvard University Press.

Lamichhane, B., Adhikari, B. M., Brosnan, S. F., & Dhamala, M. (2014). The neural basis of perceived unfairness in economic exchanges. *Brain Connectivity*, 140919140925001. doi:10.1089/brain.2014.0243.

Langford, D. J., Crager, S. E., Shehzad, Z., Smith, S. B., Sotocinal, S. G., Levenstadt, J. S., et al. (2006). Social modulation of pain as evidence for empathy in mice. *Science, 312*(5782), 1967–1970.

Machida, S. (2006). Threat calls in alliance formation by members of a captive group of Japanese macaques. *Primates, 31*(2), 205–211.

Manson, J. H., Navarrete, D. C., Silk, J. B., & Perry, S. (2004). Time-matched grooming in female primates? New analyses from two species. *Animal Behaviour, 67*(3), 493–500.

Massen, J. J., Van Den Berg, L. M., Spruijt, B. M., & Sterck, E. H. (2010). Generous leaders and selfish underdogs: Pro-sociality in despotic macaques. *PLoS One, 5*(3), e9734.

Massen, J. J., Van den Berg, L. M., Spruijt, B. M., & Sterck, E. H. (2012). Inequity aversion in relation to effort and relationship quality in long-tailed Macaques (*Macaca fascicularis*). *American Journal of Primatology, 74*(2), 145–156.

Melis, A. P., Hare, B., & Tomasello, M. (2006a). Chimpanzees recruit the best collaborators. *Science, 311*(5765), 1297–1300.

Melis, A. P., Hare, B., & Tomasello, M. (2006b). Engineering cooperation in chimpanzees: Tolerance constraints on cooperation. *Animal Behaviour, 72*(2), 275–286.

Melis, A. P., Hare, B., & Tomasello, M. (2008). Do chimpanzees reciprocate received favours? *Animal Behaviour, 76*(3), 951–962.

Mendoza, S. P., & Barchas, P. R. (1983). Behavioural processes leading to linear hierarchies following group formation in rhesus monkeys. *Journal of Human Evolution, 12*, 185–192.

shackelf@oakland.edu

Neiworth, J. J., Johnson, E. T., Whillock, K., Greenberg, J., & Brown, V. (2009). Is a sense of inequity an ancestral primate trait? Testing social inequity in cotton top tamarins (*Saguinus oedipus*). *Journal of Comparative Psychology, 123*(1), 10.

Noë, R. (1986). Lasting alliances among adult male savannah baboons. *Primate ontogeny, cognition and social behaviour* (pp. 381–392).

Nunes, P. A., & Schokkaert, E. (2003). Identifying the warm glow effect in contingent valuation. *Journal of Environmental Economics and Management, 45*(2), 231–245.

Packer, C. (1977). Reciprocal altruism in *Papio anubis*. *Nature, 265*, 441–443.

Preston, S. D., & de Waal, F. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25*(01), 1–20.

Preuschoft, S., & van Schaik, C. P. (2000). Dominance and communication. *Conflict management in various social settings* (pp. 77–105).

Price, S. A., & Brosnan, S. F. (2012). To each according to his need? Variability in the responses to inequity in non-human primates. *Social Justice Research, 25*, 140–169.

Proctor, D., Williamson, R. A., Latzman, R. D., de Waal, F. B. M., & Brosnan, S. F. (2014). Gambling primates: Reactions to a modified Iowa Gambling Task in humans, chimpanzees, and capuchin monkeys. *Animal Cognition, 17*, 983–995. doi:10.1007/s10071-014-0730-7.

Pruetz, J. D., & Lindshield, S. (2012). Plant-food and tool transfer among savanna chimpanzees at Fongoli, Senegal. *Primates, 53*(2), 133–145.

Raihani, N. J., & McAuliffe, K. (2012). Does inequity aversion motivate punishment? Cleaner fish as a model system. *Social Justice Research, 25*(2), 213–231. doi:10.1007/s11211-012-0157-8.

Raihani, N. J., Thornton, A. N., & Bshary, R. (2012). Punishment and cooperation in nature. *Trends in Ecology & Evolution*. doi:10.1016/j.tree.2011.12.004.

Range, F., Horn, L., Viranyi, Z., & Huber, L. (2009). The absence of reward induces inequity aversion in dogs. *Proceedings of the National Academy of Sciences, 106*(1), 340–345.

Range, F., Leitner, K., & Virányi, Z. (2012). The influence of the relationship and motivation on inequity aversion in dogs. *Social Justice Research, 25*(2), 170–194.

Reynolds, G. S. (1961). Behavioral contrast. *Journal of the Experimental Analysis of Behavior, 4*, 441–466.

Roma, P. G., Silberberg, A., Ruggiero, A. M., & Suomi, S. J. (2006). Capuchin monkeys, inequity aversion, and the frustration effect. *Journal of Comparative Psychology, 120*(1), 67.

Rutte, C., & Taborsky, M. (2007). Generalized reciprocity in rats. *PLoS Biology, 5*(7), e196.

Rutte, C., & Taborsky, M. (2008). The influence of social experience on cooperative behaviour of rats (*Rattus norvegicus*): Direct vs generalised reciprocity. *Behavioral Ecology and Sociobiology, 62*(4), 499–505.

Schino, G., & Aureli, F. (2008). Grooming reciprocation among female primates: A meta-analysis. *Biology Letters, 4*(1), 9–11.

Schino, G., & Aureli, F. (2010). Primate reciprocity and its cognitive requirements. *Evolutionary Anthropology: Issues, News, and Reviews, 19*(4), 130–135.

Scott, J. F. (1971). *Internalization of norms: A sociological theory of moral commitment*. Englewood Cliffs, NJ: Prentice Hall.

Seyfarth, R. M., & Cheney, D. L. (1984). Grooming, alliances and reciprocal altruism in vervet monkeys. *Nature, 308*, 541–543.

Sicotte, P. (1995). Interpositions in conflicts between males in bimale groups of mountain gorillas. *Folia Primatologica, 65*(1), 14–24.

Silberberg, A., Crescimbene, L., Addessi, E., Anderson, J. R., & Visalberghi, E. (2009). Does inequity aversion depend on a frustration effect? A test with capuchin monkeys (*Cebus apella*). *Animal Cognition, 12*(3), 505–509.

Silk, J. B. (1992). The patterning of intervention among male bonnet macaques: Reciprocity, revenge, and loyalty. *Current Anthropology, 33*, 318–325.

Silk, J. B., Brosnan, S. F., Henrich, J., Lambeth, S. P., & Schapiro, S. (2013). Chimpanzees share food for many reasons: The role of kinship, reciprocity, social bonds and harassment on food transfers. *Animal Behaviour, 85*(5), 941–947. doi:10.1016/j.anbehav.2013.02.014.

shackelf@oakland.edu

Smith, A. (1817 [1759]). *The theory of moral sentiments* (Vols. I & II). Boston: Wells and Lilly.
Smuts, B. B. (1985). *Sex and friendship in baboons*. New York: Aldine Publishing Co.
Stevens, J. R., & Hauser, M. D. (2004). Why be nice? Psychological constraints on the evolution of cooperation. *Trends in Cognitive Sciences, 8*(2), 60–65.
Talbot, C. F., Freeman, H. D., Williams, L. E., & Brosnan, S. F. (2011). Squirrel monkeys' response to inequitable outcomes indicates a behavioural convergence within the primates. *Biology Letters, 7*(5), 680–682. doi:10.1098/rsbl.2011.0211.
Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie, 20*(4), 410–433.
Tomasello, M., & Vaish, A. (2013). Origins of human cooperation and morality. *Annual Review of Psychology, 64*, 231–255.
Trivers, R. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46*(1), 35–57.
Trivers, R. (2006). Reciprocal altruism: 30 years later. In P. Kappeler & C. P. van Schaik (Eds.), *Cooperation in primates and humans* (pp. 67–83). Berlin, Heidelberg/Germany: Springer.
Trommsdorff, G., Friedlmeier, W., & Mayer, B. (2007). Sympathy, distress, and prosocial behavior of preschool children in four cultures. *International Journal of Behavioral Development, 31*(3), 284–293.
van Wolkenten, M., Brosnan, S. F., & de Waal, F. B. (2007). Inequity responses of monkeys modified by effort. *Proceedings of the National Academy of Sciences, 104*(47), 18854–18859.
Wascher, C. A., & Bugnyar, T. (2013). Behavioral responses to inequity in reward distribution and working effort in crows and ravens. *PLoS One, 8*(2), e56885.
Watanabe, K. (2001). A review of 50 years of research on the Japanese monkeys of Koshima: Status and dominance. In T. Matsuzawa (Ed.), *Primate origins of human cognition and behavior* (pp. 405–417). Tokyo: Springer.
Wechkin, S., Masserman, J. H., & Terris, W., Jr. (1964). Shock to a conspecific as an aversive stimulus. *Psychonomic Science, 1*, 47–48.
Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., et al. (1999). Cultures in chimpanzees. *Nature, 399*(6737), 682–685.
Wrangham, R. W., McGrew, W. C., de Waal, F. B. M., & Heltne, P. G. (Eds.). (1994). *Chimpanzee cultures*. Cambridge, MA: Harvard University Press.
Yamagishi, T., Horita, Y., Takagishi, H., Shinada, M., Tanida, S., & Cook, K. S. (2009). The private rejection of unfair offers and emotional commitment. *Proceedings of the National Academy of Sciences, 106*(28), 11520–11523.
Yamamoto, S., & Takimoto, A. (2012). Empathy and fairness: Psychological mechanisms for eliciting and maintaining prosociality and cooperation in primates. *Social Justice Research, 25*(3), 233–255.

shackelf@oakland.edu

Helping Another in Distress: Lessons from Rats

Peggy Mason

> Every man must decide whether he will walk in the light of creative altruism or the darkness of destructive selfishness. This is the judgment. Life's most persistent and urgent question is, What are you doing for others?
>
> – MLK Jr

As a neurobiologist, I was somewhat surprised, but also thrilled, to receive an invitation to Oakland University's Evolution of Morality conference. I was surprised because I work with rats. Although I think very highly of rats, I also recognize reality. Morality is a stretch for members of the Rodentia order. It may even be a stretch for monkeys, but I will leave that question for Sarah Brosnan and other primatologists. On the other hand, my surprise was incomplete ("somewhat surprised") because for the past several years, my work has focused on pro-social behavior, also known in less stilted language as helping, among rats.

Why study helping in rats? The answer is simple. Rats don't have religion or culture; adult rats don't teach rat pups to follow the *Golden Rule*. Therefore, a finding that rats help is evidence that helping may be an evolutionarily conserved behavior. Simply put, if rats help another in distress, then maybe at least part of what motivates us humans to help is similar to what motivates rats to help. Maybe the *Golden Rule* is not our motivation for helping. Maybe we help because of a biological inheritance that we share with rats and other mammals. In this chapter, I will make the argument that affective communication drives pro-social behavior in mammals, including rats and humans alike.

I will not make the argument that pro-social behavior is the same as moral behavior. In fact, past this introductory section, I will not discuss morality. However, I will advance the idea that helping greases the wheels of social groups, facilitating social

P. Mason (✉)
Department of Neurobiology, The University of Chicago, 947 East 58th Street, MC0926, Chicago, IL 60637, USA
e-mail: pmason@uchicago.edu

© Springer International Publishing Switzerland 2016

T.K. Shackelford, R.D. Hansen (eds.), *The Evolution of Morality*, Evolutionary Psychology, DOI 10.1007/978-3-319-19671-8_9

shackelf@oakland.edu

cohesion. And, maybe, a socially cohesive group where individuals feel connected is a particularly fecund ground for good, dare I say moral, behavior. Therefore my contribution to the conference and to this volume is framed in terms of helping and its biological roots, specifically in rats.

Mammalian Babies Need Help to Survive

Mammalian young are born helpless and incompetent to survive on their own. They depend absolutely on their mother for milk, immune defense, thermoregulation, and protection from predation. In some species such as cat, the mother must tactilely stimulate the offspring to trigger the offspring's micturition. As failure to void is lethal, this function of the mother is critical to survival. In diverse ways, the mother's ability to react appropriately to her young's needs is the difference between passing on her genes and evolutionary death. Any failure of the mother to recognize the needs of her young and react for their benefit can result in the offspring's dying. Therefore, it is no surprise that mammalian mothers are supremely attuned to their babies' affective states. Mothers feed their hungry newborns and warm their cold babies. They extend their protection beyond themselves so that the odds are tilted in favor of the survival of their offspring.

The path to mammalian survival is difficult. The parent must recognize the distress of the offspring. And the offspring who needs help must signal that need unequivocally. Ergo, babies cry. Babies' cries are imperative expressions of need and difficult to interpret as anything but distress. Consequently, parents are moved to respond. Conversely, most parents do not go to help a baby who is not crying, accepting silence as evidence that all is well and that a nap may be possible.

Biology has left little to chance. By crying, babies get attention. In fact, the sound of a crying baby is so compelling that it is hard to ignore. Because our attention is so wholly captured by a crying baby, we may experience extended crying as aversive if we are incapable of helping. However, the source of the aversion is our inability to relieve the baby's distress rather than the baby. Under natural ecological conditions, the individual who hears a baby's cry would typically be a relative and would be moved to relieve the baby's distress. In contrast, baby's cries are aversive in modern situations where we are exposed to babies we do not know and cannot comfort.

Young mammals also depend on their mother for far more lofty reasons than simply survival. Harry Harlow's work in the 1950s and 1960s dramatically illustrates that primate offsprings' requirements extend beyond nutrition. Rhesus monkeys were separated from their mother and raised with two mother substitutes, one made of cloth and one of wire (Harlow, 1958). Even when the wire mother was the only source of milk, the young spent 10 times as much time with the cloth (nonnutritive) mother than with the wire (nutritive) mother. When the cloth mother was the milk source, the monkey babies spent no time with the wire mother. A mother, Harlow concluded, is required for much more than transferring milk to her newborns. Indeed, mammalian young raised with adequate food, warmth, and protection from predation but without social contact become fearful, anxious adults with impaired social skills (Harlow, Dodsworth, & Harlow, 1965; Kanari, Kikusui, Takeuchi, & Mori, 2005).

shackelf@oakland.edu

Whereas an empathic understanding by the mammalian mother of her young's condition is critical for survival and, in turn, for evolutionary success, the utility of a social bond between individuals extends into adulthood and encompasses more than just the mother-to-offspring relationship. After weaning, many if not most mammals continue to live social lives, hitching their own fates to those of others in their group. Darwin (1871/1998, p. 126) intuited that an individual should "extend his social instincts and sympathies to all the members of the same nation, though personally unknown to him." If we take "nation" to mean tribe or herd or social group, Darwin is clearly stating that social cohesion, borne of mutually directed feelings, facilitates survival of the group and its constituent individuals. In essence, sociality benefits individual adult survival through providing protection and increasing opportunities to feed, mate, and successfully raise offspring to reproductive age.

The Extended Self

Membership in a social group brings benefits that scale with the effectiveness of the group to work together. Thus, it is social cohesion, more so than simple sociality, which most powerfully promotes survival. William James considered that "*a man's Self is the sum total of all that he* CAN *call his*, not only his body and his psychic powers, but his … friends…" (James, 1890; emphasis in original). James's idea of an extended self depends on individuals reacting to the fortunes of others as they would if the same fortune or misfortune befell one's self. To the extent that our mood soars at a friend's triumph as it does upon our own triumph or our mood plummets in reaction to harm befalling a friend, that friend is part of a Jamesian extended self. Social cohesion increases as more and more group members consider more and more other group members to be part of their extended selves.

Social Cohesion

At the core of social cohesion among mammals is the communication of affective or emotional states between individuals. When individuals respond to others' emotional states as if they were their own, this bonds the individuals to each other, thereby building social cohesion within the larger group. The communication of affect or emotion between individuals is empathy. Defined in this way, empathy is an umbrella term that includes a large range of interactions in which an emotional or affective response is elicited by the emotional or affective state of another individual. Moreover, according to this definition, empathy is neutral in that responding to another's affective state, mood, or emotion does not constrain the actions taken, if any, as a result. We may hope that an individual reacts in a way that benefits another in distress but other outcomes are possible. For example, an understanding of another's emotions can be used to motivate an act of targeted cruelty that will exacerbate a victim's distress.

shackelf@oakland.edu

How Is Affective Communication Accomplished?

Perception and behavior are linked. They are linked automatically as occurs with motor mimicry. Many of us are familiar with the phenomenon of adopting the stance of a person or people we are talking with; soon after one member of a group crosses his or her arms, another person does the same. Simply viewing another individual's actions increases the probability that the viewer will perform the same actions (Chartrand & Bargh, 1999). This holds even if the individuals are strangers. Similarly, the fundamental speech frequencies of two conversing people are modified to more closely match each other (Gregory, 1990).

The link between perceiving another's actions and our own behavior is critical to affective communication between individuals. The perception–action link allows social adjustments to occur that make the actions of two interacting people more similar to each other and serves as an affiliative signal or a kind of social glue. Passing a person who cheerfully smiles at us makes us more likely to smile. We don't reason through this process; it just happens.

Actions Influence Emotions, Embodied Emotion

Actions are not only the readout of affect but actions also influence affect so that the interaction between emotions and outward expressions is two-way (Niedenthal, 2007). In other words, just as our emotions lead to actions, our motor actions are "re-experienced" as affect. Affect and emotions are expressed through both voluntary muscles producing posture, facial expression, breathing, and gaze and autonomic processes such as a rise in heart rate or perspiring, blanching, and blushing. The influence of facial expressions upon emotional experience is particularly strong in humans. People report emotions commensurate with artificially arranged facial expressions (Laird & Bresler, 1992). Feeling happy makes us smile and smiling can make us feel happy, or at least happier. When you're feeling good and laughing with friends, just try to feel angry or sad. As long as you keep your face in a smile or laugh, feeling an incongruent emotion is nearly impossible. Deriving emotion from action, often termed embodied emotion, is the essence of the Stanislavski system of method acting in which the affect that emerges from movements provides the emotive force of a performance.

The links between perception and action and between action and affect set up a cascade whereby perception by A of B's actions ultimately results in A feeling B's mood. This cascade results in an individual experiencing an affect that matches that of another (Preston & de Waal, 2002). The affect experienced by the viewer is vicarious in nature, "caught" from the other individual. The process by which one individual catches the affect or emotion of another is termed "emotional contagion" and is a fundamental building block of more complex forms of empathy.

shackelf@oakland.edu

Personal Distress Antagonizes Helping

Emotional contagion is required but not sufficient to elicit empathically motivated helping. In humans, personal distress must be suppressed to move from emotional contagion to helping, to choose action over immobility and panic. High levels of personal distress are detrimental to helping (Batson, Fultz, & Schoenrade, 1987). Suppressing personal distress allows an individual to focus on the other over the self and leads to empathic concern, an other-oriented emotional response elicited by and congruent with the welfare of an individual in distress. By requiring that the response is congruent with the welfare of the other, this definition precludes antisocial actions; the action taken by an individual feeling empathic concern is pro-social in nature.

By helping a distressed individual, a helper resolves not only the distressed individual's predicament but also their own uncomfortable affective state, providing an internal reward (de Waal, 2008). Thus, the distress of both the helper and the distress of the beneficiary of help are dissipated by helping. That the helper benefits does not diminish the pro-social action or its effect. In fact, it is precisely because "empathy gives individuals an emotional stake in the welfare of others" that the empathy-helping connection is so effective (de Waal, 2008, p. 292).

How Can We Study Empathy in Nonhuman Animals?

Empathy is an internal experience. We can understand something about the human experience through language. However, probing a nonhuman animal for the internal feeling of empathy is more challenging. One exciting way to study empathy in a variety of animals is to look for helping behavior.

Definition of Helping

Helping actions are those that benefit another. This definition does not say anything about the helper. It does not require that the helper intends to help or even that the helper is aware that he or she is helping. How can we move from this milquetoast definition to one that incorporates the actor's internal state or intentionality? Let's consider a concrete example. If I drop a 20-dollar bill out of my pocket and a stranger picks it up hours later, the stranger is helped. But how could we determine how I feel about the action? Well, if I push my cash deeper into my pocket on the next day, I have demonstrated that I would prefer to avoid repeating my helpful action. On the other hand, if I feel happy imagining someone else's joy at finding easy money, I may go out the next day and deliberately litter the street with bills for strangers. In the former scenario, I helped once and inadvertently, whereas in the latter scenario, I purposely and repeatedly helped. The two admittedly contrived

shackelf@oakland.edu

scenarios described above highlight the distinction between a one-time event and a repeated action. Such a distinction is critical as it offers a window into motivation. In addition, we can use this distinction to come to a working definition for helping. Helping is defined as an intentional action that benefits another.

The requirement that helping is intentional allows us to distinguish between inadvertent, regretted actions that benefit another and purposeful actions that have the same outcome. We call the latter helping but not the former. This definition has utility as it reflects a common sense reality. A single act of helping may not be meaningful. It could reflect affiliative feelings and intentionality but it could, just as easily, represent an accident. On the other hand, acting repeatedly is a reasonable sign of intentionality. Furthermore, requiring that helping is an intentional action is useful as we can use this criterion to discern whether nonhuman animals and nonverbal humans help.

Primates Help

Now that we have a working definition of helping, how do we investigate the biological basis of helping? The most common approach used has been to ask whether, given the opportunity, individuals will engage in pro-social behavior such as giving money (humans only), sharing food (primarily nonhuman primates), or lending a hand (nonhuman primates and young humans including preverbal humans). Of course, adult humans help others regularly. As an example, data from a worldwide Gallup poll reveal that nearly half of the world's population reports having "helped a stranger, or someone you didn't know who needed help" in the past month (Charities Aid Foundation, 2014, p. 4). The fact that humans help each other under many, albeit not all, circumstances is sufficiently obvious as to preclude further consideration of this issue.

At this point in time, the evidence is overwhelming that preverbal humans and nonhuman primates show pro-social behavior. For example, a useful paradigm to test for helping behavior is to place an object out of the reach of an experimenter but within reach of the test subject. The subject watches the experimenter try to reach for the object unsuccessfully. The question is whether the subject will hand the object to the experimenter even though no reward is gained by doing so. In one study using this paradigm, the proportions of children (60 %) and chimps (50 %) that handed the object to the experimenter at least once in 10 trials were similar (Warneken, Hare, Melis, Hanus, & Tomasello, 2007). Children and chimps that had helped were tested again but with physical obstacles placed between the subject and the object. Although the obstacles added to the cost of helping, over half of the chimps as well as over half of the young humans helped. Thus, helping is not restricted to humans but occurs in other primate species as well.

Primates are considered "special" enough by many that establishing a trait or ability in primates is not interpreted as evidence for the trait or ability being a mammalian one. In other words, if we find tool-making in monkeys, we don't

shackelf@oakland.edu

automatically generalize this finding to conclude that tool-making is found in mammals outside the primate order. In contrast, rodents belong to a distinct superorder from primates with the divergence between these two orders occurring 70–80 million years ago. Thus, the presence of a trait or ability in rodents may be interpreted as evidence that it is a general mammalian trait or ability. With this perspective in mind, let us look at the earliest experimental work testing whether helping occurs in rodents.

Early Work on Helping in Rodents

In 1959, Church designed an experiment to test whether rats are sensitive to the distress of other rats. He trained "test" rats to press a lever to obtain food. Three groups of rats then received (1) shock paired with a "demonstrator" rat also receiving shock, (2) shock alone, or (3) no shock (controls) for two days. For the final 10 days, test rats were allowed to press the lever for food while the demonstrator rat in the adjacent chamber received shock. Test rats who had not themselves experienced shock continued to press for food; they were apparently un-bothered by the shocking of the demonstrator rat. However, test rats who had experienced shock pressed the lever for food far less. Most importantly to Church, rats who received paired shocked (with a demonstrator rat) pressed the lever for food even less than did rats who had been shocked solo. Church interpreted this result as evidence that test rats felt "sympathy" for the demonstrator rats. Another interpretation is that the rats were scared into frozen immobility.

Church's work was pioneering but inconclusive. It was pioneering to conceive that one rat may communicate affectively with another. It was even more revolutionary to suggest that rats would forego food to save a compatriot from shock, an act of costly giving. Indeed Church's work marks an important beginning in social neuroscience. It is typically cited as the first suggestion that rats express something akin to empathy (see more on this below). However, because the evidence for caring or even help stemmed from *inaction*, the results could, as noted above, be explained by immobility associated with fear.

A few years after Church, Rice and Gainer (1962) reported an experiment that circumvented the inaction problem as rats were able to help another rat through *action*. Test rats were placed in a cage with another rat that was hoisted up with its "legs hang[ing] free" (p. 123). Rice and Gainer wanted the hoisted rat to exhibit distress and were not disappointed as the rat "typically squealed and wriggled satisfactorily while suspended, and if it did not, it was prodded with a sharp pencil until it exhibited signs of discomfort" (p. 123). Test rats, regardless of whether they were naïve or had been lever trained (through shock-avoidance training), pressed a lever which lowered the hoisted rat. Control rats, either naïve or trained, did not press the lever to lower a Styrofoam block. Rice and Gainer interpreted their results as evidence of altruism. While this experiment lacked controls such as an inactive lever, the result is strongly suggestive. Nonetheless, the paradigm is sufficiently unpalatable by modern ethical standards to preclude its use today.

shackelf@oakland.edu

Giving Rats a Chance to Help

Between the 1960s and 2011, there were reports of rodent cooperation, reciprocal altruism, and emotional contagion. Yet there was no clear report of rodent helping. Connected to this lack was the absence of a paradigm that would allow a mouse or rat to help if they were so inclined. Motivated by curiosity about the biological basis for helping (if any), Inbal Ben-Ami Bartal, at that time a graduate student in the laboratory of my colleague Jean Decety, fearlessly stepped into this void. Inbal, Jean, and I joined forces to test whether rats would express pro-social behavior given the rat-friendly opportunity to do so.

The kernel for the paradigm that we developed came from simple observations of two rats, one trapped in a Plexiglas tube (termed a restrainer) and one free. In the presence of a trapped rat, a free rat became quite active; biting, digging under, and climbing on top of the restrainer. Capitalizing on the free rat's apparent distress would require presenting the free rat with a reasonably sized challenge that would allow him to liberate the trapped rat. To accomplish this, a door was fashioned that was, like Goldilocks, neither too difficult nor too easy to open but which presented an achievable challenge to the rat.

We have dubbed the paradigm that we developed the *behavioral helping test* (Ben-Ami Bartal, Decety, & Mason, 2011). We use an arena with a floor that is 50 cm × 50 cm. A restrainer is placed in the center of the arena. The restrainer has one door and that door can only be opened from the outside and, therefore, can only be opened by the free rat. At the start of a session, one rat is trapped in the restrainer; this rat is referred to as the trapped rat. The second rat is free to roam in the arena and is termed the free rat. Rats are placed in the arena for an hour per day for 12 days.

Opening the door may seem easy to the human observer but is quite difficult for a rat. The first challenge is that a free rat must venture away from the walls of the arena to the center. Rats notoriously avoid open areas. It was readily apparent that free rats tested with a trapped rat entered the arena center and spent more time in the vicinity of the restrainer than did rats tested with either empty or object-filled restrainers (Ben-Ami Bartal et al., 2011). The second challenge for the free rat is to open the restrainer. Reared in an impoverished environment, laboratory rats do not appear to easily grasp the motor challenge. They do not appear to know which part of the restrainer contains an entry point. They, of course, do not grasp the implications of the counterweight. Numerous experiments have demonstrated that Sprague–Dawley rats require 3–6 days on average to learn to open the restrainer door. This holds regardless of whether a trapped rat or a food reward lies within.

In early experiments, rats would try to open the restrainer for 2 or 3 or even 4 days. However, eventually, many would stop trying, meaning that they stopped entering the arena center and would instead curl up in a corner and go to sleep. We interpreted this behavior as a sign of learned helplessness, born of the difficulty in opening the restrainer. To prevent rats from entering into a state of learned helplessness, the restrainer door was halfway opened at a point 40 min into the hour-long testing session. When the door was propped open by 45°, either rat could complete the opening.

shackelf@oakland.edu

Thus, no opening after the halfway opening point was considered a door-opening. In other words, the free rat's opportunity to help ended at the time of halfway opening. The halfway opening procedure was originally designed to give rats "hope" so that they would continue to engage with the task on all days of testing. This procedure works well—rats try, day after day, to liberate a trapped rat—and is always used.

Now that the paradigm is fully described, let us examine what happens. When two male Sprague–Dawley rats who are cagemates are tested, the free rats begin to consistently open the door by the third to sixth day of testing (Ben-Ami Bartal et al., 2011). Repeated door-opening is a strong sign that the rats are acting intentionally. As discussed above, a single beneficial act does not make helping. Notably, most rats open the restrainer door repeatedly, on consecutive days. A minority of rats opens the restrainer as many as three times or so but does so on nonconsecutive days. This distinction has led to the definition of an opener as a rat that opens the restrainer door on consecutive days at least three times. At minimum, this requires four door-openings on four consecutive days. In reality, most rats that become openers open the restrainer door on every testing day following the first door-opening.

A second sign of intentionality is that rats open the door at progressively shorter latencies. By the final session, most rats open the restrainer door within just a few minutes (Ben-Ami Bartal et al., 2011). Another sign of intentionality is that by the second door-opening, rats almost always use the same movement to open the restrainer door, a nudge up with the head. Finally, rats freeze at the sound of the door falling over when they first open the restrainer door. However, by the sixth day of testing, rats no longer freeze, evidence that they are no longer startled by the consequences of their actions.

Why Do Rats Open the Restrainer Door?

The finding that rats open the door of a restrainer containing a trapped cagemate is not sufficient to conclude that rats are helping. It is possible that rats simply like to open restrainer doors. However, data from three control conditions demonstrate that this is not the case. In one control, rats were tested with an empty restrainer for 12 days. Rats not only did not open but also spent very little time in the arena center. The same result was obtained with a second control condition in which rats were tested with a restrainer containing an object (an Ikea toy rat). These controls demonstrate that rats are not simply interested in accessing and exploring a new space.

A third control addressed the issue of social buffering (reviewed in Kikusui, Winslow, & Mori, 2006). Essentially, social buffering refers to the beneficial effects of being in the company of a conspecific vs being alone. These beneficial effects are numerous and include a decrease in stress responses and "better recovery from aversive experiences." Whether social buffering produces benefits or social isolation causes harm is unclear. In any case, to control for the number of rats in the arena, a control was performed with two free rats and an empty restrainer. The rats were apart as in the normal paradigm but in the "2+empty" control condition, a perforated

shackelf@oakland.edu

Plexiglas wall separated the two rats and neither rat was trapped. Instead, an empty restrainer was placed on the side of one of the rats. As in the other two control conditions, rats did not become openers. It should be noted that the influence of thigmotaxis (the preference to be in contact with physical objects such as walls and therefore to avoid open spaces) on the rats' behavior was greatly lessened in the 2+empty control condition since each rat was contained in only half an arena.

In sum, rats consistently open a restrainer for a trapped cagemate but do not open a restrainer that does not contain a rat. Thus, the trapped rat is critical to motivating the free rat's door-opening behavior.

Are Rats Socially Motivated to Liberate a Trapped Rat?

A reasonable motivation for door-opening is the opportunity to socially interact with the liberated rat. Even as adults, rats spend a great deal of time playing. Therefore, we asked whether a free rat that was prevented from playing with a liberated rat following door-opening would continue to open the restrainer door. The most straightforward approach to this question would be to place the restrainer door at a divider so that the trapped rat would be liberated into a separate and therefore inaccessible space. Unfortunately, we found that this arrangement precluded the free rat from using its favorite motion, nudging up the door from the front, to open the restrainer. Instead the free rat could only open the door by tipping it over from the side. In extensive preliminary experiments, we found no evidence that male Sprague–Dawley rats could learn this action. The rats spent a great deal of time around the restrainer containing a trapped rat but never acquired the ability to open the restrainer door from the side.

To get around the motoric limitations of laboratory rats, we first "trained" rats in the normal paradigm in which the restrainer containing the trapped rat was placed in the arena center. We then tested all openers in the "separated" apparatus described above. Rats that became openers in the normal paradigm were split into two groups with one group first being exposed to a trapped rat followed by an empty restrainer and the other group experiencing the reverse order. The results were clear. Rats opened for a trapped rat but not for an empty restrainer in the separated condition. Thus, door-opening does not require the opportunity for immediate social interaction with the trapped rat.

A recent report comes to a different conclusion, suggesting that rats in the helping behavior test are simply trying to access the trapped rat for social interaction (Silberberg et al., 2014). However, this report, which purported to replicate the helping behavior test, is extraordinarily difficult to interpret in large part because the paradigm used bears little to no resemblance to that reported in Ben-Ami Bartal et al. (2011). Among the modifications are (1) a touch of the door, of unspecified strength, triggered door-opening; (2) rats were repeatedly re-trapped after being released by the free rat's touch of the door; and (3) touch of the door sometimes opened the near door, sometimes the far door, and sometimes did not open either

shackelf@oakland.edu

door. Furthermore, no evidence is presented that the rats understood this dizzying array of contingencies. The most straightforward interpretation of these experiments is that because effective liberation of the trapped rat was not possible, rats opted for physical proximity.

As eloquently described by de Waal (2008), the proximate drivers of behavior should not be confused with the ultimate causes of evolutionary success. In other words, brain circuits don't calculate the plusses and minuses of every action, releasing only those behaviors that promise the greatest gain. Such bean-counting is the domain of ecological survival and evolutionary success. In this light, we consider the tautology that social behavior requires social reward. To put it another way, if rats did not benefit from social interactions, they would not have evolved to behave socially. The benefits of sociality drive the ultimate emergence of social behaviors in mammals and other animals. Yet, rats open the door for a trapped cagemate for proximal reasons and those reasons do not include immediate social interactions.

Who Will Rats Help?

The helping behavior test is the first tractable paradigm for studying pro-social behavior in a mammal outside the primate order and as such has already paid off. Recently, we used the test to identify the social requirements of helping (Ben-Ami Bartal, Rodgers, Bernardez Sarria, Decety, & Mason, 2014). To determine if rats would help a stranger, male Sprague–Dawley rats were tested with a different, unrelated male Sprague–Dawley rat each day. In other words, each free rat was tested with 12 different trapped rats, all of whom were individually unfamiliar. Somewhat unexpectedly, rats open for strangers just as they did for cagemates. This result is less puzzling when one recalls that laboratory rats went through a genetic bottleneck in the early twentieth century. The Sprague–Dawley stock started from breeding Wistar females to one male of mixed (and unknown) parentage in 1925 (Lindsay & Baker, 2006). Hundreds of generations later, individual Sprague–Dawley rats remain genetically related but not identical conspecifics.

Given the genetic relatedness of Sprague–Dawley rats, a better way to ask whether rats will help a stranger is to ask if rats will help a stranger from a different stock. [Stocks are commonly, although erroneously, referred to as "outbred strains" and are genetically heterogeneous although reproductively isolated populations.] Therefore, male Sprague–Dawley rats were tested with trapped male rats of the Long-Evans stock (Ben-Ami Bartal et al., 2014). The Long-Evans stock was founded in 1915 by mating several Wistar females with a wild male rat (Lindsay & Baker, 2006). A large number of breeding pairs (~250) and less than 1 % in-breeding have been used to separately maintain the Long-Evans and Sprague–Dawley rat stocks since their inception (Lindsay & Baker, 2006). In essence, Long-Evans and Sprague–Dawley rats are genetically isolated populations of the same species. Long-Evans and Sprague–Dawley rats appear different to human observers as the former have a black cape pattern on a white hair background, while the latter are albino (all white with pink eyes) rats.

shackelf@oakland.edu

We found that Sprague–Dawley rats help a Long-Evans cagemate but not a Long-Evans stranger (Ben-Ami Bartal et al., 2014). This result was not easily interpreted, as the Sprague–Dawley rat tested with a Long-Evans stranger was unfamiliar with both the individual trapped rat and the kind of rat in the restrainer. One possibility was that stock familiarity was required but individual familiarity was not. To test this idea, Sprague–Dawley rats were housed for 2 weeks with one Long-Evans rat, after which they were re-housed with each other. These Sprague–Dawley rats were then tested with Long-Evans strangers. Remarkably, Sprague–Dawley rats that previously lived with a single Long-Evans rat opened the restrainer door for Long-Evans strangers. These results show that helping is extended to rats of a familiar stock, regardless of whether the individual rat is known.

For rats, the target of helping is easily modified by experience. A mere two weeks of cohousing was enough to open up an entire stock of rats as potential recipients of pro-social behavior. If environment is so powerful, is there any role for genetics? Is genetic relatedness enough of a biological connection to support helping? This is a question that cannot be answered with human experimentation. Luckily the helping behavior test allowed us to use rats to empirically address the issue.

Sprague–Dawley rats were transferred on the day of birth to a Long-Evans litter (Ben-Ami Bartal et al., 2014). The fostered Sprague–Dawley rats grew up in a Long-Evans world, knowing only a Long-Evans mother and only Long-Evans siblings. Importantly, these fostered Sprague–Dawley rats never saw or interacted with another Sprague–Dawley rat. When the fostered rats were adults, they were tested with either Long-Evans or Sprague–Dawley strangers. Since the fostered Sprague–Dawley rats had spent their entire lives, not just two weeks, surrounded by Long-Evans rats, it was expected that the fostered Sprague–Dawley rats would help Long-Evans strangers. Indeed that is exactly what happened.

The big question then was: would Sprague–Dawley rats help Sprague–Dawley strangers? Amazingly, the answer was no (Ben-Ami Bartal et al., 2014). Sprague–Dawley rats that had never encountered another Sprague–Dawley rat did not help Sprague–Dawley strangers. These findings suggest that rats do not inherit genetic instructions to help others of their own kind. Instead, they learn which individuals to help from their social environment. Thus, when it comes to targeting helping, environmental experiences trump genetics.

Socially selective helping in rats may be mistakenly taken as evidence for a biological basis for social bias, stock-ism if you will. However, the results are more consistent with a biological basis for group-ism that is targeted to groups through social experience. Humans readily form strong affiliations to groups that are based on "minimal-group" criteria, such as an arbitrary assignment to one of two meaningless markers (e.g., red vs blue wrist bands; Tajfel, 1970). The finding that rats raised without socially interacting with others of their own stock do not help strangers of their own stock demonstrates that group affiliation, with respect to helping, is fluid, based on experience and not genetically determined. Jane Elliott's classic experiment (http://www.pbs.org/wgbh/pages/frontline/shows/divided/etc/view.html) demonstrates the same fluid group membership among school children and illustrates the biological power of social cohesion to drive discriminatory behavior, even when based on hollow characteristics such as eye color.

shackelf@oakland.edu

What Prevents Helping?

Not all humans show empathy or express helping behavior. Similarly, about 25 % of the rats that we have tested in standard conditions (i.e., with a trapped rat from a familiar stock) do not show helping behavior. The predominant reason for not helping appears to be an excessive amount of personal anxiety. Bonobo apes who show more anxiety, measured by how much they scratch themselves, and who take longer to recover from a stressful event show less consolation behavior toward other bonobos in distress (Clay & de Waal, 2013). This finding dovetails beautifully with human literature suggesting that to use empathy for helping or caring, an individual must overcome personal distress, a process typically termed self- or down-regulation. For example, people with a specific genetic variation who show greater social anxiety also demonstrate less helping behavior (Stoltenberg, Christ, & Carlo, 2013). Thus, rather than showing a lack of empathy, many individuals who do not help may be unable to suppress the anxiety associated with catching another's feeling of distress.

In professions that involve repeated exposure to human suffering, such as medicine, strong down-regulation is highly adaptive in counteracting the development of burnout. Indeed, physicians have a down-regulated response to noxious events that are common in medical practice. For example, an image of a needlestick evokes a lower assessment of pain intensity and unpleasantness by physicians than controls (Decety, Yang, & Cheng, 2010).

In light of the vulnerability of helping to personal distress, it is remarkable that rats often help another rat in distress. We would expect that through emotional contagion, the free rat would experience at least some of the distress expressed by the trapped rat. Yet rather than freeze, the typical reaction of a rat in distress, rats act intentionally to open the restrainer door (Ben-Ami Bartal et al., 2011). This suggests that the helper rat recognizes that his distress is vicarious in origin. In other words, the rat is able to attribute his personally felt distress to the trapped rat's condition and to distinguish that from his own condition. Such recognition, at least at the operational level, of the distinction between self and other is unexpected in a rat.

Multidimensional Space of Helping

There are limits to what rats and the helping behavior test can tell us about humans. To understand these limits, let us imagine that helping occurs within a multidimensional space. Axes within this space include the cost of helping, the severity of the need, and the familiarity of the individual in need of help. Helping may or may not occur depending on where a given situation sits within this space. For example, one's own child asking for money to see a movie is far more likely to be helped than if a stranger asks for money to see the same film. On the other hand, when the need for help is possibly life-threatening, the requirement for familiarity is tossed out the window. The news is replete with stories of heroic individuals who help an injured stranger in a time of acute crisis.

shackelf@oakland.edu

The helping behavior test sits within the multidimensional space of helping at a point that represents low cost, moderate need, and familiarity with "type" although not necessarily with the individual. There are also likely to be factors that influence the likelihood of helping in humans but not in rats. One such example comes from the influence of witnesses or bystanders. This topic was explored in a classic book, *The Unresponsive Bystander*: *Why Doesn't He Help*?, by Latané and Darley. Latané and Darley set out to understand what makes the modern bystander so apparently apathetic and callous, watching but not helping as others are hurt, maimed, and even die. They ask whether urbanization has created such extreme apathy and alienation that people no longer care about each other. The proximal motivation for their inquiries were two horrific crimes: the killing of teenager Andrew Mormille, witnessed by many but helped by none, on a New York City subway, and the 1964 stabbing, rape, and murder of Catherine "Kitty" Genovese.

Despite the fact that the version of the Genovese story repeated by Latané and Darley is now recognized to be apocryphal, there are plenty of anecdotes, then and now, of groups doing little or nothing to help others in distress. This apparent diminution of social affiliation evidenced by a seeming lack of concern for others is the starting point for Latané and Darley. Latané and Darley argue, reasonably so, that a change in the "basic motives underlying altruism" is unlikely. Evolution of a seismic shift in psychological motives over the last few hundred years due to the industrial revolution and the growth of urban living appears to them, and me, exceedingly unlikely. They therefore try to understand the psychology of helping through a series of experiments.

Latané and Darley begin with several innocuous forms of "help" that incur a very low cost. New Yorkers acquiesced to strangers' requests for the time (85 %) or directions (84 %). Yet, only a third responded when asked, "Excuse me, I wonder if you could tell me what your name is?". So far, so good; for low cost, non-dangerous, and nonintrusive requests, New Yorkers are helpful at a high rate. The finding that New Yorkers balk at intrusive requests appears reasonable if one considers that yielding one's privacy is, in and of itself, a cost.

Latané and Darley then conducted a series of experiments in which subjects (college students) went into a waiting room where they started to fill out a questionnaire. In all of the experiments, there was a cover story, a ruse if you will, for the experiment. For example, subjects might be told that the experimenters were interested in the students' reactions to New York. Then the subjects were asked to fill out preparatory forms prior to testing. The actual experiments took place during this time, when the subjects believed that they were simply waiting for the experimental interview to occur.

In one of the most compelling experiments, white smoke started to flow into the room and continued to do so for 4 min. Within this time period, 75 % of the subjects who were in the room alone (18/24) left the room to report the problem but only three of 24 (13 %) people in the room as part of eight triplets did so. The difference between 75 and 13 % is both large and significant. Furthermore, consider that if the three people in the triplets acted independently, then there would be only a 1.6 % chance that none of the triplet subjects would have reported the smoke.

shackelf@oakland.edu

However, 87 % of the triplets failed to report the smoke! Thus, the presence of two strangers decreases the likelihood of helping from an expected 98 to 13 %. Latané and Darley find the same effect—far more helping by single people than by individuals within groups—in a series of experiments employing a variety of other scenarios such as a beer being stolen from a liquor store, a man having a seizure, and so on.

When the in-house experiments (all but the liquor store experiment) were over, the experimenters always talked to the participants, asking them whether they "had experienced any difficulty while filling out the questionnaire." The responses in the smoke experiment were exemplary. Participants that reported the smoke said that they thought something was wrong, that there may be a fire, and that in any case, checking into the situation was warranted. In contrast, participants that did not report did not believe that the smoke was due to a fire. They had other explanations for the smoke. Some of these explanations were truly bizarre; not one but two subjects shared their thought that the smoke was a truth gas intended to make sure that they accurately filled out the questionnaire. In the other experiments, there were similar justifications for the strange goings-on, but only from those that did not act. Thus, it appears that the stories that reporters and non-reporters tell themselves were different. Could this be because reporters and non-reporters perceived the incidents differently?

A clue that reporters and non-reporters indeed processed identical situations differently came from the smoke experiment. In this experiment, it was easy to tell the moment when subjects noticed the smoke: they turned their head and looked up toward the vent. Of subjects tested alone, 63 % noticed the smoke within 5 s, whereas only 26 % of subjects tested in groups did so. Could the presence of others have put a damper on helping by changing the processing of olfactory/visual stimuli and thus olfactory/visual perception?

The idea that the perceptions of reporters and non-reporters could differ came into sharp relief in an experiment on children fighting. In this experiment, subjects (again college students) came in purportedly for market testing on games. When the subjects arrived, they were ushered into a room marked Student's Testing Center, which was separated by a heavy curtain from a room marked Children's Testing Center. Shortly after entering the room, the subject (who was alone) heard a rapidly escalating fight over a toy. [The fight was previously recorded and played on a tape recorder.] At one point, one boy said "…Get off me. You cut me. It hurts so much. Help me, please. I'm bleeding." The other boy's voice said, "You asked for it. I'm going to really beat you up." And so on.

Only one of 12 subjects reported the fight. But what is really interesting is that most (9/12) participants in this study explained afterwards that they did not believe that it was a fight! They thought that the boys were listening to a TV show or that it was a recording. These results suggested to Latané and Darley the possibility that subjects "were motivated to disbelieve the reality of the situation by their desires to avoid involvement." They then directly tested this idea. They set up a condition that was just the same except that subjects were told of the children, "Somebody is in there testing them…they do tend to get a little noisy but please leave them alone."

shackelf@oakland.edu

Thereby, *absolved of responsibility*, only one subject (1/8) failed to believe that the fight was real. The difference between 75 % (responsibility) and 12 % (no responsibility) of the subjects perceiving the situation as faked was significant. This is strong evidence that expectations of responsibility can "distort" perceptions.

There was only one situation—the one involving a person having a seizure—in which the non-reporters consistently believed that there was a genuine emergency and were distressed. In other words, the non-reporters perceived the situation in the same way as did reporters. The non-reporters said into microphones (that they believed were turned off), "Oh, what should I do?," and asked about the health of the seizing confederate immediately when the experimenters came in to terminate the session. Latané and Darley conclude that rather than having decided not to help, these non-responders simply had not figured out how to help.

In the end, Latané and Darley suggested four ways in which the presence and appearance of others suppress helping by an individual in an emergency situation:

- By acting, we risk looking foolish in front of others. This may not be a conscious influence. In the bullying experiment, the burden of responsibility led to internal pressure to interpret the situation as not requiring intervention.
- Others' inactivity guides us to act similarly. Whereas others may only be inactive transiently, en route to figuring out what to do, the influence of this inactivity reinforces collective inaction.
- The effect of the above two influences is additive so that an individual's perception of situations is heavily influenced by the apparent lack of reaction shown by others. In other words, *perception itself is altered by the social circumstances*, enabling a bystander to interpret a situation as less serious than it is and lessening the pressure to act.
- The presence of others diffuses an individual's sense of responsibility, making it more acceptable to not act because it is perceived that others may act.

Is there a particular personality that is associated with increased likelihood of helping? Or are we are all capable of watching others in distress without helping just as we are all capable of helping in other situations? What are the factors that influence us, all of us, to help or not help in specific circumstances? Such factors surely exist although, for example, subjects in the studies of Latané and Darley consistently believed that the presence of others had no influence on their actions. Latané and Darley speculate that if we all had an understanding of the social influences on helping, we may be able to counteract our natural tendencies and act more pro-socially, even in crowds.

Helping Is Mammalian

Despite their enormous differences, humans and rats share commonalities inherited through a shared phylogenetic history. As mammals, both species are capable of detecting and responding helpfully to the distress of others. Humans often do this

shackelf@oakland.edu

using language but rats use actions that move physical objects in the world. Although popular culture suggests that a chasm separates saying and doing, this idea is not supported by neurobiology. The muscles of speech—laryngeal, abdominal, and upper airway muscles—are skeletal muscles that require input from motoneurons just as the biceps, quadriceps, pectoral muscles are. In essence, speaking (and writing) is no more and no less an action than lifting a weight or opening a door. There is nothing more accurate or honest about an action using laryngeal muscles than one using limb muscles. With this motor equivalence in mind, I would argue that rats' actions speak as loudly as ours do. Therefore, we would do well to listen to the lessons that rats can teach us.

References

Batson, C. D., Fultz, J., & Schoenrade, P. A. (1987). Distress and empathy: Two qualitatively distinct vicarious emotions with different motivational consequences. *Journal of Personality, 55*, 19–39.

Ben-Ami Bartal, I., Decety, J., & Mason, P. (2011). Empathy and pro-social behavior in rats. *Science, 334*, 1427–1430.

Ben-Ami Bartal, I., Rodgers, D. A., Bernardez Sarria, M. S., Decety, J., & Mason, P. (2014). Pro-social behavior in rats is modulated by social experience. *eLife, 3*, e01385.

Charities Aid Foundation. (2014). *World giving index 2014*. http://www.cafamerica.org/media/wgi-2014/

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology, 76*, 893–910.

Church, R. M. (1959). Emotional reactions of rats to the pain of others. *Journal of Comparative and Physiological Psychology, 52*(2), 132–134.

Clay, Z., & de Waal, F. B. (2013). Development of socio-emotional competence in bonobos. *Proceedings of the National Academy of Sciences, 110*, 18121–18126.

Darwin, C. (1871/1998). *The descent of man*. Amherst, NY: Prometheus Books.

de Waal, F. B. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology, 59*, 279–300.

Decety, J., Yang, C. Y., & Cheng, Y. (2010). Physicians down-regulate their pain empathy response: An event-related brain potential study. *NeuroImage, 50*, 1676–1682.

Gregory, S. W. (1990). Analysis of fundamental frequency reveals covariation in interview partners' speech. *Journal of Nonverbal Behavior, 14*, 237–251.

Harlow, H. F. (1958). The nature of love. *American Psychologist, 13*, 673–685.

Harlow, H. F., Dodsworth, R. O., & Harlow, M. K. (1965). Total social isolation in monkeys. *Proceedings of the National Academy of Sciences, 54*, 90–97.

James, W. (1890). *The principles of psychology*. http://psychclassics.asu.edu/James/Principles/prin10.htm

Kanari, K., Kikusui, T., Takeuchi, Y., & Mori, Y. (2005). Multidimensional structure of anxiety-related behavior in early-weaned rats. *Behavioural Brain Research, 156*, 45–52.

Kikusui, T., Winslow, J. T., & Mori, Y. (2006). Social buffering: Relief from stress and anxiety. *Philosophical Transactions of the Royal Society B: Biological Sciences, 361*, 2215–2228.

Laird, J. D., & Bresler, C. (1992). The process of emotional experience: A self-perception theory. In M. S. Clark (Ed.), *Emotion. Review of personality and social psychology*, No. 13 (pp. 213–234). Thousand Oaks, CA: Sage Publications, Inc.

Latané, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?* Upper Saddle River, NJ: Prentice-Hall.

shackelf@oakland.edu

Lindsay, J. R., & Baker, H. J. (2006). Historical foundations. In M. A. Suckow, S. H. Weisbroth, & C. L. Franklin (Eds.), *The laboratory rat* (pp. 1–52). Amsterdam: Elsevier.

Niedenthal, P. M. (2007). Embodying emotion. *Science, 316*, 1002–1005.

Preston, S. D., & de Waal, F. B. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25*, 1–20.

Rice, G. E., & Gainer, P. (1962). "Altruism" in the albino rat. *Journal of Comparative and Physiological Psychology, 55*, 123–125.

Silberberg, A., Allouch, C., Sandfort, S., Kearns, D., Karpel, H., & Slotnick, B. (2014). Desire for social contact, not empathy, may explain "rescue" behavior in rats. *Animal Cognition, 17*, 609–618.

Stoltenberg, S. F., Christ, C. C., & Carlo, G. (2013). Afraid to help: Social anxiety partially mediates the association between 5-HTTLPR triallelic genotype and prosocial behavior. *Social Neuroscience, 8*, 400–406.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American, 223*, 96–102.

Warneken, F., Hare, B., Melis, A. P., Hanus, D., & Tomasello, M. (2007). Spontaneous altruism by chimpanzees and young children. *PLoS Biology, 5*, e184.

shackelf@oakland.edu

Part IV
Religious Beliefs and Behavior

shackelf@oakland.edu

A Moral Guide to Depravity: Religiously Motivated Violence and Sexual Selection

Yael Sela, Todd K. Shackelford, and James R. Liddle

Many people around the world consider religion and morality to be inseparably linked. A recent study by the Pew Research Center (2014) found that in 22 out of 39 countries surveyed, a large majority of respondents believed that being moral and having good values requires belief in God. This view was less prevalent in richer nations, but was still held by 53 % of US respondents. This perceived link between religion and morality is problematic, particularly because of the types of behaviors that can be considered "morally right" based on one's religious beliefs and one's interpretation of religious texts. Although religion can motivate charity, cooperation, and forgiveness, it is vital to acknowledge that religion also motivates violence. This chapter explores the link between religion and violence, with the goal of better understanding this link through an evolutionary psychological perspective.

Human violence research has benefited tremendously from an evolutionary psychological perspective (e.g., Buss, 2005; Kaighobadi, Shackelford, & Goetz, 2009; Shackelford & Weekes-Shackelford, 2012). The most general contribution of this research is the understanding that violence is the behavioral output of evolved psychological mechanisms designed to solve particular adaptive problems and therefore occurs in predictable contexts (Buss & Shackelford, 1997b). This perspective has informed explanations of violence throughout the animal kingdom (Newton-Fisher & Thompson, 2012), and it is reasonable to argue that many cases of violence in humans have an underlying evolutionary function. We now understand some of

Y. Sela (✉)
Department of Psychology, Oakland University,
130 Pryale Hall, Rochester, MI 48309-4401, USA
e-mail: ysela@oakland.edu

T.K. Shackelford
Oakland University, Rochester, MI, USA

J.R. Liddle
Florida Atlantic University, Boca Raton, FL, USA

© Springer International Publishing Switzerland 2016

T.K. Shackelford, R.D. Hansen (eds.), *The Evolution of Morality*,
Evolutionary Psychology, DOI 10.1007/978-3-319-19671-8_10

shackelf@oakland.edu

the contexts in which violent behavior is most likely to occur, as well as the function it may have served over our evolutionary history (Liddle, Shackelford, & Weekes-Shackelford, 2012). Might there be circumstances that exacerbate violence within these contexts?

In this chapter, we argue that the short answer is "yes" and that the longer answer implicates religious beliefs in motivating and exacerbating violence, arguably by being shaped by preexisting, violence-promoting mechanisms in evolved psychology. In other words, we expect religiously motivated violence to occur in contexts that would elicit violent responses as revealed by past research, especially research from an evolutionary perspective (e.g., in the context of sexual selection). At best, the rates and degrees of violence should be equal in religious and nonreligious circumstances. However, we argue that it is more likely that religion makes it worse. That is, religious beliefs, practices, and norms may lower the threshold for violence, as well as explicitly promote and conveniently justify violent actions in evolutionarily relevant contexts.

Evolutionary Psychological Perspectives on Violence

Human violence is not a novel phenomenon. Evidence indicates that violence has been a part of human life throughout our evolutionary history. Archeological evidence of violent deaths among early humans, such as indicators of clubbing (Ferguson, 1997), arrowheads and barbs (Smith, 2007), and mass killings via blows to the head (Keeley, 1996), suggest that humans have experienced violence throughout our evolutionary history. Further evidence of our violent history comes from specialized tools whose function was not just hunting for food but also inflicting violence on people (Smith, 2007), the writings of many societies worldwide that describe violence (Keeley, 1996), and costly fortification structures that humans erected to defend themselves once they transitioned from nomadicity to permanent settlements (Smith, 2007). Given our history of violence, it is reasonable to suspect an evolutionary rationale behind such behavior.

A substantial amount of violence occurs within species, and an evolutionary perspective can help explain why this occurs. Much of this violence, and the conditions in which it occurs, can be explained with reference to two evolutionary theoretical models: sexual selection theory and parental investment theory (reviewed in Liddle et al., 2012).

Sexual Selection

Sexual selection (Darwin, 1871) is an important element of natural selection (Sela & Shackelford, 2015). Sexual selection theory is concerned with "the advantages that certain individuals have over others of the same sex and species, in exclusive

shackelf@oakland.edu

relation to reproduction" (Darwin, 1871, p. 256). Sexual selection is directly related to the reproductive success of individuals within a species, arising from sexual competition among individuals for access to mates and resulting in traits such as the antlers of stags, the horns of antelopes, and the tail of peacocks (*Pavo cristatus*). Although they facilitate reproductive success, such traits are costly, often reduce survival prospects, and can only be maintained by sexual selection.

Competition for reproductive success is realized in intersexual selection and intrasexual selection. Intrasexual selection refers to competition between individuals of the same sex for mating opportunities with members of the other sex. Intrasexual selection has resulted in the evolution of weaponry such as antlers and horns and—in humans—increased male body size, along with sex-specific patterns of aggression (Daly & Wilson, 1988), jealousy (Buss, Larsen, Westen, & Semmelroth, 1992; Easton & Shackelford, 2009), and many other traits (see Schmitt, 2003). Intersexual selection refers to the process by which the mate preferences of one sex (primarily females, who are typically the choosier sex) influence the selection of traits in the opposite sex. Intersexual selection has resulted in the evolution of traits such as the peacock's elaborate tail, vibrant coloration in birds and fish, and many kinds of bird vocalizations (Andersson, 1994) and courting behaviors (Coleman, Patricelli, & Borgia, 2004). As we discuss in this chapter, both of these components of sexual selection can help explain much of the violence that occurs within species, and specifically they can help explain religiously motivated violence in humans.

Parental Investment

Sexual selection alone provides an incomplete understanding of violent behavior. For example, in most species, males are more likely to engage in violent intrasexual competition than are females. To explain this robust sex difference, we must refer to parental investment theory (Trivers, 1972). Parental investment theory refers to the allocation of resources to offspring at the expense of other potential resource allocations (e.g., survival, mating effort, additional offspring). Essential to parental investment theory is the (often) large difference in the minimum obligatory investment that males and females provide their offspring. This differential investment can explain sex differences in regard to reproductive strategies, such as engaging in risky competition (e.g., physical violence) for mates. Parental investment theory predicts that the sex that exhibits greater obligatory investment (typically females, due to internal fertilization, gestation, nursing, etc.) will tend to avoid violence because of the costs associated with injury and death (i.e., being unable to care for current or future offspring). Additionally, the more investing sex should be choosier when selecting a mate because of the greater costs associated with making a poor selection, such as potentially investing valuable and finite resources into offspring with a low probability of survival and/or future reproductive success due to lower-quality genes. Meanwhile, the less investing sex (typically males, whose minimum investment may be a contribution of sperm) will tend to pursue risky competitive

shackelf@oakland.edu

behavior because of the benefits of successfully competing with same-sex rivals (i.e., access to the choosier sex for mating).

This risky competition often involves violence, most likely because of how effective violence can be at overcoming one's rivals relative to other forms of competition. Parental investment theory is supported by the fact that the males of most animal species, particularly mammals, are overwhelmingly more prone to violence than are females (Ghiglieri, 1999). Additionally, the theory's predictive power is particularly evident when examining certain species in which the discrepancy in minimum obligatory investment is reversed between the sexes. Examples of such sex-role-reversed species include Australian cassowaries (Ghiglieri, 1999), Mormon crickets, pipefish sea horses, and Panamanian poison arrow frogs (Trivers, 1985). In these species, males invest more than females in their offspring and—as parental investment theory predicts—males are choosier when selecting mates, whereas females compete with each other, often violently, for sexual access to males.

Sexual Selection and Parental Investment Explain Violence

Liddle et al. (2012) review and apply the theories of sexual selection and parental investment to violence. They discuss evidence to suggest that violence among nonhuman animals is not arbitrary or pointless, but instead appears to be determined by unconscious calculations of costs and benefits, and that these calculations also drive the majority of human violence. In other words, evolved psychological mechanisms for violence exist in human and nonhuman animals, and these mechanisms generate violent behavior in response to specific inputs indicating that—on average, over time—the potential costs of violence are outweighed by the potential benefits. Liddle et al. (2012) describe what some of these inputs may be and how these inputs determine the likelihood of engaging in violent behavior.

Paternal care in humans surpasses that of many other species, but women still exhibit greater obligatory investment. Women are therefore, on average, the choosier sex, and female mate preferences are a limiting factor for male reproductive success. Sexual dimorphism [e.g., men are heavier (Ghiglieri, 1999) and taller (Holden & Mace, 1999) than women] suggests a history of effective polygyny in which reproductive variability has been greater in men than women (Buss & Shackelford, 1997). Lower paternal care and increased reproductive variability in men (vs. women) lead to fiercer and riskier intrasexual competition for mates (Archer, 2009; Campbell, 2005). Indeed, around the world, men are much more likely to be the perpetrators, and targets, of violence (Archer, 2004, 2009; Burbank, 1992; Buss, 2005; Daly & Wilson, 1988, Ghiglieri, 1999; Lester, 1991).

Both intrasexual competition and intersexual selection help explain how violence translates into mating opportunities (Archer, 2009; Daly & Wilson, 1988; Liddle et al., 2012). Male intrasexual competition is often directly related to status or reputation. Women prefer partners of high status, most likely because status

shackelf@oakland.edu

serves as an honest signal of a man's ability (but not necessarily intent) to provide for a woman and her offspring. High status has historically translated into having access to more food, better territory, and greater social support (Buss, 2005). Men who perpetrate violence against other men have higher social status and greater sexual access to women, a phenomenon that has been documented among many tribal societies (Chagnon, 1988, 1992; 1992; Daly & Wilson, 1988) and in the United States (Buss, 2011; Campbell, 1993; Ghiglieri, 1999; Palmer & Tilly, 1995).

Intersexual selection and parental investment theory help explain the circumstances in which men perpetrate violence against women (Goetz, Shackelford, Starratt, & McKibbin, 2008). One consequence of substantial paternal investment in humans, relative to most mammalian species, is the increased fitness cost of being victimized by cuckoldry (i.e., unwitting investment in genetically unrelated offspring produced by a partner's infidelity). Cuckoldry has posed a serious adaptive problem for men throughout human evolutionary history; not only does a cuckold's investment in unrelated offspring benefit his rival, but also his own fitness suffers due to the diversion of his finite resources away from his current or future genetic offspring. Contemporary estimates of worldwide cuckoldry (ranging from 1 to 30 %; Goetz, Shackelford, Platek, Starratt, & McKibbin, 2007) suggest that cuckoldry occurred often enough throughout our evolutionary history to have imposed strong selection pressures on men. Therefore, we expect male psychology to have been shaped by this adaptive problem, resulting in psychological mechanisms for detecting and preventing cuckoldry. When an adaptive problem carries such extraordinary costs, violence may emerge as a viable solution in relevant contexts. Next, we discuss religiously motivated violence in the mating context.

Sexual Selection, Parental Investment, and Religiously Motivated Violence

The theories of sexual selection and parental investment allow us to explain the evolutionary rationale behind many acts of human violence, as well as the contexts in which such acts are most likely to occur. It is worth noting that, as with any evolutionary explanation for particular behaviors, it is unnecessary and indeed unlikely that individuals be aware of the evolutionary rationale behind their actions. If psychological mechanisms that promote violent behavior in response to particular environmental inputs (e.g., threats to one's status, indications of partner infidelity) historically lead to greater fitness, on average, then these mechanisms will be selected for, regardless if individuals understand the fitness-enhancing effects of such mechanisms. With this in mind, we now turn to an examination of specific religious beliefs and how they may have been shaped by preexisting psychological mechanisms designed to promote context-dependent violence. These beliefs, in turn, may serve to facilitate the functions of violence-promoting psychological mechanisms.

shackelf@oakland.edu

Wife Beating, Killing, and Raping

Men are much more violent than women due to lesser obligatory parental investment and greater reproductive variability. Male violence is not reserved only for male rivals—it is also directed at women and especially romantic partners: In 2001, 20 % of reported incidents of nonfatal violence against women age 12 or older were perpetrated by an intimate partner (Bureau of Justice Statistics, 2003), and between 1976 and 2005, 30 % of female homicide ("femicide") victims were killed by an intimate partner, making it the largest class of victim-offender relationship (Bureau of Justice Statistics, 2007). Previous research has identified contexts in which men are more likely to be violent, such as perceived risk of partner infidelity (reviewed in Kaighobadi et al., 2009). Accordingly, we expect men's religiously motivated, partner-directed violence to occur under similar conditions and the religious justifications for doing so to be explicit and situation specific.

Mate Guarding and Controlling Behaviors Mate retention behaviors can range from checking up on a partner to physical violence (Buss, 1988; Buss & Shackelford, 1997a). Although both men and women perform mate retention behaviors, men use more violent behaviors than women when guarding their partner (Buss, 2005; Daly & Wilson, 1988). If partner-directed violence does not deter partner infidelity or defection, a man may kill his partner as a last resort to prevent other men from gaining sexual access to her (Buss, 2000, 2005). Mate-killing also can repair a man's reputation. In many cultures (e.g., cultures of honor; Nisbett & Cohen, 1996), cuckolded men are viewed as emasculated, and killing an unfaithful partner may be the only way a man can repair his reputation (Buss, 2005; Daly & Wilson, 1988). Men direct violence at their partner, ultimately, to avoid the devastating consequences of cuckoldry. Male sexual jealousy is often the proximate mechanism motivating men's partner-directed violence. Indeed, it is among the most frequently cited causes of men's partner-directed violence, both physical and sexual (e.g., Buss, 2000; Daly & Wilson, 1988; Daly, Wilson, & Weghorst, 1982; Dobash & Dobash, 1979; Dutton, 1998; Frieze, 1983; Gage & Hutchinson, 2006; Russell, 1982; Walker, 1979). Therefore, we expect partner-directed violence (fueled by male sexual jealousy) to be sanctioned by certain religions, especially in the contexts of sexual conflict (e.g., real or perceived infidelity, offspring from previous relationships, a wife's refusal to have sex with her husband). Do religious texts promote male partner-directed violence in these contexts?

- Leviticus 20:10–12[1]: "And the man that committeth adultery with another man's wife, even he that committeth adultery with his neighbour's wife, the adulterer and the adulteress shall surely be put to death. And the man that lieth with his father's wife hath uncovered his father's nakedness: both of them shall surely be put to death; their blood shall be upon them. And if a man lie with his daughter in law, both of them shall surely be put to death: they have wrought confusion; their blood shall be upon them. Both parties in adultery shall be executed."

[1] All biblical quotations are from the King James Version.

shackelf@oakland.edu

- Deuteronomy 22:22: "If a man be found lying with a woman married to a husband, then they shall both of them die."
- Matthew 5:31–32: "It hath been said, Whosoever shall put away his wife, let him give her a writing of divorcement. But I say unto you, That whosoever shall put away his wife, saving for the cause of fornication, causeth her to commit adultery: and whosoever shall marry her that is divorced committeth adultery."
- Quran 4:15[2]: "[T]hose of your women who commit illegal sexual intercourse, take the evidence of four witnesses from amongst you against them; and if they testify, confine them (i.e., women) to houses until death comes to them or Allah ordains for them some (other) way."
- Quran 4:34: "[T]he righteous women are devoutly obedient (to Allah and to their husbands), and guard in the husband's absence what Allah orders them to guard (e.g. their chastity, their husband's property). As to those women on whose part you see ill-conduct, admonish them (first), (next), refuse to share their beds, (and last) beat them (lightly, if it is useful)."

Religious texts such as the Bible and Quran instruct men to stone, burn, torture, and poison women who are suspected of extramarital sex. Men are instructed to punish wives that commit infidelity (real or suspected) and exhibit—vaguely defined—promiscuous behavior [e.g., "play the harlot" or "commit whoredome," Ezekiel 23:1–49; If they do not "lower their gaze (from looking at forbidden things), and protect their private parts (from illegal sexual acts)… draw their veils all over *Juyubihinna* (i.e., their bodies, faces, necks and bosoms)," Quran 24:31] such that both conviction of the crime and execution of the punishment is at men's discretion (e.g., Quran 24:6). Some current manifestations of these practices are brutal beatings and killing of wives for showing skin (e.g., a husband recurrently beat his wife and finally killed her with a knife and cut her into pieces after she refused to cover her face with a veil outside the house, saying this was the best way to "punish [her] for rebelling against Allah's orders"; Jafri, 2013b), leaving the house without permission (e.g., a husband axed his wife after she "insulted him" by staying out overnight. He asked her to loudly recite the Kalimas—Islamic texts—before killing her and told the police he killed his wife to make her "a lesson for other women who do not obey their spouses"; Anonymous, 2013).

Religious codes of conduct are also enforced on women by designated groups of men (convenient for the husbands while they are away from their wives). For example, every year "chastity squads" of the morality police unit in Iran forcefully arrest and fine thousands of citizens, especially women and adolescents, for not following the Islamic dress code (e.g., Cohen, 2011). Further, extramarital sex is a public offence in Iran, punishable by stoning to death (Razavi, 2006; public stoning videos of alleged adulteresses are available online[3,4]). Another example is the Jewish modesty patrol in ultra-Orthodox neighborhoods in Israel. The modesty

[2] All quotes from the Quran are taken from the Hilali-Khan translation, available online from http://muttaqun.com/files/PDF/The_Holy_Quran_English_Arabic.pdf.

[3] Stoning to death of a couple in Afghanistan—http://www.youtube.com/watch?v=jJkYJ3cbxh0.

[4] Afghan woman executed for adultery—http://www.youtube.com/watch?v=9ed0TZN2Egk.

shackelf@oakland.edu

patrol beat, spit, and hurl stones at women and girls who wear clothing deemed provocative or that have allegedly consorted with men other than their husband (Associated Press, 2008).

It is socially acceptable and even expected for men to be overt in their violent mate guarding behaviors. In this way, some men gain an advantage in intersexual competition (by more effectively controlling their partner and preventing her infidelity). Additionally, men may gain an advantage in intrasexual competition by being allowed to retaliate more harshly (and with fewer reputational costs) against male rivals than would otherwise be permissible (e.g., "Both parties in adultery shall be executed"; Leviticus 20:10–14). By endorsing partner-directed violence, these religious texts offer men a socially acceptable solution to the adaptive problem of cuckoldry (by way of increasing paternity assurance). Men's religiously motivated, partner-directed violence also reduces male–male conflict over women because religiously prescribed violence limits women's behavior.

A husband's right to confine, beat, torture, and murder his wife for her infidelity is just one set of examples of religiously motivated violence that supports an unfortunate suite of men's evolved psychology. Another example is genital mutilation, to which we turn next.

Genital Mutilation

Male and female genital mutilation (MGM and FGM, respectively) is any permanent modification of the external genitalia that involves the ablation of tissue. MGM includes superincision (longitudinal bisection of the dorsal skin), circumcision (removal of the entire foreskin), and castration (removal or crushing of one or both testes or the penis). The most common MGM today is male circumcision, which is mandated in the Bible (e.g., Genesis 17:10–14, Exodus 12:48, Josh 5:2) and instructed by Islamic hadiths[5] and fatwas[6] (e.g., Al-Munajjid, n.d.a, n.d.b) and is therefore practiced by Jews, Muslims, and Christians.

FGM includes hoodectomy (removing the clitoral hood), vaginal infibulation (narrowing of the vaginal opening by cutting and repositioning the inner or outer labia), excision (partial or total removal of the clitoris and the labia minora, with or without excision of the labia majora), and clitoridectomy (partial or total removal of the clitoris) (WHO, 2013). FGM is prevalent worldwide, especially in Africa. FGM

[5] The hadith is a record of traditions or sayings of the Prophet Muhammad and his companions. This collective body of Islamic traditions is revered and received as a major source of religious law and moral guidance, second only to the Quran (Hadith. 2013. In Britanica.com. Retrieved September 1, 2013, from http://www.britannica.com/EBchecked/topic/251132/Hadith).

[6] A fatwa is a formal legal opinion given by an Islamic legal authority (mufti) in answer to an inquiry by a private individual or judge (Fatwa. 2013. In Britanica.com. Retrieved September 1, 2013, from http://www.britannica.com/EBchecked/topic/202671/fatwa).

shackelf@oakland.edu

is permissible, and some say encouraged, in the following hadiths: "Circumcision is Sunnah for men and an honorable thing for women" (Musnad Ahmad); "Cut off only the foreskin (outer fold of skin over the clitoris; the prepuce) but do not cut off deeply (i.e., the clitoris itself), for this is brighter for the face (of the girl) and more favorable with the husband" (Mùjam al-Tabarânî al-Awsat[7]); "A woman used to perform circumcision in Medina. The Prophet (peace be upon him) said to her: Do not cut severely as that is better for a woman and more desirable for a husband" (Book 41, Number 5251[8]). Religious authorities in countries where female circumcision is common encourage the practice by conveying to their (often illiterate female) audience that female circumcision is a religious requirement (Slack, 1988).

Despite the variability in genital mutilation types, each presents health risks (e.g., bleeding and infections) and reproductive costs (e.g., infertility). Further, it is unlikely that damaging or removing mechanically, neurally, and endocrinologically specialized, healthy sexual tissue is neutral with respect to its evolved function (Wilson, 2008). If genital mutilation impairs success in sperm competition (e.g., by affecting sperm delivery; see Wilson, 2008, c.f. Morris & Krieger, 2013), which is pivotal in sexual conflict, we expect an associated fitness benefit somewhere else. It might advantage men to reduce the functionality of their sperm competition adaptations (e.g., penis anatomy, ejaculate adjustment by way of arousal, and even reduced female partner satisfaction) if men thereby secured some other benefits that would help them later and without which they would be worse off.

Wilson (2008) argues that there is an underlying and common function to genital mutilation: providing hard-to-fake signals of compliance with the social assignment of reproduction. That is, conflict over paternity may favor men who invest preferentially in spouses with FGM and cooperate preferentially with peers who submit to MGM (Wilson, 2008). Wilson generated and tested several predictions from this functional, sexual conflict hypothesis of male and female genital mutilation. He found that MGM was associated with polygynous societies at high risk for extramarital sex and that MGM appeared to reduce this risk. He also found that MGM is performed by a nonrelative in public view of other men and that a genitally mutilated man gains access to social and sexual privileges that outweigh the costs of genital mutilation. Wilson's findings suggest that genital mutilations may impair the evolved capacity for extra-pair fertilizations, thereby decreasing paternity uncertainty and reproductive conflict, and that the benefits of trust and social investment from powerful married men outweigh the costs of mutilation.

[7] Al-Tabarani, quoted in Al-Albani, Muhammad N., Silsilat Al-Ahadeeth Al-Sahihah, A1 Maktab Al-Islami, Beirut, Lebanon, 1983, vol. 2, Hadeeth no. 722, pp. 356–357. See also Keller, N. H. M. trans.(1997). Reliance of the Traveller A Classic Manual of Islamic Sacred law by Ibn Naqib Misri. e4.3.

[8] http://www.usc.edu/org/cmje/religious-texts/hadith/abudawud/041-sat.php.

shackelf@oakland.edu

Honor Killing

Honor killing is the murder of a family member (usually female) by one or more other members (usually male), because they have brought shame to the family by "straying from the righteous path of God." Women have been killed for refusing to enter a prearranged marriage (e.g., Jafri, 2013a; Mirza, 2008), committing adultery (actual or alleged), being in a relationship that displeased their relatives (e.g., Spolar, 2005), or being raped (often by another family member; e.g., Mirza, 2008). This male bias in killers is consistent with men being more violent than women in general and with fathers being more violent than mothers due to paternity uncertainty. The methods of honor killing often reflect "overkill" (using excessive methods, more than would be reasonably necessary to kill) and include bludgeoning, mutilation, burning, and dozens of stabbings per victim (Chesler, 2009, 2010; Mirza, 2008).

Intense torture and overkill suggest that there may be a unique psychology to honor killing which may require a special justification (i.e., religious motive[9]). That is, an especially strong emotional motivation may be needed to commit an honor killing because there are psychological mechanisms "designed" for caring for and protecting offspring or other kin under most circumstances. That is what makes honor killing appear deeply counterintuitive: Parents, siblings, cousins, and uncles killing their kin is puzzling from an evolutionary perspective. However, as with genital mutilation, there may be associated reproductive benefits (to the killers). Overkill suggests that there may also be an important communicative element to honor killing, such as upholding the honor, and therefore the status, of men in the family and community. If a man's reproductive success depended ancestrally on his (and his family's) status and reputation, and these are jeopardized (e.g., his daughter is accused of having sex before marriage, besmirching the father's and family's reputation; Quran 24:2, 17:32), then we expect him to attempt to repair his own and his family's reputation, sometimes by any means necessary—including killing the offender. If honor killing is an acceptable practice in the community, then this is a socially sanctioned way of restoring a man's social standing. Several passages in the Bible and the Quran appear to justify, even require, honor killings:

- Deuteronomy 22:13–21: If a man decides that he "hates" his wife, he can claim she wasn't a virgin when they were married. If her father can't produce the "tokens of her virginity" (bloody sheets), "then they shall bring out the damsel to the door of her father's house, and the men of her city shall stone her with stones that she die."
- Deuteronomy 22:23–24: "If a damsel that is a virgin be betrothed unto an husband, and a man find her in the city, and lie with her [t]hen ye shall bring them both out unto the gate of that city, and ye shall stone them with stones that they die."

[9] We did not find a single case of an honor killing that was *not* committed by a religious adherent and for a *non*religious reason.

shackelf@oakland.edu

In other words, if an engaged virgin is raped in the city and doesn't cry out loud enough, then the men of the city must stone her to death.

- Ezekiel 23:1–49: Two sisters were guilty of committing "whoredoms in their youth" by pressing their breasts and bruising "the teats of their virginity." One of the sisters "played the harlot" as she "doted her lovers… Thus she committed her whoredoms with them, with all them that were the chosen men of Assyria and with all on whom she doted. With all their idols she defiled herself. Neither left she her whoredoms brought from Egypt; for in her youth they lay with her, and they bruised the breasts of her virginity and poured their whoredom upon her," etc. As a punishment (they "executed judgment upon her"), one sister was stripped, her children were taken from her, and she was killed with a sword. The other sister was tortured by cutting her nose and ears off, and she was made to "pluck off" her own breasts, then she was raped and mutilated, and finally, stoned to death.
- Leviticus 21:9 – "And the daughter of any priest, if she profane herself by playing the whore, she profaneth her father: she shall be burnt with fire."
- Quran 4:15, 34 (see examples in *Mate guarding and controlling behaviors*).
- Quran 24:2 "The fornicatress and the fornicator, flog each of them with a hundred stripes. Let not pity withhold you in their case … if you believe in Allah … [a]nd let a party of the believers witness their punishment. (This punishment is for unmarried persons guilty of the above crime, but if married persons commit it (illegal sex), the punishment is to stone them to death, according to Allah's Law)."
- Mirza (2008) reviews several relevant sahih hadiths[10] that include statements such as "No one commits adultery while still remaining a believer, for faith is more precious unto Allah than such an evil act!" and "A woman came to the prophet and asked for purification by seeking punishment… [She] admitted she was pregnant… When the day arrived for the child to take solid food, Muhammad handed the child over to the community… [H]e had given command over her and she was put in a hole up to her breast, he ordered the people to stone her. Khalid b. al-Walid came forward with a stone which he threw at her head, and when the blood spurted on her face he cursed her."

According to these religious texts, adulterers and fornicators lose their rights and human value, bring great shame on their families, and should be punished with flogging, mutilation, and stoning to death. The victims' children are doomed as well, eliminating any reproductive success that forbidden copulation might otherwise offer. Further, the punishment is public and hard to fake. Honor killers will often announce their actions and goals publicly. For example, after 16-year-old Jordanian girl Kifaya Husayn was raped by her 21-year-old brother, her uncles persuaded another brother that she must die because she had disgraced their family by being raped. This 32-year-old brother bound her to a chair, told her to recite an Islamic

[10] Used in classification of the hadiths, it is the highest level of authenticity given to a narration (Sahih. 2013. In Wikiislam.net. Retrieved September 1, 2013, from http://wikiislam.net/wiki/Sahih).

shackelf@oakland.edu

prayer, and then slashed her throat. He then ran out into the street, waving the bloody knife, crying: "I have killed my sister to cleanse my honor" (Choo, 1998).

It may be important to note that the goal of honor killings, unlike intimate partner violence, for example, is to save or repair the honor and reputation of the victim's family and not for any animosity or wealth (Mirza, 2008). The honor killers love the girl as their own (daughter, sister, niece, etc.), but carry out the killing because they view it to be their moral obligation to save their family honor, erase damaging stigmas, and restore their religious piety (Mirza, 2008). After the honor killing, family members usually mourn and cry for the victim (even the killers themselves), but feel their actions are justified and necessary for all parties involved (Mirza, 2008).

Child Abuse and Filicide

Children are at the greatest risk of physical abuse and murder if they live with a stepparent, even after controlling for potential confounds such as socioeconomic status (Daly & Wilson, 1985, 1988, 1998; Wilson, Daly, & Weghorst, 1980). This increased risk of abuse and filicide by stepparents has been documented across diverse cultures (Bjorklund & Pellegrini, 2002; Daly & Wilson, 1988, 1998). Further, the risk is greater for children living with a stepfather than a stepmother (e.g., Daly & Wilson, 1994; Weekes-Shackelford & Shackelford, 2004). Accordingly, we expect that among religiously motivated cases of child abuse and filicide, there will be a higher proportion of stepparent perpetrators than genetic parent perpetrators, and fathers will be more violent than mothers. If stepparents are already motivated to harm (or care less for) stepchildren, might they exhibit this behavior even more intensely if they have religious justification? Here are some religious guidelines about parenting:

- Exodus 21:15: "And he that smiteth his father, or his mother, shall be surely put to death."
- Exodus 21:17: "And he that curseth his father, or his mother, shall surely be put to death."
- 2 Kings 2:23–24: "[A]nd as he was going up by the way, there came forth little children out of the city, and mocked him, and said unto him, Go up, thou bald head; go up, thou bald head. And he turned back, and looked on them, and cursed them in the name of the LORD. And there came forth two she bears out of the wood, and tare forty and two children of them."
- Proverbs 13:24: "He that spareth his rod hateth his son: but he that loveth him chasteneth him betimes."
- Proverbs 19:18: "Chasten thy son while there is hope, and let not thy soul spare for his crying."
- Proverbs 22:15: "Foolishness is bound in the heart of a child; but the rod of correction shall drive it far from him."

shackelf@oakland.edu

- Proverbs 23:13–14: "Withhold not correction from the child: for if thou beatest him with the rod, he shall not die. Thou shalt beat him with the rod, and shalt deliver his soul from hell."
- Proverbs 29:15: "The rod and reproof give wisdom: but a child left to himself bringeth his mother to shame."
- Hebrews 11:17: "By faith Abraham, when he was tried, offered up Isaac: and he that had received the promises offered up his only begotten son."

The Bible instructs parents to beat their children with rods and even kill them, as a sign of parental love, discipline, salvation, and devotion to God. Therefore, it may be reasonable to expect higher rates of child-directed violence in religious (vs. nonreligious) contexts. Further, high-risk situations such as cohabiting with a stepparent may be even riskier in religious (vs. nonreligious) contexts.

Religiously motivated child-directed violence is underreported (e.g., there is no national statistic of this phenomenon), perhaps because it is more acceptable to use physical punishments in religious communities (Heimlich, 2011). For example, Ellison and Bradshaw (2009) found that conservative Protestants spank their children more frequently than do other Christian believers and that religious beliefs (in a hierarchical God and in Hell) are associated with this difference. Heimlich reviews the urgent yet understudied topic of religious childhood maltreatment, including physical, emotional, and sexual abuse, and medical neglect.

Another context-specific risk for child-directed violence is maternal infanticide: a genetic mother killing her infant. High-risk factors include a mother's young age, low paternal investment prospects (mother is unpartnered or with a man other than the infant's father), mother's perception of the infant's low quality (e.g., illness, deformities), and mother's assessment of unfavorable child-rearing circumstances (e.g., multiple births, economic hardships) (Daly & Wilson, 1988). Young, poor, unmarried mothers are most likely to kill their newborns (e.g., d'Orban, 1979; Putkonen, Weizmann-Henelius, Collander, Santtila, & Eronen, 2007), a pattern that reflects unconscious, evolved decision rules in women about resource allocation (Daly & Wilson, 1988). However, even in societies where infanticide is lenient, killing her infant is extremely upsetting for the mother (Daly & Wilson, 1988). Although evolved psychological mechanisms influence infanticide, having a proximate justification may help facilitate the mother's behavioral decision and emotional consequences. We would expect religiously motivated infanticide to occur in similar conditions, and perhaps at a higher frequency, due to increased social acceptance and support for the mother. We would also expect some infanticidal mothers to provide explicitly religious explanations for the killings.

Other examples of religious doctrines exacerbating violence in contexts relevant to sexual selection are masturbation and birth control, rape of women during war, and terrorism. For example, in the book of Genesis (38:8–10), Judah tells Onan to "go in unto thy brother's wife" after the brother is killed. But "Onan knew that the seed should not be his; and … when he went in unto his brother's wife … he spilled it on the ground … And the thing which he did displeased the Lord; wherefore he slew him also." In other words, both masturbation and birth control are punishable by

shackelf@oakland.edu

death, according to this passage. If masturbation functions to increase one's sperm-competition ability (e.g., Baker & Bellis, 1995), it may be in the interest of men to condemn this behavior in others while practicing it themselves. If men secure their exclusive sexual access to their partner (as we've seen religious rules do), then we expect them to object to any form of contraception because the likelihood of cuckoldry is low and, therefore, their investment in their partner's offspring is more likely to provide fitness benefits rather than costs.

The following passages provide guidelines for behavior during war. Specifically, they instruct the soldiers to kill all the men, children, and nonvirgin women. Young, virgin women, on the other hand, should be raped and kept as sex slaves. Men are already motivated to eliminate their competitors, young and old (especially out-group members), and to copulate with fertile women. The following examples provide a convenient justification for this behavior:

- Numbers 31:1–54: Under God's direction, Moses' army defeats the Midianites. They kill all the adult males, but take the women and children captive. When Moses learns that they left some alive, he angrily says: "Have you saved all the women alive? Kill every male among the little ones, and kill every woman that hath known man by lying with him. But all the women children, that have not known a man by lying with him, keep alive for yourselves."
- Leviticus 20:13–14: In the cities that God "delivers into thine hands, thou shalt smite every male thereof with the edge of the sword: But the women, and the little ones … shalt thou take unto thyself."
- Zechariah 14:1–2, JKV: "God will make 'all nations' fight against Jerusalem. The women will be 'ravished' and half its people enslaved."
- Judges 21:11–23: "Ye shall utterly destroy every male, and every woman that hath lain by man. And they found among the inhabitants of Jabeshgilead four hundred young virgins, that had known no man by lying with any male: and they brought them unto the camp… [A]nd they [the Israelites] gave them [the Benjamites] wives which they had saved alive… and yet so they sufficed them [the Benjamites] not." To complete the number of virgins (one per Benjamite), they were instructed to ambush "the daughters of Shiloh [when they] come out to dance in dances, then come ye out of the vineyards, and catch you every man his wife of the daughters of Shiloh. And the children of Benjamin did so, and took them wives, according to their number, of them that danced, whom they caught."

The last passage is explicit about the conditions under which kidnapping and raping women of an out-group is required—when there is a shortage of women. Sex ratio is an important cue to sperm competition, and men are expected to be especially violent when there are indicators of a shortage in women (and, therefore, fiercer male intrasexual competition).

If men are capable of such severe violence for mating opportunities, we should not be surprised at the extremities of their violence when promised the ultimate, supernatural reward: dozens of beautiful virgins at their disposal. The Quran promises the ultimate reward for participating in terrorism (including suicide terrorism; see Sela & Shackelford, 2014):

shackelf@oakland.edu

- 3:151–157: "We shall cast terror into the hearts of those who disbelieve… O you who believe! Be not like those who disbelieve (hypocrites) and who say to their brethren when they travel through the earth or go out to fight: 'If they had stayed with us, they would not have died or been killed,' … It is Allah that gives life and causes death. And if you are killed or die in the Way of Allah, forgiveness and mercy from Allah are far better than all that they amass (of worldly wealths)."
- 4:74: "Let those (believers) who sell the life of this world for the Hereafter fight in the Cause of Allah, and whoso fights in the Cause of Allah, and is killed or gets victory, We shall bestow on him a great reward."
- 4:91: "You will find others that wish to have security from you and security from their people. Every time they are sent back to temptation, they yield thereto. If they withdraw not from you, nor offer you peace, nor restrain their hands, take (hold of) them and kill them wherever you find them. In their case, We have provided you with a clear warrant against them."
- 59:2: "Allah's (Torment) reached them [disbelievers] from a place whereof they expected it not, and He cast terror into their hearts so that they destroyed their own dwellings with their own hands and the hands of the believers."
- 4:95: "Allah has preferred in grades those who strive hard and fight with their wealth and their lives above those who sit (at home). Unto each, Allah has promised good (Paradise), but Allah has preferred those who strive hard and fight, above those who sit (at home) by a huge reward."
- 37:40–48: "the chosen slaves of Allah (i.e., the true believers of Islamic Monotheism) For them there will be a known provision (in Paradise), … In the Gardens of delight (Paradise)… [B]eside them will be Qdsirdt-at-Tarf [chaste females (wives), restraining their glances (desiring none except their husbands)], with wide and beautiful eyes."
- 44:51–54: The pious (faithful Muslims) "will be in place of Security (Paradise)… And we shall marry them to Hur (very fair females created by Allah as such, not from the offspring of Adam, with intense black irises of their eyes and intense white sclera) with wide, lovely eyes."

The descriptive details of young, beautiful, virgin women with healthy-looking eyes that grant exclusive sexual access to the devoted Muslim who fights with his life for Allah are consistent with male mating preferences from an evolutionary perspective.

Future Directions

The study of violent behaviors from an evolutionary perspective has been fruitful and has focused on identifying the psychological mechanisms, and their relevant contextual inputs, that lead to violence. Appreciating the role religious beliefs play in mediating violence is a crucial piece to understanding and, ultimately, minimizing violence. We have provided preliminary theoretical groundwork for thinking about how religion exacerbates violent tendencies, but such an endeavor is only useful if

shackelf@oakland.edu

it opens up new avenues for empirical study. We now provide some examples of how an evolutionary perspective in general, and a sexual selection framework in particular, can be applied to future research.

Several examples we have discussed relate to reputation and honor. It has previously been suggested that "cultures of honor" contribute to the differences in regional violence (e.g., in the United States; Nisbett & Cohen, 1996). Nisbett and colleagues have speculated that the underlying reason that gave rise to the psychological mechanisms underlying cultures of honor may have been theft of property. However, using a sexual selection analysis, Shackelford (2005) suggested that the adaptive problem may have been theft of wives, not property, placing this phenomenon in the mating context. Mechanisms of mate retention may be recruited in displaying behaviors consistent with a culture of honor. If manifest behavioral indicators of a culture of honor are the output of evolved psychological mate retention mechanisms, then to the extent that wifely infidelity rates remain higher in the southern United States than elsewhere in the nation, this may help to account for the persistence of a culture of honor to the present (Shackelford, 2005). In addition to the social mechanisms identified in recent research [e.g., collective representations that condone violence, such as laws (Cohen, 1996) and media representations (Cohen and Nisbett, 1997), and institutional non-stigmatization of violence (Cohen and Nisbett, 1997)], regional differences in recent female infidelity rates might help to account for the persistence of the southern culture of honor to the present. There are also known regional differences in religiosity (Newport, 2013; Putnam & Campbell, 2010). Might religion play a role in the persistence of cultures of honor?

We have focused on men's religiously motivated violent behaviors, but women can be violent too. Although female aggression is more often indirect and social rather than direct and physical (e.g., Archer, 2004; Björkqvist, 1994; Campbell, 1999), in some contexts women are the main perpetrators of violence (e.g., infanticide). It is important to consider the adaptive problems women have faced in sexual selection and parental investment, and possible solutions to these problems, when examining the possible influence of religion on female violence. Might religious women be guided to execute violence under similar conditions as nonreligious women and even to a greater degree? Do religious communities excuse female violence under these circumstances? Do religious texts promote female social aggression such as gossip and derogation of other women for not being religious or virtuous enough, as predicted by intrasexual competition?

Conclusion

Human violence in general is rarely committed arbitrarily. Psychopathic mass murder (along with other forms of violence stemming from psychological abnormalities) notwithstanding, violent behaviors often denote an evolutionary rationale, with violent behavior most likely to occur when the potential fitness benefits (on average, throughout our evolutionary history) of violence outweigh the potential costs.

shackelf@oakland.edu

Similarly, violence committed in the name of religion is rarely arbitrary, but rather conforms to the same evolutionarily rationale. That religious texts condone, or even command, violence in evolutionarily relevant contexts (e.g., cases of infidelity, threats to one's status, reproductive success, and/or fitness in general) is unlikely to be coincidental, but rather a reflection of the evolved psychological mechanisms of those who created such texts.

Furthermore, the status of religious texts—as a moral guide to life—provides further justification for particular forms of violence, which may serve as additional input to the collection of psychological mechanisms that ultimately determine one's course of action. Evolved psychological mechanisms that predispose one to commit violence (given specific environmental inputs) do not function in a vacuum; humans have evolved predispositions for cooperation as well, and we should expect conflict between psychological mechanisms that promote conflicting behaviors (Kurzban, 2010). Although the complex interactions between evolved psychological mechanisms are only beginning to be understood, it is probable that mechanisms promoting violence must "compete" with mechanisms that deter violence. If religious texts and beliefs serve as input to additional mechanisms involved in decision-making, these inputs may increase the likelihood of violence (as well as cooperation) in the appropriate contexts. Although there is still much that we don't understand regarding the interaction of religious beliefs and evolved psychological mechanisms, it is clear that we cannot afford to ignore the role that religion plays in exacerbating violence, and an evolutionary perspective may provide us with a better understanding of how and why this occurs.

Acknowledgment This chapter is based on Sela, Shackelford, and Liddle (2015).

References

Al-Munajjid, M. S. (n.d.a). Fatwa #9412: Circumcision: How it is done and the rulings on it. *Islam Q&A.* Retrieved August 29, 2013, from http://www.islam-qa.com/en/ref/9412

Al-Munajjid, M. S. (n.d.b). Fatwa #7073: The health and religious benefits of circumcision. *Islam Q&A.* Retrieved August 29, 2013, from http://www.islam-qa.com/en/ref/7073

Andersson, M. (1994). *Sexual selection*. Princeton, NJ: Princeton University Press.

Anonymous. (2013, June 8). In cold blood: Killed for stepping out. *The Express Tribune*. Retrieved August 29, 2013, from http://tribune.com.pk/story/560597/in-cold-blood-killed-for-stepping-out/

Archer, J. (2004). Sex differences in aggression in real world settings: A meta-analytic review. *Review of General Psychology, 8*, 291–322. doi:10.1037/1089-2680.8.4.291.

Archer, J. (2009). Does sexual selection explain human sex differences in aggression? *Behavioral and Brain Sciences, 32*, 249–311.

Associated Press. (2008, October 6). Jewish "modesty patrols" sow fear in Israel. *Ynet News.* Retrieved from http://www.ynetnews.com/articles/0,7340,L-3605707,00.html

Baker, R. R., & Bellis, M. A. (1995). *Human sperm competition: Copulation, masturbation and infidelity*. London: Chapman & Hall.

Björkqvist, K. (1994). Sex differences in physical, verbal, and indirect aggression: A review of recent research. *Sex roles, 30*, 177–188.

shackelf@oakland.edu

Bjorklund, D. F., & Pellegrini, A. D. (2002). *The origins of human nature: Evolutionary developmental psychology*. Washington, DC: American Psychological Association.
Bureau of Justice Statistics. (2003, February). *Intimate partner violence, 1993–2001*. Washington, DC: U.S. Department of Justice, Office of Justice Programs.
Bureau of Justice Statistics. (2007, December). *Intimate partner violence: Victim characteristics, 1976–2005.* Washington, DC: U.S. Department of Justice, Office of Justice Programs.
Burbank, V. K. (1992). Sex, gender, and difference. *Human Nature, 3*, 251–277.
Buss, D. M. (2000). *The dangerous passion*. New York: Free Press.
Buss, D. M. (2005). *The murderer next door: Why the mind is designed to kill*. New York: Penguin.
Buss, D. M. (2011). *Evolutionary psychology: The new science of the mind* (4th ed.). Boston: Pearson.
Buss, D. M., Larsen, R. J., Westen, D., & Semmelroth, J. (1992). Sex differences in jealousy: Evolution, physiology, and psychology. *Psychological Science, 3*, 251–255.
Buss, D. M., & Shackelford, T. K. (1997a). From vigilance to violence: Mate retention tactics in married couples. *Journal of Personality and Social Psychology, 72*, 346–361.
Buss, D. M., & Shackelford, T. K. (1997b). Human aggression in evolutionary psychological perspective. *Clinical Psychology Review, 17*, 605–619. doi:10.1016/S0272-7358(97)00037-8.
Campbell, A. (1993). Men, women, and aggression. New York, NY: Basic Books.
Campbell, A. (1999). Staying alive: Evolution, culture, and women's intrasexual aggression. *Behavioral and Brain Sciences, 22*, 203–214.
Campbell, A. (2005). Aggression. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 628–652). Hoboken, NJ: Wiley.
Chagnon, N. A. (1988). Life histories, blood revenge, and warfare in a tribal population. *Science, 239*, 985–992. doi:10.1126/science.239.4843.985.
Chagnon, N. A. (1992). Yanomamö (4th ed.). Fort Worth, TX: Harcourt Brace Jovanovich College Publishers.
Chesler, P. (2009). Are honor killings simply domestic violence? *Middle East Quarterly, XVI*, 61–69. Retrieved from http://www.meforum.org/2067/are-honorkillings-simply-domestic-violence
Chesler, P. (2010). Worldwide trends in honor killings. *Middle East Quarterly, XVII*, 3–11. Retrieved from http://www.meforum.org/2646/worldwide-trends-in-honorkillings
Choo, K. (1998, May 3). Star reporter: Probe of "honor killings" of women shocks Jordanians. *Chicago Tribune Lifestyle.* Retrieved from http://articles.chicagotribune.com/1998-05-03/features/9805030063_1_honorkillings-jordan-times-crime
Coleman, S. W., Patricelli, G. L., & Borgia, G. (2004). Variable female preferences drive complex male displays. *Nature, 428*, 742–745.
Cohen, D. (1996). Law, social policy, and violence: The impact of regional cultures. *Journal of Personality and Social Psychology, 70*, 961–978.
Cohen, D. & Nisbett, R. E. (1997). Field experiments examining the culture of honor: The role of institutions in perpetuating norms about violence. *Personality and Social Psychology Bulletin, 23*, 1188–1199.
Cohen, D. (2011, July 17). Morality police hit Tehran streets. *Ynet News.* Retrieved from http://www.ynetnews.com/articles/0,7340,L-4096487,00.html
d'Orban, P. T. (1979). Women who kill their children. *The British Journal of Psychiatry, 134*, 560–571.
Daly, M., & Wilson, M. (1985). Child abuse and other risks of not living with both parents. *Ethology and Sociobiology, 6*, 197–210.
Daly, M., & Wilson, M. (1988). *Homicide*. Hawthorne, NY: Aldine de Gruyter.
Daly, M., & Wilson, M. I. (1994). Some differential attributes of lethal assaults on small children by stepfathers versus genetic fathers. *Ethology and Sociobiology, 15*, 207–217.
Daly, M., & Wilson, M. (1998). *The truth about Cinderella*. London: Weidenfeld & Nicolson.
Daly, M., Wilson, M., & Weghorst, J. (1982). Male sexual jealousy. *Ethology and Sociobiology, 3*, 11–27.
Darwin, C. (1871). *The descent of man, and selection in relation to sex*. London: John Murray.
Dobash, R. E., & Dobash, R. P. (1979). *Violence against wives*. New York: Free Press.

shackelf@oakland.edu

Dutton, D. G. (1998). *The abusive personality*. New York: Guilford Press.
Ellison, C. G., & Bradshaw, M. (2009). Religious beliefs, sociopolitical ideology, and attitudes toward corporal punishment. *Journal of Family Issues, 30*, 320–340.
Easton, J. A., & Shackelford, T. K. (2009). Morbid jealousy and sex differences in partner-directed violence. *Human Nature, 20*, 342–350.
Ferguson, R. B. (1997). Violence and war in prehistory. In D. L. Martin & D. W. Frayer (Eds.), *Troubled times: Violence and war in the past. War and society* (Vol. 3, pp. 321–356). Amsterdam, The Netherlands: Gordon Breach.
Frieze, I. H. (1983). Investigating the causes and consequences of marital rape. *Signs: Journal of Women in Culture and Society, 8*, 532–553.
Gage, A. J., & Hutchinson, P. L. (2006). Power, control, and intimate partner sexual violence in Haiti. *Archives of Sexual Behavior, 35*, 11–24.
Ghiglieri, M. P. (1999). *The dark side of man: Tracing the origins of male violence*. New York: Basic Books.
Goetz, A. T., Shackelford, T. K., Platek, S. M., Starratt, V. G., & McKibbin, W. F. (2007). Sperm competition in humans: Implications for male sexual psychology, physiology, anatomy, and behavior. *Annual Review of Sex Research, 18*, 1–22.
Goetz, A. T., Shackelford, T. K., Starratt, V. G., & McKibbin, W. F. (2008). Intimate partner violence. In J. D. Duntley & T. K. Shackelford (Eds.), *Evolutionary forensic psychology* (pp. 65–78). New York: Oxford University Press.
Heimlich, J. (2011). *Breaking their will: Shedding light on religious child maltreatment*. New York: Prometheus Books.
Holden, C., & Mace, R. (1999). Sexual dimorphism in stature and women's work: A phylogenetic cross-cultural analysis. *American Journal of Physical Anthropology, 110*, 27–45.
Jafri, O. (2013a, August 17). Free will: Man kills daughter over disobedience. *The Express Tribune.* Retrieved from http://tribune.com.pk/story/591237/free-will-man-kills-daughter-over-disobedience/
Jafri, O. (2013b, August 22). Cut into pieces: "She challenged God's orders". *The Express Tribune.* Retrieved from http://tribune.com.pk/story/593466/cut-into-pieces-she-challenged-gods-orders/
Kaighobadi, F., Shackelford, T. K., & Goetz, A. T. (2009). From mate retention to murder: Evolutionary psychological perspectives on men's partner-directed violence. *Review of General Psychology, 13*, 327–334.
Keeley, L. H. (1996). *War before civilization: The myth of the peaceful savage*. New York: Oxford University Press.
Kurzban, R. (2010). *Why everyone (else) is a hypocrite*. Princeton, NJ: Princeton University Press.
Lester, D. (1991). *Questions and answers about murder*. Philadelphia: Charles Press.
Liddle, J. R., Shackelford, T. K., & Weekes-Shackelford, V. A. (2012). Why can't we all just get along? Evolutionary perspectives on violence, homicide, and war. *Review of General Psychology, 16*, 24–36. http://www.toddkshackelford.com/downloads/Liddle-Shackelford-Weekes-Shackelford-RGP-2012.pdf
Mirza, S. K. (2008, January 16). Honor killing is absolutely Islamic. *Islam Watch* (Original article published online 2005). Retrieved from http://www.islam-watch.org/syedkamranmirza/honor_killing.htm
Morris, B. J., & Krieger, J. N. (2013). Does male circumcision affect sexual function, sensitivity, or satisfaction?—a systematic review. *The journal of sexual medicine, 10*, 2644–2657.
Newport, F. (2013, February 13). *Mississippi maintains hold as most religious U.S. state*. Retrieved from http://www.gallup.com/poll/160415/mississippi-maintains-hold-religious-state.aspx
Newton-Fisher, N. E., & Thompson, M. E. (2012). Comparative evolutionary perspectives on violence. In T. K. Shackelford & V. A. Weekes-Shackelford (Eds.), *The Oxford handbook of evolutionary perspectives on violence, homicide, and war* (pp. 41–60). New York: Oxford University Press.
Nisbett, R. E., & Cohen, D. (1996). *Culture of honor*. Boulder, CO: Westview Press.
Palmer, C. T., & Tilly, C. F. (1995). Sexual access to females as a motivation for joining gangs: An evolutionary approach. *Journal of Sex Research, 32*, 213–217. doi:10.1080/00224499509551792.

shackelf@oakland.edu

Pew Research Center. (2014, May). *Worldwide, many see belief in God as essential to morality*. Retrieved from http://www.pewglobal.org/files/2014/05/Pew-Research-Center-Global-Attitudes-Project-Belief-in-God-Report-REVISED-MAY-27-2014.pdf

Putkonen, H., Weizmann-Henelius, G., Collander, J., Santtila, P., & Eronen, M. (2007). Neonaticides may be more preventable and heterogeneous than previously thought – neonaticides in Finland 1980–2000. *Archives of Women's Mental Health, 10*, 15–23.

Putnam, R. D., & Campbell, D. E. (2010). *American grace: How religion divides and unites us*. New York: Simon & Schuster.

Razavi, S. (2006). Islamic politics, human rights and women's claims for equality in Iran. *Third World Quarterly, 27*, 1223–1237.

Russell, D. E. H. (1982). *Rape in marriage*. New York: Macmillan.

Schmitt, D. P. (2003). Universal sex differences in the desire for sexual variety: Tests from 52 nations, 6 continents, and 13 islands. *Journal of Personality and Social Psychology, 85*, 85–104.

Sela, Y., & Shackelford, T.K. (2015). Evolution. In R. Cautin and S. Lilienfeld (Eds.), *The encyclopedia of clinical psychology* (pp. 1-7). Hoboken, NJ: Wiley-Blackwell.

Sela, Y., & Shackelford, T. K. (2014). The myth of the myth of martyrdom. *Behavioral and Brain Sciences, 37*, 376–377.

Sela, Y., Shackelford, T. K., & Liddle, J. R. (2015). When religion makes it worse: Religiously-motivated violence as a sexual selection weapon. In J. Slone & J. Van Slyke (Eds.), *The attraction of religion* (pp. 111–131). London: Bloomsbury Academic.

Shackelford, T. K. (2005). An evolutionary psychological analysis of cultures of honor. *Evolutionary Psychology, 3*, 381–391.

Shackelford, T. K., & Weekes-Shackelford, V. A. (Eds.). (2012). *The Oxford handbook of evolutionary perspectives on violence, homicide, and war*. New York: Oxford University Press.

Slack, A. T. (1988). Female circumcision: A critical appraisal. *Human Rights Quarterly, 10*, 437–486.

Smith, D. L. (2007). *The most dangerous animal. Human nature and the origins of war*. New York: St. Martin's Press.

Spolar, C. (2005, November 17). For family honor, she had to die. *Chicago Tribune News*. Retrieved from http://articles.chicagotribune.com/2005-11-17/news/0511170188_1_honorkillings-family-honor hatun-surucu

Trivers, R. L. (1972). Parental investment and sexual selection. In B. Campbell (Ed.), *Sexual selection and the descent of man 1871–1971* (pp. 136–179). Chicago: Aldine.

Trivers, R. L. (1985). *Social evolution*. Menlo Park, CA: Benjamin/Cummings.

Walker, L. E. (1979). *The battered woman*. New York: Harper & Row.

Weekes-Shackelford, V. A., & Shackelford, T. K. (2004). Methods of filicide: Stepparents and genetic parents kill differently. *Violence and Victims, 19*, 75–81.

Wilson, M., Daly, M., & Weghorst, S. J. (1980). Household composition and the risk of child abuse and neglect. *Journal of Biosocial Science, 12*, 333–340.

Wilson, C. G. (2008). Male genital mutilation: an adaptation to sexual conflict. *Evolution and Human Behavior, 29*(3), 149–164.

World Health Organization (2014). Fact sheet N°241 – Female Genital Mutilation. Retrieved June 17, 2015, from http://www.who.int/mediacentre/factsheets/fs241/en/.

shackelf@oakland.edu

Disentangling Religion and Morality: An Analysis of Religiosity in the United States

James R. Liddle

Introduction

Religious beliefs and behaviors, although they exist in many forms (Moro & Myers, 2010; Smith, 1991) and may not always be viewed as a formal "religion" by their practitioners, appear to be human universals (Atran, 2002; Norenzayan, 2010). Although there is debate regarding how to define religion, the assertion of religion's universality is based on a definition by Atran (2002, p. 13) that is broad enough to avoid excluding less "traditional" religions and comprises four components:

1. Widespread counterfactual belief in supernatural agents (gods, ghosts, goblins, etc.)
2. Hard-to-fake public expressions of costly material commitments to supernatural agents—that is, sacrifice (offerings of goods, time, other lives, one's own life, etc.)
3. A central focus of supernatural agents on dealing with people's existential anxieties (death, disease, catastrophe, pain, loneliness, injustice, want, loss, etc.)
4. Ritualized and often rhythmic coordination of 1, 2, and 3—that is, communion (congregation, intimate fellowship, etc.)

These components exist in all known societies, converging into what may reasonably be defined as "religion."

Despite the apparent universality of religion, the *degree* of religiousness (i.e., religiosity) varies across individuals, societies, and time. For example, Sweden and Denmark are among the least religious nations (excluding nations with state-imposed atheism, which do not accurately represent the religiosity of the populace), with studies reporting a range of 46–85 % of Swedes and 43–80 % of Danes as nonbelievers in God (Zuckerman, 2007). These percentages may be large when compared to other nations, but they also highlight the substantial variability in

J.R. Liddle (✉)
Department of Psychology, Florida Atlantic University, Boca Raton, FL 33431, USA
e-mail: james.r.liddle@gmail.com; jliddle1@fau.edu

© Springer International Publishing Switzerland 2016

T.K. Shackelford, R.D. Hansen (eds.), *The Evolution of Morality*,
Evolutionary Psychology, DOI 10.1007/978-3-319-19671-8_11

shackelf@oakland.edu

religiosity at the individual level. Even though these societies are highly secular, there is a substantial portion of highly religious individuals in these populations (Zuckerman, 2008). Furthermore, this predominant secularism has not been constant throughout these nations' histories, as Zuckerman (2008) notes regarding the late 1700s and 1800s:

> …there is no question that heartfelt, faithful Christianity was discernibly pervasive in various parts of Denmark and Sweden," and "…ever since sociologists began collecting data on religion in Denmark and Sweden—which, admittedly, really wasn't that long ago—the clear pattern has been that of decline, in both belief and participation. (p. 125)

Besides Denmark and Sweden, developed, postindustrial nations tend to exhibit lower religiosity than less-developed nations (Norris & Inglehart, 2004; Zuckerman, 2009). The United States, however, appears to be an exception. Recent estimates of the percentage of atheists, agnostics, or nonbelievers in the United States range from 3 to 9 %, which places the United States 44th—between Portugal and Albania—on a list of the top 50 countries with citizens self-identified as such (Zuckerman, 2007). Nevertheless, United States religiosity has been declining in recent decades. An analysis of the 1990, 2001, and 2008 waves of the American Religious Identification Survey (ARIS) indicates that the percentage of Americans designated as religious "nones" (i.e., those who do not identify with any particular religion) increased from 14.3 million (about 8 % of the population) in 1990 to 34.2 million (15 %) in 2008 (Kosmin & Keysar, 2009).

How can we explain this individual, societal, and temporal variability in religiosity? Although historical and cultural factors play a role in shaping individual and societal religiosity over time, it can also be useful to analyze religion from an evolutionary perspective.

Evolutionary Psychology and Its Application to Religion

Evolutionary psychology is not a subdiscipline of psychology, such as social psychology or personality psychology, but rather an *approach* to psychology that applies evolutionary theory. Evolutionary psychology is founded on the premise that the brain, like every other organ, has evolved and is therefore open to analysis from an evolutionary perspective, which means that the products of the brain (i.e., thoughts, feelings, behaviors) are open to evolutionary analysis. For example, an evolutionary psychological approach has proven useful in examining social behavior (Cosmides, 1989; Kenrick, Maner, & Li, 2005), learning (MacDonald, 2007; Weber & Depew, 2003), memory (McBurney, Gaulin, Devineni, & Adams, 1997), development (Bjorklund & Pellegrini, 2000), and perception (Rhodes, 2006), to name a few diverse topics. All aspects of psychology have the potential to be better understood by examining them from an evolutionary perspective, and religious beliefs and behaviors are no exception.

Evolutionary psychologists argue that the mind is composed of a large number of evolved mechanisms shaped throughout evolutionary history to solve specific,

shackelf@oakland.edu

recurrent adaptive problems of survival and reproduction. Although the number of evolved mechanisms that exist and the typical scope of such mechanisms (i.e., domain specific vs. domain general) are subjects of debate, a less controversial aspect of this view is the description of evolved mechanisms as information-processing mechanisms sensitive to specific types of information (i.e., environmental stimuli, physiological activity, output from other parts of the brain), and this information is processed, resulting in a specific type of output (i.e., physiological activity, input to other mechanisms, or manifest behavior) (Buss, 2011).

From this conceptual foundation, one can examine religious beliefs and behaviors by considering the possible evolved psychological mechanisms that produce such output. Identifying the types of information—particularly environmental input—that influence religiosity can aid in identifying the mechanisms involved in producing religious beliefs and behaviors and the functions of these mechanisms, which can aid in determining whether religious beliefs and behaviors are produced by specialized mechanisms or as the byproducts of other evolved mechanisms (i.e., output that is merely a consequence of a mechanism's design rather than output that a mechanism is specialized to produce).

Although this chapter presents a test of a theory of religiosity that was not developed by evolutionary psychologists, an evolutionary perspective can be useful in understanding the theory's predictions and interpreting findings related to this theory.

Secure Society Theory

Norris and Inglehart (2004) proposed a theory of secularization that may explain the variability in religiosity between nations as well as societal changes in religiosity over time. They argue that a key factor driving secularization is the level of security provided by a society, which influences individuals' "existential security" or "…the feeling that survival is secure enough that it can be taken for granted" (p. 4). Their theory is referred to throughout this chapter as "Secure Society Theory."

Secure Society Theory is built on two premises, referred to by Norris and Inglehart (2004) as "the Security Axiom" and "the Cultural Traditions Axiom" (pp. 13–18). The Cultural Traditions Axiom does not play a large role in the theory's explanation of secularization, but rather emphasizes that the religious worldviews of a society continue to influence that society's culture even as that society moves toward secularization (e.g., the Protestant work ethic). More relevant to explaining religious variability over time is the Security Axiom, which states that variability exists between societies with regard to the level of security (i.e., people's vulnerability to risks and dangers, such as environmental disasters, diseases, crime, human rights violations, poverty, etc.) and that societal shifts from agrarian to industrial and from industrial to postindustrial improve societal security. The first stage of modernization (agrarian to industrial) lifts developing nations out of extreme poverty, aiding the most vulnerable portions of the population and improving

shackelf@oakland.edu

the standard of living. As societies develop, there are improvements in nutrition, sanitation, access to clean water, healthcare, and education, as well as improved mass communication. These changes have a positive impact on individuals' perceptions of security (i.e., existential security).

Norris and Inglehart (2004) acknowledge that societal development does not inevitably lead to greater security, at least not for all citizens. This makes sense when considering Secure Society Theory from an evolutionary perspective. If the psychological mechanisms that produce religious beliefs and behaviors are sensitive to societal security, then even a prosperous, developed nation can have a religious populace if cues to insecurity are present. For example, certain events can have significant negative impacts on any nation regardless of the level of development (e.g., natural disasters, war, recession) and will impact individuals' perceptions of security. Economic inequality is also is an important variable, as a substantial portion of the population may continue to suffer from threats to their security while a small "elite" class of citizens reaps the benefits of development.

Using the Security Axiom and Cultural Traditions Axiom as their foundation, Norris and Inglehart (2004) hypothesize that the variability in security between societies, resulting from varying levels of development and historical events, can partially explain the variability in religiosity between societies, with greater security leading to increased secularization. More specifically, they predict that greater security will result in weaker religious beliefs, values, and participation and that differences in religiosity will be most pronounced between agrarian, industrial, and postindustrial societies.

This relationship between religiosity and security is based on the argument that as individuals' perceptions of security increase, their need for religion decreases, as religion—particularly supernatural beliefs—provides a coping mechanism for living in less secure and unpredictable conditions. That supernatural beliefs afford coping with uncertainty was hypothesized by Malinowski (1954) and has since been supported by a variety of studies. For example, regular church attendance is linked to a reduced incidence of depression, suggesting greater ability to cope with stress (McCullough & Larson, 1999), and several studies indicate that individuals compensate for uncertainty, ambiguity, and reduced feelings of control through superstitious (Burger & Lynn, 2005; Case, Fitness, Cairns, & Stevenson, 2004; Keinan, 2002) and religious (Frijters & Barón, 2010; Kay, Gaucher, McGregor, & Nash, 2010; Kay, Moscovitch, & Laurin, 2010; Kay, Whitson, Gaucher, & Galinsky, 2009) beliefs and behaviors. Whitson and Galinsky (2008) have even shown that the experience of lacking control can increase the perception of illusory patterns, including developing superstitions and forming illusory correlations about stock market data. In evolutionary terms, these findings suggest that religion functions as an adaptation (i.e., that religious beliefs and behaviors are designed output of psychological mechanisms responding to environmental threats), a view implicit in Secure Society Theory and further considered in the Discussion below.

Norris and Inglehart (2004) expand on the hypothesized relationship between security and religiosity by noting that societal changes in security are not expected to have an immediate impact on individuals' religiosity, but rather that these effects

shackelf@oakland.edu

should take time. If Secure Society Theory is correct, changes in religiosity over time may correspond to societal changes in security. Identifying this relationship requires consistently gathered data over an adequate length of time. Evidence regarding the hypothesized relationship between security and religiosity is discussed next.

Evidence Supporting Secure Society Theory

Norris and Inglehart (2004) conducted a series of analyses to test Secure Society Theory. Their primary source of data for religiosity was the pooled World Values Survey/European Values Survey conducted in four waves from 1981 to 2001. This data set provided information about 76 nation-states. However, not all nations were included in each wave, so time series analyses were limited to 20 societies.

The specific measures of religiosity included religious participation, both collective (attending religious services) and personal (prayer frequency), religious values (the importance of religion in one's life), and religious beliefs (belief in God, heaven/hell, life after death, and existence of the soul). Regarding societal security, Norris and Inglehart (2004) categorized societies as agrarian ($n = 23$), industrial ($n = 33$), and postindustrial ($n = 23$) based on the Human Development Index, a 100-point scale of societal modernization published annually by the United Nations Development Programme. This measure combines levels of knowledge (adult literacy and education), health (life expectancy at birth), and standard of living (real per capita GDP). Additional measures of security/development drawn from a variety of sources included the proportion of the population living in rural and urban areas, the Gini coefficient of economic inequality, access to mass communications, the number of HIV/AIDS cases, access to an improved water source, immunization rates, the distribution of physicians, and average life expectancy at birth.

Norris and Inglehart (2004) first examined differences in religiosity between agrarian, industrial, and postindustrial nations. As hypothesized by Secure Society Theory, religious participation, values, and beliefs were strongest in agrarian societies and weakest in postindustrial societies. For example, 54 % of respondents in agrarian societies reported praying every day, compared to 34 % and 26 % of those living in industrial and postindustrial societies, respectively.

Norris and Inglehart (2004) then conducted correlational analyses between the various measures of security/development and religious behavior (attending religious services and prayer frequency). These results also supported Secure Society Theory, with each societal indicator correlating with both religious participation and prayer frequency in the predicted direction: As societal conditions improve, religiosity decreases. The correlations ranged in strength from 0.41 to −0.74. For example, the Human Development Index was negatively correlated with both religious participation and prayer frequency, $rs = -0.53$, $ps < 0.001$.

Norris and Inglehart (2004) also examined historical trends regarding religiosity, specifically the annual trends in regular (weekly) religious service participation from 1970 to 1998 for 13 European societies. For each society, the year of the survey

shackelf@oakland.edu

was regressed on the proportion of respondents reporting weekly religious service attendance. Every model resulted in a negative regression coefficient, and this result was statistically significant for nine of the societies. Although these results show that religiosity is in decline in these European societies, the analyses did not address the causal effect of societal security on these declines.

Finally, Norris and Inglehart (2004) note that the United States appears to represent an exception to their theory. Although the United States religiosity is declining, it remains an outlier compared to most postindustrial nations. The high level of religiosity observed in the United States seemingly contradicts Secure Society Theory, as the United States is a successful postindustrial nation. However, when the United States is analyzed in terms of societal indicators of *security*, the high rates of religiosity are less anomalous. For example, Norris and Inglehart highlight that the United States exhibits greater economic inequality (as measured by the Gini coefficient) than any other postindustrial nation included in their analyses. They further state that:

> Many American families, even in the professional middle classes, face risks of unemployment, the dangers of sudden ill health without adequate private medical insurance, vulnerability to becoming a victim of crime, and the problems of paying for long-term care of the elderly. (p. 108)

Although Norris and Inglehart do not analyze these additional factors and their relationship to religiosity statistically, they make a strong case for the value of Secure Society Theory for understanding religiosity in the United States, as well as throughout the world. Nevertheless, it is important to consider additional evidence regarding the validity of Secure Society Theory.

Additional Supporting Evidence of Secure Society Theory

Since Norris and Inglehart's (2004) initial presentation of their theory, several researchers have further tested Secure Society Theory either implicitly or explicitly. For example, Paul (2005) tested the hypothesis that popular religiosity is beneficial to society by examining rates of religious belief and practice along with several indicators of societal health and dysfunction (homicide, youth suicide, sexually transmitted disease prevalence, teen pregnancy and birth, and abortion rates) in 18 developed democracies, including the United States. Paul (p. 7) concluded that "higher rates of belief in and worship of a creator correlate with higher rates of homicide, juvenile and early adult mortality, STD infection rates, teen pregnancy, and abortion in the prosperous democracies." Furthermore, the United States is an outlier regarding most societal indicators of dysfunction, with homicide rates, STD infection rates, early adolescent pregnancies, and abortion rates much higher than in the other countries analyzed.

Given that the United States is an outlier among prosperous democracies on several indicators of societal security, it is important to consider whether the observed relationship between religiosity and security is being driven primarily by

shackelf@oakland.edu

the inclusion of the United States in statistical analyses. This possibility was considered in a later study by Paul (2009), in which analyses similar to those conducted earlier (Paul, 2005) were performed with and without the United States included. The newer study also included the creation of a "Successful Societies Scale" (SSS) based on over two dozen indicators—several of which were used in Paul's (2005) study—and a "Popular Religiosity Versus Secularism Scale" (PRVSS) comprising seven measures of religiosity and secularism (absolute belief in a supernatural creator deity, biblical literalism, religious service attendance, prayer frequency, belief in an afterlife, self-reported agnosticism/atheism, and acceptance of human descent from animals). Higher scores on the SSS indicated less societal dysfunction, and higher scores on the PRVSS indicated higher levels of secularization. Results indicated that scores on the SSS positively correlated with scores on the PRVSS, with the United States included ($r=0.71$, $p<0.001$) and excluded ($r=0.53$, $p<0.01$), although the relationship is larger with United States inclusion. These results are consistent with Secure Society Theory, as many of the societal measures are indicators of societal security (e.g., homicides, incarcerations, life expectancy, infant mortality, Human Poverty Index), and the relationship with religiosity is not driven by data from the United States.

A study by Rees (2009) further tested Secure Society Theory with an analysis of 55 countries. This study also tested alternative explanations for changes in religiosity: the traditional modernization theory of secularization and Rational Choice Theory, the latter of which states that secularization occurs "due to competition for attention from secular services and the provision of unattractive products by the monopoly of religious providers" (Rees, 2009; p. 2). After establishing economic inequality—measured by the Gini coefficient—to be a reasonable proxy for personal insecurity given its correlation with several societal indicators of security, Rees developed a model with economic equality and variables related to the alternative explanations of secularization (governmental and social regulation of religion, religious fractionalization, and per capita GDP) as predictors of religiosity. The results supported Secure Society Theory, indicating that while taking into account other variables, economic inequality remains a unique predictor and was indeed the strongest predictor of religiosity.

Barber (2011) provided further support for the link between economic inequality—as well as other variables likely to influence existential security—and religiosity. As a measure of religious *disbelief*, Barber relied on the proportion of the population reporting that they do not believe in God, as compiled by Zuckerman (2007) for 137 countries. Barber controlled for the effect of living in Communist societies (where religious beliefs are criminalized) and Islamic states that observe Sharia law (where atheism is criminalized). Independent variables included economic development measured in terms of the proportion of the labor force employed in agriculture and third-level education enrollment, economic security measured in terms of the Gini coefficient and the level of personal taxation (a proxy for the extent of the welfare state), and health security measured in terms of the severity of 22 parasites (i.e., "pathogen prevalence") as reported by Fincher and Thornhill (2008). The results indicated that religious disbelief was correlated with all the independent and control

shackelf@oakland.edu

variables, and these variables were all predictors of religious disbelief in regression analyses, explaining 75 % of the variance in disbelief.

Pesta, McDaniel, and Bertsch (2010) provided an indirect test of Secure Society Theory by creating an index of well-being for the United States, using the 50 states as the units of analysis. They identified six "sub-domains" of subjective well-being for which state-level data are available: religiosity, health, crime, education, income level, and *g*, or general intelligence. Although religiosity was included because of its documented positive effect on well-being (see Pesta et al., 2010), correlational analyses indicated that religiosity was positively correlated with the only sub-domain representing lower well-being—crime—and negatively correlated with every other sub-domain representing greater well-being. In other words, although religiosity has beneficial effects on well-being at the individual level, higher levels of religiosity are associated with lower-state-level well-being. This apparent contradiction is consistent with Secure Society Theory, as individuals living in states with stronger indicators of lower well-being (e.g., higher crime rates, lower health, lower education, etc.) are expected to display greater religiosity in an effort to cope with these conditions.

Finally, a study by Solt, Habel, and Grant (2011) analyzed economic inequality and religiosity over time, providing a test of the temporal component of Secure Society Theory (i.e., changes in societal security result in changes in religiosity over time). Solt et al. analyzed data over a 50-year period, from 1955 to 2005. Grant's (2008) Aggregate Religiosity Index (ARI), which provides a single value of national religiosity for each year based on available survey data, was used as the measure of religiosity, and economic inequality was measured by the Gini coefficient. GDP per capita was also included in their analysis. To test the effects that these variables have on each other over time, Solt et al. analyzed these data with vector autoregression, a form of time series analysis that comprises a series of regression equations. Each variable under consideration serves as the predictor in one of the equations. More specifically, a time-lagged version of each variable serves as a predictor to determine whether it can predict future values of the other variables. The analyses indicated that GDP per capita negatively predicts future religiosity, whereas economic inequality positively predicts future religiosity (i.e., as inequality increases, future religiosity increases). However, the time lag considered in this analysis was only 1 year, as preliminary analyses indicated this was the most appropriate lag. Therefore, although the results are in the direction hypothesized by Secure Society Theory, such a brief time lag may not provide a proper test of Secure Society Theory's hypothesis regarding gradual historical change in religiosity.

The Present Study

A major limitation of the previous research on Secure Society Theory is the lack of direct data regarding *existential security* (i.e., people's personal *perceptions* of security). Instead, religiosity has been linked to societal indicators of security. Though

shackelf@oakland.edu

this type of investigation is important for testing Secure Society Theory, as existential security should be correlated with societal conditions, it is necessary to demonstrate the relationship hypothesized by Norris and Inglehart (2004) between personal perceptions of security and religiosity.

Another limitation of previous studies is the lack of time series analyses, because Secure Society Theory hypothesizes that changes in security over time will lead to changes in religiosity. Although Norris and Inglehart (2004) analyzed historical changes in religiosity, they did not investigate the hypothesized causal relationship between security and religiosity over time. Solt et al. (2011) attempted to address this, but their time series analysis relied on a time lag of just 1 year. Although their results supported the hypothesized causal relationship between changes in security and religiosity, their brief time lag is not consistent with the gradual (i.e., generational) changes hypothesized by Secure Society Theory.

The present study attempts to address these limitations as well as replicate previous findings regarding Secure Society Theory by analyzing religiosity, societal security, and perceptions of security in the United States from 1972 to 2012 through use of the General Social Survey (GSS) (Smith, Marsden, Hout, & Kim, 2013) and data provided by the Federal Bureau of Investigation and the United States Census. Because Secure Society Theory hypothesizes a negative relationship between societal security and religiosity, another way of framing this is that Secure Society Theory hypothesizes a positive relationship between societal *insecurity* and religiosity. This is how the hypothesized relationship was framed in the present study, given the variables available for analysis.

The present study tested three hypotheses. By including a measure of personal perception of insecurity, the present study tested the hypothesis that personal perceptions of insecurity are positively related to self-reported religiosity (Hypothesis 1). This personal measure was analyzed along with societal indicators of insecurity used in previous studies (crime rates and poverty rates) to test the hypothesis that both personal perceptions and societal indicators of insecurity uniquely predict self-reported religiosity (Hypothesis 2). Finally, religiosity and insecurity were analyzed at the national level over time to test the hypothesis that insecurity positively predicts future religiosity (Hypothesis 3).

Method

Data Sets and Variable Selection

To test the Secure Society Theory as it applies to religiosity in the United States, data were obtained from several data sets. These data sets, and the relevant variables from these data sets that are analyzed in the present study, are described below.

The General Social Survey. For variables regarding religiosity and one variable regarding perceptions of societal insecurity, the present study relied on data obtained

shackelf@oakland.edu

from the GSS. This sociological survey was conducted almost annually from 1972 to 1994 (excluding 1979, 1981, and 1992, due to funding limitations) and biennially from 1994 to the present (although data from 2014 were not yet available at the time of this study), by the University of Chicago's National Opinion Research Center (NORC), and provides a valuable source of time series data on American demographic characteristics and attitudes on a wide range of topics.

The GSS is a 90-min, in-house interview of a probability-based sample of noninstitutionalized US adults at least 18 years old. From 1972 to 1993, for each year the survey was conducted, the target sample size was 1500 participants, with actual sample sizes ranging from 1372 in 1990 to 1613 in 1972. Since 1994, the GSS has been administered to two samples each year the survey is conducted, each with a target size of 1500 participants. Total sample sizes range from 2765 in 2002 to 2992 in 1994. Aside from an oversampling of black participants in 1982 and 1987 (statistically controlled for in the present study's analyses), there has been no oversampling in other periods.

The 1972–2012 GSS data set consists of a total of 56,355 participants (25,804 men, 30,551 women) after correcting for an oversampling of black participants. The mean age of participants is 44.37 years (SD = 17.00, range = 18–89). The majority of participants (82.2 %) are identified as white (n = 46,328), with 6906 participants (12.3 %) identified as black and 3120 participants (5.5 %) identified as "other." The primary religious identifications of participants are as follows: 32,289 Protestant, 14,533 Catholic, 1111 Jewish, and 5994 "none." The residence of participants is coded into nine regions; this unfortunately reduces the amount of regional variability in religiosity and security that can be analyzed, but provides more detail than data at the national level.

The questions in the GSS are of two broad types: Participants either simply provide an answer and the interviewer is responsible for assigning the appropriate code to the response, or participants are given a hand card with a list of possible responses from which to choose. In 2002, the GSS switched from printed questionnaires to computer-assisted personal interviewing (CAPI), but hand cards are still provided to participants for relevant questions.

Measures of Religiosity. The following variables from the GSS were used as measures of religiosity for the present study. Although other variables related to religion are available in the GSS, the current variables were chosen because they were included in the greatest number of surveys from 1972 to 2012.

Religious Attendance. Participants were asked "How often do you attend religious services?" and responses were coded on a scale of 0 to 9 (0 = "never," 1 = "less than once a year," 2 = "about once or twice a year," 3 = "several times a year," 4 = "about once a month," 5 = "2–3 times a month," 6 = "nearly every week," 7 = "every week," 8 = "several times a week," 9 = "don't know" or no answer). Participants were not provided a hand card with these options, but interviewers were instructed to use these categories as probes if necessary. This question has been asked every year the GSS has been administered, resulting in 29 years for which data have been collected over a 40-year period (valid n = 55,821; 534 cases of "don't know" or no answer).

shackelf@oakland.edu

Participants tend to overstate the frequency of church attendance, in both the GSS and other surveys, such as Gallup polls (Hadaway, Marler, & Chaves, 1993). The principal investigator of the GSS has acknowledged this issue and explains it as a result of three factors: social desirability bias, telescoping, and participants relying on a broader interpretation of "attend[ing] religious service" (Smith, 1996). However, these data can still be used to analyze changes over time and relationships with other variables because the difference between reported attendance and actual attendance (which has been more accurately measured with time diary studies) has remained consistent in recent decades (Chaves, 2011).

Prayer Frequency. Participants were asked "About how often do you pray?" and responses were coded on a scale of 1 to 6 (1 = "several times a day," 2 = "once a day," 3 = "several times a week," 4 = "once a week," 5 = "less than once a week," 6 = "never"). Participants were not provided a hand card with these options, but interviewers were instructed to use these categories as probes if necessary. The GSS began asking this question in 1983, but it was not included in 1986 and 1991, resulting in 18 years for which data have been collected over a 29-year period (valid $n = 27{,}816$; 324 cases of "Don't know" or no answer).

These data were reverse-coded in the present study so that greater values indicate higher levels of prayer frequency. Additionally, for data from 1983, "Never" responses were collapsed with "Less than once a week." This was recommended by Smith (1988) due to an unusually high number of "Never" responses coded in that year. Although the wording of the question and instructions for interviewing and coding were not changed, it is possible that the discrepancy was due to inadequate interviewer training or inadequate care by the coding supervisor.

Biblical Fundamentalism. Participants were asked "Which of these statements comes closest to describing your feelings about the Bible?" and responses were coded on a scale of 1 to 4 (1 = "the Bible is the actual word of God and is to be taken literally, word for word"; 2 = "the Bible is the inspired word of God, but not everything in it should be taken literally, word for word"; 3 = "the Bible is an ancient book of fables, legends, history, and moral precepts recorded by men"; 4 = "Other" [volunteered]). The first three options were provided to participants on a hand card. The GSS began asking this question in 1984, but it was not included in 1986, resulting in 18 years for which data have been collected over a 28-year period (valid $n = 27{,}618$; 349 cases of "Other," 644 cases of "don't know" or no answer). These data were reverse-coded in the present study so that greater values indicate higher levels of fundamentalism. Responses of "Other" were excluded from analyses due to a lack of information about how to interpret these responses.

Strength of Affiliation. Participants who indicated any religious preference were asked "Would you call yourself a strong (preference named) or a not very strong (preference named)?" and responses were coded on a scale of 1 to 4 (1 = "strong," 2 = "not very strong," 3 = "somewhat strong" [volunteered], 4 = "no religion"). Participants were not provided a hand card with these options, but interviewers were instructed to use these categories as probes, if necessary. Interviewers were instructed to refer to the religious preference previously identified by the participant

shackelf@oakland.edu

when asking this question. If participants indicated that they follow no religion, this question was not asked. The GSS began asking this question in 1974, and it has been asked every year since, resulting in 27 years for which data have been collected over a 38-year period (valid $n = 51{,}436$; 1797 cases of "don't know" or no answer). These data were reverse-coded in the present study so that greater values indicate a greater strength of affiliation. Also, the order of the "not very strong" and "somewhat strong" categories were switched to more accurately reflect a scale of increasing religiosity. Therefore, the final coding as used in the present study is as follows: 1 = "no religion," 2 = "not very strong," 3 = "somewhat strong," and 4 = "strong."

Religiosity Component. To facilitate analyzing the effects of several variables on religiosity, the present study considered whether religious attendance, prayer frequency, biblical fundamentalism, and strength of affiliation could be combined into a single religiosity composite variable. A principal components analysis was conducted on these variables, with extraction based on eigenvalues greater than 1. The number of factors to extract was not fixed beforehand. Two measures of sampling adequacy, Kaiser–Myer–Olkin (KMO) and Bartlett's test of sphericity, indicated that it was appropriate to proceed with principal components analysis, as KMO was greater than 0.5 (KMO = 0.760) and Bartlett's test of sphericity was significant ($p < 0.001$).

The four religiosity variables are all correlated. A principal components analysis yielded a single component with an eigenvalue of 2.44, explaining 61.04 % of the variance. The second largest eigenvalue was 0.69 and, therefore, was not extracted. The principal component communalities were 0.690 for religious attendance, 0.629 for prayer frequency, 0.436 for biblical fundamentalism, and 0.687 for strength of affiliation. Thus, the majority of the variance in these variables is accounted for by a one-component solution, although variance in biblical fundamentalism is not accounted for as strongly as the other variables. Nevertheless, it is reasonable to consider these four variables as part of a single religiosity component. Therefore, a single religiosity variable was constructed by first calculating z-scores for each of the four GSS variables (since their scales of measurement are not uniform) and then calculating the mean of these z-scores. The four standardized variables exhibited strong internal consistency (Cronbach's $\alpha = 0.79$). For each valid case in the GSS data set, a single "religiosity" value was calculated. Unless otherwise stated, all subsequent analyses rely on this religiosity composite variable.

The number of original religiosity variables used to create the values for the composite variable varies by year, since some surveys from 1972 to 2012 only include a subset of the four variables. As a result, the religiosity composite variable is most strongly influenced by religious attendance ($n = 55{,}821$), followed by strength of affiliation ($n = 51{,}436$), prayer frequency ($n = 27{,}816$), and biblical fundamentalism ($n = 27{,}618$). This may also partially explain why biblical fundamentalism has the least amount of variance accounted for by the one-component solution.

Measure of Societal Insecurity. One variable from the GSS was used to assess participants' perceptions of societal security. Specifically, participants were asked "Is there any area right around here—that is, within a mile—where you would be afraid

shackelf@oakland.edu

to walk alone at night?" and responses were simply "yes" (coded as 1) or "no" (coded as 2). The GSS began asking this question in 1973, but it was not included in 1975, 1978, 1983, and 1986, resulting in 24 years for which data were collected over a 39-year period (valid $n = 33{,}652$; 253 cases of "Don't know" or no answer). The data for this "fear" variable were reverse-coded in the present study to match the other variables assessing societal insecurity (described below), with higher values indicating a greater degree of insecurity.

Control Variables. The following control variables were included in a subset of the present study's analyses because of their relationships with religiosity: sex, race, and age. These relationships were tested in the present study through a series of preliminary analyses (performed using SPSS version 21), which are described below.

Several studies indicate that women are, on average, more religious than men; women express a greater interest in religion (Sasaki, 1979), are more strongly committed to their religions (Bensen, Donahue, & Erickson, 1989), and report more frequent religious attendance (Batson, Schoenrade, & Ventis, 1993). These trends are consistent across denomination and type of religious belief system (Stark & Bainbridge, 1985). This relationship is also found in the 1972–2012 GSS data. An analysis of sex and the religiosity composite variable indicated that women are more religious ($M = 0.14$, SD = 0.82) than men [($M = -0.15$, SD = 0.84), $t(56{,}188) = -40.52$, $p < 0.001$, mean difference = −0.29, 95 % CI [−0.30, −0.27], Cohen's $d = 0.34$].

Although the relationship between race and religiosity has not been examined as extensively as sex and religiosity, researchers have found evidence of differences between African Americans and white Americans, with African Americans exhibiting greater religiosity (see Levin, Taylor, & Chatters, 1994). This relationship was tested in the 1972–2012 GSS data. A one-way Analysis of Variance indicated a difference in religiosity between races, $F(2, 56{,}187) = 437.75$, $p < 0.001$. However, the effect size is very small, $\eta^2 = 0.02$. Nevertheless, a post hoc Tukey test indicated that African Americans reported higher religiosity ($M = 0.28$, SD = 0.77, 95 % CI [0.27, 0.30]) than both white Americans ($M = -0.03$, SD = 0.85, 95 % CI [−0.04, −0.02]) and nonwhite others ($M = -0.07$, SD = 0.81, 95 % CI [−0.10, −0.04]).

Finally, several studies have found a relationship between age and religiosity in the United States (Bahr, 1970; Chaves, 1991; Firebaugh & Harley, 1991; Hout & Greeley, 1990), although there is debate as to what is driving this relationship (see Argue, Johnson, & White, 1999). An analysis of the 1972–2012 GSS data indicated a correlation between age and religiosity, $r(56{,}040) = 0.18$, $p < 0.001$. However, as with race, the effect size is quite small, $r^2 = 0.03$.

Uniform Crime Reports. In addition to the fear variable from the GSS, societal insecurity was assessed by measuring regional crime rates, as the frequency of crime in one's surrounding area arguably influences one's perception of societal insecurity. These crime rates were obtained from *Crime in the United States* (CIUS), an annual report published as part of the Federal Bureau of Investigation's Uniform Crime Report program (UCR; FBI, 2014). Like the variables from the GSS, these data were ideal for the purposes of the current study because of the availability of data

shackelf@oakland.edu

over a substantial time period, as well as the consistency of measurement over time. These data are available as early as 1930, but the present study only used data corresponding to the time period for which religiosity data from the GSS are available (i.e., 1972–2012).

The UCR divides reported crimes into two broad categories: violent crimes and property crimes. Violent crimes are defined in the UCR as offenses involving force or the threat of force, and this category consists of murder and nonnegligent manslaughter, forcible rape, robbery, and aggravated assault. Property crime consists of burglary, larceny–theft, motor vehicle theft, and arson, specifically when there is no force or threat of force against the victims. However, statistics regarding arson are not included in the UCR summary data of property crime because of limited participation by local law enforcement agencies, as well as variance in data collection procedures by agencies that do participate. Although the CIUS reports provide data on each type of crime listed above (except arson), the present study relied on data from the broader categories of violent crime and property crime.

To control for varying population sizes between states, the current study relied on crime rate data reported in CIUS, rather than the absolute number of crimes reported. For each state and each year, CIUS provides a violent crime rate and property crime rate that is calculated as the total number of crimes reported in each category divided by the total population. These data are presented as the rate per 100,000 inhabitants.

These data sets do not provide a perfect record of how many crimes are committed, as they are limited to those crimes discovered by or reported to law enforcement agencies. Therefore, changes in the values reported over time only partially describe changes in the actual frequency of crimes committed.

Data from the GSS are not provided by state, but rather by region. Therefore, prior to analysis, state-level data on violent crime and property crime were converted to region-level data by calculating means for the states corresponding to each region (e.g., for each year, violent crime data from New York, New Jersey, and Pennsylvania were converted to a single mean value for the "Middle Atlantic" region). These converted data were then added to the GSS data set.

Current Population Survey. Data regarding the percentage of the United States population in poverty, by state and by year, were obtained from The Annual Social and Economic Supplement (ASEC; United States Census Bureau, 2013) of the Current Population Survey (CPS). Poverty is calculated based on a series of dollar value thresholds, and these thresholds are determined by family size and the number of resident children under 18 years old (DeNavas-Walt, Proctor, & Smith, 2013). To be categorized as "in poverty," the family's total income (before taxes and tax credits) must be less than the applicable threshold, which is updated annually based on the Consumer Price Index to account for inflation. If the family's total income is below the threshold, then every individual in the family is considered to be in poverty.

The ASEC provides data from 1980 to 2012 on the number of people in poverty in each state and the poverty rate for each state (i.e., the number of poor divided by the state's population). The poverty rate is presented as the rate per 100,000 inhabitants.

shackelf@oakland.edu

As with the crime data, the poverty rates were converted to region-level data by calculating mean rates for the states corresponding to each GSS region. These converted data were then added to the GSS data set.

Time Series Data Set. To test Hypothesis 3 of the present study, regarding societal insecurity predicting *future* levels of religiosity, a separate data set was constructed using several of the variables described above. In this data set, each "case" was a year, resulting in 41 cases representing the years 1972–2012. Violent crime rate, property crime rate, and poverty rate data were added to the data set by calculating the mean of the state values for each year. The neighborhood fear variable from the GSS was also added to this data set by calculating the mean responses for each year in the 1972–2012 GSS data set. The religiosity composite variable was added to this data set by calculating the mean values for each year in the 1972–2012 GSS data set.

Given the already limited number of cases available in this data set for time series analysis, the fact that there are several years in which the GSS has not been administered is a considerable problem. Therefore, missing values of neighborhood fear and religiosity in the time series data set were replaced by calculating the median of the nearest value above and below the missing value. For example, a religiosity value of 0.0457 was created for the year 1992 by calculating the median of the existing values from 1991 to 1993. This method of missing value replacement was chosen based on the assumption that values for these variables are unlikely to change substantially in the short term, from year to year, and therefore the missing value for any given year can be reasonably estimated as falling in between the values of neighboring years.

Results

Religiosity and Individual Perceptions of Societal Insecurity

To test Hypothesis 1, that individuals' perceptions of societal insecurity are positively related to their reported religiosity, an independent sample *t*-test was conducted (using SPSS 21) to determine the effect of neighborhood fear (i.e., one's fear of going outside at night in their neighborhood) on the religiosity composite variable. This analysis was conducted to take advantage of the greatest number of cases in the GSS data set, as the subsequent regression analyses described below, which use several variables, are limited by the number of cases with valid data for all variables. The *t*-test was significant, $t(33{,}554) = -15.57$, $p < 0.001$, mean difference $= -0.15$, 95 % CI [−0.16, −0.13], Cohen's $d = 0.17$. Participants who indicated they are afraid to walk around their neighborhood at night had higher religiosity scores ($M = 0.09$, SD = 0.83) than those who were not afraid ($M = -0.05$, SD = 0.83). In other words, as hypothesized, higher perceptions of societal insecurity are related to higher levels of religiosity.

shackelf@oakland.edu

To further test Hypothesis 1, a series of chi-square tests of independence was conducted using SPSS 21 to investigate whether the relationship between the religiosity composite and fear would obtain for each of the four original religiosity variables from the GSS (religious attendance, prayer frequency, biblical fundamentalism, and strength of affiliation).

To facilitate the analysis of religious attendance and interpretation of the chi-square test results, the attendance categories "never," "less than once a year," and "about once or twice a year" were collapsed into a "low attendance" category, and the categories "nearly every week," "every week," and "several times a week" were collapsed into a "high attendance" category. Other categories of attendance were not included in the chi-square test, as they arguably represent "moderate" levels of religiosity.

The chi-square test of independence for religious attendance was significant, $X^2(1,\ N=23{,}681)=101.51,\ p<0.001$. However, the effect size, as measured by Cramer's V (φ_c), was quite small, $\varphi_c=0.07$. Those who indicated that they *are* afraid to walk around their neighborhood at night were more likely than expected to report high levels of religious attendance ($O=4747$, $E=4350$; std. residual = 5.7), and they were less likely than expected to report low levels of religious attendance ($O=4496$, $E=4873$; std. residual = −5.4). Those who indicated that they are *not* afraid were more likely than expected to report low levels of religious attendance ($O=8017$, $E=7640$; std. residual = 4.3), and they were less likely than expected to report high levels of religious attendance ($O=6441$, $E=6818$; std. residual = −4.6). In short, security and religiosity are not independent of each other, and as hypothesized, higher perceptions of societal insecurity are related to higher levels of religious attendance.

To facilitate the analysis of prayer frequency and interpretation of the chi-square test results, the categories of "never," "less than once a week," and "once a week" were collapsed into a "low prayer frequency" category, and the categories of "several times a week," "once a day," and "several times a day" were collapsed into a "high prayer frequency" category.

The chi-square test of independence for prayer frequency was significant, $X^2(1,\ N=17{,}835)=157.58,\ p<0.001$. Again, the effect size was small, $\varphi_c=0.09$. Those who indicated that they *are* afraid to walk around their neighborhood at night were more likely than expected to report high prayer frequency ($O=4939$, $E=4567$; std. residual = 5.5), and they were less likely than expected to report low prayer frequency ($O=1626$, $E=1998$; std. residual = −8.3). Those who indicated that they are *not* afraid were more likely than expected to report low prayer frequency ($O=3802$, $E=3430$; std. residual = 6.4), and they were less likely than expected to report high prayer frequency ($O=7468$, $E=7840$; std. residual = −4.2). In short, as hypothesized, higher perceptions of societal insecurity are related to higher prayer frequency.

Regarding biblical fundamentalism, it was not necessary to collapse any categories because responses were recorded in three categories: belief that the Bible is (1) the literal word of God, (2) the inspired word of God, or (3) a book of fables. The chi-square test of independence for biblical fundamentalism was significant, $X^2(2,\ N=17{,}550)=78.89,\ p<0.001$. The effect size was small, $\varphi_c=0.07$. Those who indicated

shackelf@oakland.edu

that they *are* afraid to walk around their neighborhood at night were more likely than expected to view the Bible as the literal word of God ($O = 2462$, $E = 2204$; std. residual = 5.5) and less likely than expected to view the Bible as a book of fables ($O = 972$, $E = 1095$; std. residual = −3.7). Those who indicated that they are *not* afraid to walk around their neighborhood at night were more likely than expected to view the Bible as a book of fables ($O = 1975$, $E = 1852$; std. residual = 2.9) and less likely than expected to view the Bible as the literal word of God ($O = 3469$, $E = 3727$; std, residual = −4.2). The "inspired word of God" category showed a significant difference for the "afraid" group, but no significant difference for the "unafraid" group. It is difficult to interpret the effects for this category because it arguably represents a moderate level of religiosity, but for the less ambiguous categories, as hypothesized, higher perceptions of societal insecurity are related to greater biblical fundamentalism.

To facilitate the analysis of strength of religious affiliation, the categories of "no religion" and "not very strong" were collapsed into a "low strength" category, and the "strong" category was unchanged and used as the "high strength" category. The "somewhat strong" category was not included in this analysis, as it represents a vague, volunteered response, and few participants belong to this category in the first place.

The chi-square test of independence for strength of affiliation was significant, $X^2(1, N = 27{,}787) = 155.66$, $p < 0.001$. The effect size was small, $\varphi_c = 0.07$. Those who indicated that they *are* afraid to walk around their neighborhood at night were more likely than expected to report strong religious affiliations ($O = 5110$, $E = 4608$; std. residual = 7.4), and they were less likely than expected to report weak religious affiliations ($O = 5766$, $E = 6268$; std. residual = −6.3). Those who indicated that they were *not* afraid were more likely than expected to report weak religious affiliations ($O = 10{,}247$, $E = 9745$; std. residual = 5.1) and less likely than expected to report strong religious affiliations ($O = 6664$, $E = 7166$; std. residual = −5.9). In short, as hypothesized, higher perceptions of societal insecurity are related to stronger religious affiliations.

Does Societal Insecurity, at the Individual and Regional Level, Predict Religiosity?

To test Hypothesis 2, that both individual perceptions of societal insecurity and regional factors indicative of societal insecurity positively predict religiosity, multiple linear regression analyses were conducted using SPSS 21. Two regression models were computed to determine whether regional societal insecurity as measured by violent crime rate, property crime rate, and poverty rate, along with individual perceptions of societal insecurity as measured by the GSS fear variable, predict religiosity when controlling for age, sex, and race (dummy coded with "African American" and "other" entered into the model, and "white" omitted to serve as the reference category). Two models were computed to avoid issues of

shackelf@oakland.edu

multicollinearity because of the large correlation between violent crime rate and property crime rate, $r(27,526) = 0.58$, $p < 0.001$.

Model 1 excluded property crime rate. The variables were entered into the model in two blocks: The first block (the partial model) included the control variables, and the second block (the full model) introduced the societal insecurity variables. This allowed for determining whether the inclusion of the societal insecurity variables contributed to the model fit. The partial model predicted and explained roughly 8 % of the variance in religiosity, adjusted $R^2 = 0.081$, SE = 0.79, $F(4, 27,523) = 607.66$, $p < 0.001$. All of the control variables contributed to the model, and their coefficients were all positive, which is consistent with the previous literature and preliminary analyses described in the Method section. The standardized coefficients (β) indicated that age was the strongest predictor ($\beta = 0.185$, $p < 0.001$). The next strongest predictor was sex ($\beta = 0.168$, $p < 0.001$), followed by the African American racial category ($\beta = 0.133$, $p < 0.001$), and, finally, the "other" racial category ($\beta = 0.012$, $p = 0.046$).

The full model was also significant and explained about 10 % of the variance in religiosity, adjusted $R^2 = 0.097$, SE = 0.78, $F(7, 27,520) = 425.57$, $p < 0.001$. Using a criterion of VIF values less than 10, there were no apparent issues of multicollinearity in the full model, as all VIFs were less than 1.2. The change in explained variance was significant, $\Delta R^2 = 0.017$, $\Delta F(3, 27,520) = 168.04$, $p < 0.001$, indicating that the collection of societal insecurity variables contributed to the model beyond the control variables. The control variables all remained significant predictors in the full model, and although their coefficients changed slightly, their relative strengths as predictors remained the same. Aside from the "other" racial category ($\beta = 0.016$, $p = 0.005$), all of the control variables were stronger predictors of religiosity than the societal insecurity variables. Among the societal insecurity variables, poverty rate was the strongest predictor ($\beta = 0.119$, $p < 0.001$), followed by violent crime rate ($\beta = 0.027$, $p < 0.001$) and neighborhood fear ($\beta = 0.013$, $p = 0.027$). As hypothesized, all of the coefficients for the societal insecurity variables were positive, indicating that as both individual perceptions of societal insecurity and regional factors indicative of societal insecurity increase, religiosity increases.

Model 2 excluded violent crime rate. The variables were again entered into the model in two blocks, following the same procedure as Model 1. As block 1 was identical to that of Model 1, the results of the partial model were identical and are not repeated here. The full model was significant and, like Model 1, explained about 10 % of the variance in religiosity, adjusted $R^2 = 0.098$, SE = 0.78, $F(7, 27,520) = 425.96$, $p < 0.001$. Using a criterion of VIF values less than 10, there were no apparent issues of multicollinearity in the full model, as all VIFs were less than 1.2. The change in explained variance was significant, $\Delta R^2 = 0.017$, $\Delta F(3, 27,520) = 168.67$, $p < 0.001$. The control variables all remained significant predictors in the full model, and although their coefficients changed slightly, their relative strengths as predictors remained the same. Aside from the "other" racial category ($\beta = 0.017$, $p = 0.003$), all of the control variables were stronger predictors of religiosity than the societal insecurity variables. Among the societal insecurity variables, poverty rate was again the strongest

shackelf@oakland.edu

predictor ($\beta = 0.122$, $p < 0.001$), followed by property crime rate ($\beta = 0.028$, $p < 0.001$) and neighborhood fear ($\beta = 0.013$, $p = 0.035$). As hypothesized, all of the coefficients for the societal insecurity variables were positive, again indicating that as both individual perceptions of societal insecurity and regional factors indicative of societal insecurity increase, religiosity increases.

Does Societal Insecurity Predict Future Religiosity?

To test Hypothesis 3, that both perceptions of societal insecurity and factors indicative of societal insecurity can predict *future* levels of religiosity at the national level, a time series analysis was conducted. More specifically, this hypothesis was tested through the use of vector autoregression (VAR; Sims, 1980)—following the methodology of Solt et al. (2011)—using the statistical software program STATA 12.1.

The VAR equation is similar to that used in multiple linear regression. However, the VAR model consists of n equations, where n is the number of variables being investigated (Stock & Watson, 2001). In each equation, one of the variables takes a turn as the criterion and is predicted by lagged values of itself, all other variables, and a serially uncorrelated error term. Each equation is then estimated by ordinary least squares (OLS) regression. This is referred to as a standard, "reduced-form" VAR, which is the type of VAR used in the present study.

For these analyses, violent crime rate and property crime rate were combined into a "total crime rate" variable to avoid issues of multicollinearity, as running separate models with each type of crime is less feasible in this case because of the number of equations involved in a single model. The total crime rate variable was constructed by first standardizing the violent crime and property crime data in the time series data set into z-scores. The total crime variable was then computed as the mean of the two standardized variables.

Model selection criteria are used to determine the best lag length for the VAR model. The most common criteria used are the Akaike (AIC), Schwarz–Bayesian (BIC), and Hannan–Quinn (HQ) (Zivot & Wang, 2003). These criteria are produced in STATA after specifying the variables to be included in the VAR and inputting a set number of potential lags. The recommended procedure is to choose the lag for which the selection criteria values are minimized (Lütkepohl, 2005). For the present study, although a lag of 8 years was found to best minimize the selection criteria (AIC = 3.55, HQIC = 3.82, SBIC = 4.52), this model failed to pass the Lagrange multiplier (LM) test for autocorrelation, $X^2(16, n = 25) = 36.28$, $p = 0.003$. Despite collapsing violent and property crime into a single variable, there was still an issue of multicollinearity. The LM test output indicated that a lag of 10 years should be chosen. Although the selection criteria values were slightly larger (AIC = 3.99, HQIC = 4.24, SBIC = 4.98), this model passed the LM test, $X^2(16, n = 23) = 17.82$, $p = 0.33$.

shackelf@oakland.edu

These analyses relied on a separate time series data set, with 41 cases representing the years 1972–2012. However, with a lag of 10 years and poverty data only going back to 1980, the VAR model was limited to analyzing 23 cases, from 1990 to 2012. The variables entered into the VAR equations were total crime rate, poverty rate, neighborhood fear, and the religiosity composite. In this data set, these variables reflect values at the national level, as the values are means of state- or region-level data for each year. The VAR model can be illustrated as follows:

$$\begin{aligned}\text{Religiosity}_t = {} & a_{10} + a_{11}\text{Religiosity}_{t-10} + a_{12}\text{TotalCrime}_{t-10} + a_{13}\text{Poverty}_{t-10} \\ & + a_{14}\text{Fear}_{t-10} + e_{1t}\end{aligned}$$

$$\begin{aligned}\text{TotalCrime}_t = {} & a_{20} + a_{21}\text{Religiosity}_{t-10} + a_{22}\text{TotalCrime}_{t-10} + a_{23}\text{Poverty}_{t-10} \\ & + a_{24}\text{Fear}_{t-10} + e_{2t}\end{aligned}$$

$$\begin{aligned}\text{Poverty}_t = {} & a_{30} + a_{31}\text{Religiosity}_{t-10} + a_{32}\text{TotalCrime}_{t-10} + a_{33}\text{Poverty}_{t-10} \\ & + a_{34}\text{Fear}_{t-10} + e_{3t}\end{aligned}$$

$$\begin{aligned}\text{Fear}_t = {} & a_{40} + a_{41}\text{Religiosity}_{t-10} + a_{42}\text{TotalCrime}_{t-10} + a_{43}\text{Poverty}_{t-10} \\ & + a_{44}\text{Fear}_{t-10} + e_{4t}\end{aligned}$$

Because STATA does not provide standardized coefficients as part of the output for VAR analyses, prior to running the VAR model, the poverty rate, fear, and religiosity variables were transformed into *z*-scores. Thus, the unstandardized coefficients provided by STATA can be interpreted in the same way as standardized coefficients (i.e., how many *SD*s the criterion changes for every 1 *SD* change in the predictor). The total crime rate variable was not transformed as it already represents the mean values of two standardized variables and thus approximates a standardized variable itself ($M = 0.00$, $\text{SD} = 0.97$).

The results of the VAR model indicated that all four equations were significant predictors of their respective criteria, explaining 76 % of the variance in religiosity, 73 % of the variance in total crime, 72 % of the variance in poverty rate, and 55 % of the variance in neighborhood fear (all $ps < 0.001$). The primary portion of the model to consider for the present study is the equation for predicting religiosity. For this equation, the strongest predictor was neighborhood fear ($\beta = 0.46$, $p = 0.003$), followed by poverty rate ($\beta = 0.30$, $p = 0.03$). Both coefficients were positive, indicating that, as hypothesized, increases in fear and poverty predict an increase in future religiosity (specifically, religiosity 10 years later). However, counter to expectations, total crime rate did not predict future religiosity. Thus, the results provide only partial support for Hypothesis 3.

It is also worth noting from the other equations that religiosity is a significant predictor of future increases in crime ($\beta = 0.58$, $p = 0.001$) and future increases in neighborhood fear ($\beta = 0.89$, $p = 0.003$). These results were not hypothesized, yet they suggest the possibility of a feedback loop between religiosity and societal security, which is considered in more detail in the Discussion section.

shackelf@oakland.edu

Discussion

The present study tested three hypotheses derived from Norris and Inglehart's (2004) Secure Society Theory of religiosity, with the goal of explaining variations in religiosity within the United States at both the individual and societal level. Hypothesis 1 was supported; individual perceptions of societal insecurity, as measured by reporting whether one was afraid to walk around their neighborhood at night, were positively related to religiosity. Individuals who reported that they were afraid were more religious than those reporting they were not afraid, as indicated by a religiosity composite score comprising religious attendance, prayer frequency, biblical fundamentalism, and strength of religious affiliation. This relationship remained when examining each religiosity measure separately. Hypothesis 2 was also supported, providing further evidence for the positive relationship between societal insecurity and religiosity. Neighborhood fear, violent crime rate, property crime rate, and poverty rate all predicted religiosity, even after controlling for the effects of each other and the variables of sex, age, and race.

Hypothesis 3 was partially supported, in that neighborhood fear and poverty rate (but not crime rate) predicted levels of religiosity 10 years later. As with the previous analyses, there was a positive relationship, indicating that increases in societal insecurity predict increases in future religiosity. Interestingly, religiosity was also positively related to future crime rate and neighborhood fear. These results were not predicted by Secure Society Theory, and it is unclear why increased religiosity would lead to increases in actual societal insecurity (i.e., higher crime rates). However, the relationship between religiosity and perceived societal insecurity (i.e., neighborhood fear) is more readily interpretable as indicative of a feedback loop, in which religiosity is both influenced by perceptions of insecurity and influences those perceptions. In other words, it may be that not only are people who perceive greater societal insecurity likely to be more religious, but people who are more religious are likely to perceive greater insecurity.

Given the overall support found for the hypotheses in the present study, what conclusions can be reached regarding the relationship between societal insecurity and religiosity? Overall, the apparent relationship is consistent with the Secure Society Theory. Perceptions of insecurity, whether assessed directly (through neighborhood fear) or indirectly (through regional crime rates and poverty rates, which are likely to influence the perceptions of people living in those regions), are positively related to religiosity: The less secure one perceives society to be, the more religious they are, now and 10 years later.

However, the present results do not allow one to make strong inferences regarding causality. The chi-square tests of independence indicated that religiosity and perceptions of societal insecurity are not independent of each other, but this does not mean that a causal relationship necessarily exists. Likewise, the multiple regression analyses indicated that societal insecurity variables uniquely predict religiosity, but their predictive power is merely an indication of the relationship between the variables, not an indication that the predictors *cause* changes in religiosity. The time

shackelf@oakland.edu

series analysis provides the strongest evidence for causality in the present study by showing that some aspects of insecurity predict future levels of religiosity, but technically the results only indicate that the variables are related, as with the other multiple regression analyses performed earlier.

From an evolutionary perspective, the present study's results indicate that the mechanisms (or a subset of the mechanisms) associated with religious beliefs and behaviors are sensitive to environmental input regarding societal insecurity, and these mechanisms respond to increased insecurity by strengthening religiosity. This is consistent with Norris and Inglehart's (2004) proposed function of religion as a coping mechanism in the face of societal insecurity. However, this does not necessarily support the idea that religion is produced by a specialized psychological adaptation. A possible alternative explanation is that there are mechanisms designed to cope with unpredictable and unsafe environments by increasing one's perception of control, as this may serve as a buffer against helplessness or negative affect (Case et al., 2004). This may occur regardless of the existence of religious beliefs; indeed, experimental manipulations of perceived control have been found to elicit increased superstitious behavior (Keinan, 2002) and illusory pattern perception (Whitson & Galinsky, 2008). However, given the existence of religious beliefs, particularly beliefs that provide meaning for events or circumstances that may otherwise seem unpredictable, religion in this context may be best viewed as an "exaptation," serving as a form of compensatory control (Kay et al., 2010) even though religious beliefs and behaviors were not necessarily originally selected for this purpose.

This interpretation leads to the question: Why would the *illusory* perception of increased control be adaptive? Norris and Inglehart (2004) suggest that perceptions of control are useful in unpredictable and unsafe environments because "Individuals experiencing stress have a need for rigid, predictable rules. They need to be sure of what is going to happen because they are in danger—their margin for error is slender and they need maximum predictability" (p. 19). This does not explain what is going on, however, because the superstitions, perceptions of illusory patterns, and religious beliefs do not provide "maximum predictability." They provide the illusion of increased control without *actually* increasing one's control over events. If one's margin for error is slim, it seems maladaptive to invent meaning behind events just to reduce negative affect. After all, natural selection does not care how happy or sad you are, and these feelings are only useful to the extent to which they motivate solutions to adaptive problems.

When dealing with threats in the natural environment, the most accurate interpretation of one's control would be favored, as this would allow one to allocate limited resources appropriately. However, Kurzban (2010) offers an interesting evolutionary explanation for the phenomena described above, which are examples of what he calls "strategic ignorance." The key to explaining why inaccurate information may sometimes be favored is the fact that humans deal with more than just the natural environment: We are an extremely social species, and we have spent much of our evolutionary history living in small groups in which we depended on each other for survival. It is important that others view you as being a valuable member of the group; therefore, it is important to persuade others that you are more valuable than

shackelf@oakland.edu

you actually are. Rather than consciously lying, an effective method of persuasion is to believe the inaccurate information yourself. Although this explanation is speculative, there are a variety of phenomena that it can potentially explain (see Kurzban, 2010), including the illusory perception of control in unpredictable and unsafe environments.

Limitations of the Present Study

Although it is not technically a *limitation* of the present study, this section begins with addressing the small effect sizes obtained from the tests of Hypotheses 1 and 2, because these small effects may be partially explained by some of the limitations of this study. For example, one of the weaknesses of using archival data is that the data were collected without the present study's hypotheses in mind. As a result, the variables used in the present study regarding societal insecurity were not ideal. The question assessing insecurity in the GSS (neighborhood fear) was very specific and, arguably, only assesses a small portion of overall perception of societal insecurity. Therefore, the effect of neighborhood fear on religiosity should not be misconstrued as the effect of *overall* perceptions of societal insecurity on religiosity. More comprehensive measures of perceived societal insecurity would provide a clearer picture of how this perception is related to religiosity, and would possibly result in a larger effect size.

Another limitation that may have reduced the effect sizes is the lack of state-level data from the GSS. Because the GSS participants were only identified as belonging to one of nine regions, the state-level data on crime and poverty rates obtained elsewhere had to be converted to the same nine regions for analysis. To the extent that crime and poverty affect one's perceptions of insecurity, it is likely that the effect weakens the further away the crime and poverty are from the individual. Because some of the GSS regions contain as many as eight states, one should expect the crime and poverty data to only partially represent one's perception of insecurity in their local environment. It is therefore encouraging that, despite this lack of precision, the insecurity variables in the present study were all significantly related to religiosity in the hypothesized direction, and it is likely that data more specific to one's local environment would produce a larger effect size.

Although the effect sizes identified in the time series analysis were larger, there was a limitation to this portion of the study as well, namely, the restricted number of observations. Although the data used in the present study span several decades, the time lag of 10 years used in the vector autoregression model resulted in 23 observations. Therefore, the results of these analyses should be interpreted with caution, as they may not accurately describe the relationship between religiosity and insecurity over time. The results were promising in that they were, for the most part, in the hypothesized direction, but studies analyzing this relationship over a longer period of time are necessary before making any strong conclusions.

shackelf@oakland.edu

Future Directions

The results of the present study support the hypothesis that perceptions of societal insecurity influence religiosity, which serves as a useful foundation for future studies to investigate the possible causal nature of this relationship. Several possibilities for building on the present results are described below.

The present study relied on measures that arguably assess one's perceptions of societal insecurity. An important next step for future studies would be to develop a more direct measure of this perception. By constructing a scale that is specifically focused on measuring perceptions of societal insecurity, researchers could assess the relationship between these perceptions and religiosity more accurately. Additionally, assuming such a scale is constructed and validated, researchers could design experiments to test whether these perceptions *cause* changes in religiosity. For example, researchers could prime participants to have an increased perception of insecurity, perhaps by exposing them to either real or fabricated news stories regarding crime in their local environment. The effectiveness of the prime could be assessed by measuring differences between groups on the societal insecurity scale, and researchers could then investigate whether such primes lead to an increase in reported religiosity relative to participants who are primed to have a decreased perception of insecurity.

Similarly, given the possible feedback loop between religiosity and perceptions of insecurity suggested by the results of the time series analysis, researchers could prime participants' religiosity and subsequently have them complete the societal insecurity scale. This would allow one to test for causality in the opposite direction—i.e., to test for an increase in perceived insecurity relative to participants who are not primed with religious ideas and concepts.

Finally, future studies can investigate the possible benefits of increased religiosity as a result of increased societal insecurity. Assuming an effect on religiosity is found when priming societal insecurity, one could investigate whether this effect is strengthened in a group setting. If Kurzban's (2010) argument regarding strategic ignorance applies to this phenomenon (i.e., increased religiosity leads to increased perceptions of control, which makes one appear more valuable to the group), one could hypothesize that the presence of others may strengthen this effect. This line of reasoning could also be extended to investigating effects in the context of cooperative/competitive games; perhaps people are more willing to cooperate with individuals who react more strongly with compensatory control in the face of insecurity.

Conclusion

This chapter presented evidence consistent with the Secure Society Theory of religiosity. There appears to be a relationship between societal insecurity and religiosity in the United States over the last 40 years. Although the nature of this relationship is still unclear, the present study builds on previous research by illustrating that this

shackelf@oakland.edu

relationship is not driven solely by societal conditions, but by individuals' *perceptions* of societal insecurity. Given the number of studies that now reliably demonstrate a relationship between societal indicators of insecurity and religiosity, it is important to move forward by more directly assessing perceptions of societal insecurity at the individual level. The present study serves as a useful foundation for moving research forward in this way. Although religiosity is influenced by a variety of factors, further investigation of perceptions of societal insecurity may provide insights regarding variations in religiosity at the individual and societal level.

Acknowledgment This chapter is based on Liddle (2015).

References

Argue, A., Johnson, D. R., & White, L. K. (1999). Age and religiosity: Evidence from a three-wave panel analysis. *Journal for the Scientific Study of Religion, 38*, 423–435.

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

Bahr, H. (1970). Aging and religious disaffiliation. *Social Forces, 49*, 60–71.

Barber, N. (2011). A cross-national test of the uncertainty hypothesis of religious belief. *Cross-Cultural Research, 45*, 318–333.

Batson, C. D., Schoenrade, P., & Ventis, W. L. (1993). *Religion and the individual: A social-psychological perspective*. New York: Oxford University Press.

Bensen, P. L., Donahue, M. J., & Erickson, J. A. (1989). Adolescence and religion: A review of the literature from 1970–1986. *Research in the Social Scientific Study of Religion, 1*, 153–181.

Bjorklund, D. F., & Pellegrini, A. D. (2000). Child development and evolutionary psychology. *Child Development, 71*, 1687–1708.

Burger, J. M., & Lynn, A. L. (2005). Superstitious behavior among American and Japanese professional baseball players. *Basic and Applied Social Psychology, 27*, 71–76.

Buss, D. M. (2011). *Evolutionary psychology: The new science of the mind* (4th ed.). Boston: Pearson.

Case, T. I., Fitness, J., Cairns, D. R., & Stevenson, R. J. (2004). Coping with uncertainty: Superstitious strategies and secondary control. *Journal of Applied Social Psychology, 34*, 848–871.

Chaves, M. (1991). Family structure and protestant church attendance: The sociological basis of cohort and age effects. *Journal for the Scientific Study of Religion, 30*, 501–514.

Chaves, M. (2011). *American religion: Contemporary trends*. Princeton, NJ: Princeton University Press.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason Selection Task. *Cognition, 31*, 187–276.

DeNavas-Walt, C., Proctor, B. D., & Smith, J. C. (2013). *Income, poverty, and health insurance coverage in the United States, 2012 (Current population reports, P60-245)*. Washington, DC: U.S. Census Bureau.

Fincher, C. L., & Thornhill, R. (2008). Assortative sociality, limited dispersal, infectious disease and the genesis of the global pattern of religious diversity. *Proceedings of the Royal Society of London, 275*, 2587–2594.

Firebaugh, G., & Harley, B. (1991). Trends in U.S. church attendance: Secularization and revival, or merely life cycle effects? *Journal for the Scientific Study of Religion, 30*, 487–500.

Frijters, P., & Barón, J. D. (2010, April). The cult of Theoi: Economic uncertainty and religion. *IZA Discussion Paper No. 4902*.

shackelf@oakland.edu

Grant, J.T. (2008). Measuring aggregate religiosity in the United States, 1952–2005. *Sociological Spectrum, 28*, 460–476.
Hadaway, C. K., Marler, P. L., & Chaves, M. (1993). What the polls don't show: A closer look at U.S. church attendance. *American Sociological Review, 58*, 741–752.
Hout, M., & Greeley, A. M. (1990). The center doesn't hold: Church attendance in the Unites States, 1940–1984. *American Sociological Review, 52*, 325–345.
Kay, A. C., Gaucher, D., McGregor, I., & Nash, K. (2010). Religious belief as compensatory control. *Personality and Social Psychology Review, 14*, 37–48.
Kay, A. C., Moscovitch, D. A., & Laurin, K. (2010). Randomness, attributions of arousal, and belief in God. *Psychological Science, 21*, 216–218.
Kay, A. C., Whitson, J. A., Gaucher, D., & Galinsky, A. D. (2009). Compensatory control: Achieving order through the mind, our institutions, and the heavens. *Current Directions in Psychological Science, 18*, 264–268.
Keinan, G. (2002). The effects of stress and desire for control on superstitious behavior. *Personality and Psychology Bulletin, 28*, 102–108.
Kenrick, D. T., Maner, J. K., & Li, N. P. (2005). Evolutionary social psychology. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 803–827). Hoboken, NJ: Wiley.
Kosmin, B. A., & Keysar, A. (2009). *American Religious Identification Survey (ARIS 2008) summary report*. Hartford, CT: Institute for the Study of Secularism in Society & Culture.
Kurzban, R. (2010). *Why everyone (else) is a hypocrite*. Princeton, NJ: Princeton University Press.
Levin, J. S., Taylor, R. J., & Chatters, L. M. (1994). Race and gender differences in religiosity among older adults: Findings from four national surveys. *Journal of Gerontology, 49*, S137–S145.
Liddle, J. R. (2015). *An analysis of religiosity in the United States: Testing the Secure Society Theory* Manuscript submitted for publication.
Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Germany: Springer.
MacDonald, K. B. (2007). Cross-cultural comparison of learning in human hunting: Implications for life-history evolution. *Human Nature, 18*, 386–402.
Malinowski, B. (1954). *Magic, science and religion, and other essays*. Garden City, NY: Doubleday.
McBurney, D. H., Gaulin, S. J. C., Devineni, T., & Adams, C. (1997). Superior spatial memory of women: Stronger evidence for the gathering hypothesis. *Evolution and Human Behavior, 18*, 165–174.
McCullough, M. E., & Larson, D. B. (1999). Religion and depression: A review of the literature. *Twin Research, 2*, 126–136.
Moro, P. A., & Myers, J. E. (2010). *Magic, witchcraft, and religion: A reader in the anthropology of religion* (8th ed.). New York: McGraw-Hill.
Norenzayan, A. (2010). Why we believe: Religion as a human universal. In H. Høgh-Olesen (Ed.), *Human morality and sociality: Evolutionary and comparative perspectives* (pp. 58–71). New York: Palgrave MacMillan.
Norris, P., & Inglehart, R. (2004). *Sacred and secular: Religion and worldwide politics*. New York: Cambridge University Press.
Paul, G. S. (2005). Cross-national correlations of quantifiable societal health with popular religiosity and secularism in the prosperous democracies. *Journal of Religion & Society, 7*, 1–17.
Paul, G. S. (2009). The chronic dependence of popular religiosity upon dysfunctional psychosociological conditions. *Evolutionary Psychology, 7*, 398–441.
Pesta, B. J., McDaniel, M. A., & Bertsch, S. (2010). Toward an index of well-being for the fifty U.S. states. *Intelligence, 38*, 160–168.
Rees, T. J. (2009). Is personal insecurity a cause of cross-national differences in the intensity of religious belief? *Journal of Religion and Society, 11*, 1–24.
Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology, 57*, 199–226.
Sasaki, M. (1979). Status inconsistency and religious commitment. In R. Wuthnow (Ed.), *The religious dimension: New directions in quantitative research* (pp. 35–156). New York: Academic.
Sims, C. A. (1980). Macroeconomics and reality. *Econometrica, 48*, 1–48.

shackelf@oakland.edu

Smith, H. (1991). *The world's religions* (revised edn.). New York: HarperCollins.
Smith, T. W. (1988). *Timely artifacts: A review of measurement variation in the 1972–1988 GSS* (GSS methodological report No. 56). Chicago: National Opinion Research Center.
Smith, T. W. (1996). *A review of church attendance measures (GSS methodological report No. 88)*. Chicago: National Opinion Research Center.
Smith, T. W., Marsden, P., Hout, M., & Kim, J. (2013). *General social surveys, 1972–2012: Cumulative codebook*. Chicago: National Opinion Research Center.
Solt, F., Habel, P., & Grant, J. T. (2011). Economic inequality, relative power, and religiosity. *Social Science Quarterly, 92*, 447–465.
Stark, R., & Bainbridge, W. (1985). *The future of religions*. Berkeley, CA: University of California Press.
Stock, J. H., & Watson, M. W. (2001). Vector autoregressions. *Journal of Economic Perspectives, 15*, 101–115.
U.S. Department of Commerce. (2013). *Current population survey, 2013 annual social and economic (ASEC) supplement*. Washington, DC: U.S. Census Bureau.
U.S. Department of Justice, Federal Bureau of Investigation. (2014). *Uniform crime reports: Crime in the United States*. Retrieved from http://www.fbi.gov/about-us/cjis/ucr/ucr-publications#Crime
Weber, B. H., & Depew, D. J. (2003). *Evolution and learning: The Baldwin Effect reconsidered*. Cambridge, MA: MIT Press.
Whitson, J. A., & Galinsky, A. D. (2008). Lacking control increases illusory pattern perception. *Science, 322*, 115–117.
Zivot, E., & Wang, J. (2003). *Modeling financial time series with S-PLUS*. New York: Springer.
Zuckerman, P. (2007). Atheism: Contemporary rates and patterns. In M. Martin (Ed.), *Cambridge companion to atheism* (pp. 47–67). New York: Cambridge University Press.
Zuckerman, P. (2008). *Society without god: What the least religious nations can tell us about contentment*. New York: New York University Press.
Zuckerman, P. (2009). Atheism, secularity, and well-being: How the findings of social science counter negative stereotypes and assumptions. *Sociology Compass, 3*(6), 949–971.

shackelf@oakland.edu

Part V
Politics, Law, and Game Theory

shackelf@oakland.edu

The Evolved Functions of Procedural Fairness: An Adaptation for Politics

Troels Bøggild and Michael Bang Petersen

Politics is the process of determining resource access (Petersen, 2015). In a classical definition, the political scientist Easton (1953) thus defined politics as the process of determining the "authoritative allocation of values for a society." In more plain terms, another classical definition described politics as about "who gets what" (Lasswell, 1950). These definitions highlight that politics is ultimately about *outcomes*: who gets a resource and who does not. Some researchers of politics have taken this to imply that outcomes are all people care about in politics. As Ulbig (2002, p. 793) notes, the traditional notion has been that "when people get what they want, they do not care how they get it." Popkin (1991, p. 99), for example, holds that people "care about ends not means; they judge government by results and are ... indifferent about the methods by which the results were obtained."

A brief look at the reporting of political news, however, makes it clear that outcomes are just one part of what receives attention in politics. Reports on the substances of political policies (who is affected by the policies and how) are often outnumbered by reports on the political process that leads to the policies (see Aalberg, Strömbäck, & de Vreese, 2011; Patterson, 1994): who are the key actors, why do they promote or oppose a specific policy, what are their interests, and what is their strategy for achieving those interests? Research has documented that such a focus is, at least in part, demand driven: people crave information about the process leading to outcomes and the strategic aspects involved (Smith, Larimer, Littvay, & Hibbing, 2007; Trussler & Soroka, 2014). This appetite for information about what goes on behind the scenes of political outcomes extends even to the personal lives of politicians: who are they, whom are they personally affiliated with, and do they practice what they preach politically? This latter observation illustrates an important

T. Bøggild (✉) • M.B. Petersen
Department of Political Science, Aarhus University,
Bartholins Allé 7, Aarhus C 8000, Denmark
e-mail: tboeggild@ps.au.dk; michael@ps.au.dk

© Springer International Publishing Switzerland 2016

T.K. Shackelford, R.D. Hansen (eds.), *The Evolution of Morality*,
Evolutionary Psychology, DOI 10.1007/978-3-319-19671-8_12

shackelf@oakland.edu

point: people's political interest not only extends beyond a narrow interest in the substance of policies; this interest is also moral in nature. Public outrage and strong moral condemnation meet the hypocritical and self-interested politician (Bøggild, 2014; Cappella & Jamieson, 1996; Hibbing & Alford, 2004).

The observation that people are morally preoccupied with not just political outcomes but also the processes leading to these outcomes is supported by decades of research in social psychology. This research has predominantly focused on the perceived legitimacy of decisions in the context of everyday life rather than politics. In the contexts of everyday life, it has been demonstrated (1) that people are interested in many more features of decisions than outcomes and (2) that this interest is, in part, moralistic (Charness & Levine, 2007; Falk, Fehr, & Fischbacher, 2008; Ong, Riyanto, & Sheffrin, 2012; Tyler, 1990; Van den Bos, Wilke, & Lind, 1998). Many of these studies have been organized around the concept of procedural fairness, which highlights both the orientation toward procedures and the moral nature of this orientation. Empirical findings include that people view decisions as more legitimate if they were allowed to voice their opinions before the decision was made, if the decision-maker was impartial and lacked personal interest in the decision, and if all relevant parties were included in the decision-making process (e.g., Tyler, 1990). The overall finding that people care not only about the substance and outcomes of decisions but also about *how decisions come about* when evaluating group leaders and their decisions has been characterized as one of the most replicated findings in social psychology (Van den Bos, Wilke, & Lind, 1998, p. 1449). Recent empirical studies have extended the concept of procedural fairness to politics, directly confirming the observation that people care about political process as well as political outcomes. It has been shown that whether political institutions follow fair procedures is an important determinant of citizens' trust in national-level political and legal institutions and political representatives and also that political decisions that emerge from fair procedures receive greater public support (Bøggild, 2014; Grimes, 2006; Tyler, 1994; Ulbig, 2008).

In this chapter, we ask: why are people morally preoccupied with features of the processes leading to political outcomes rather than just with the outcomes themselves? Although the concept of procedural fairness has received widespread attention across the social sciences over the past four decades, it remains contested at the theoretical level. As argued by Smith et al. (2007, p. 288), while the empirical support for people's attention to procedural fairness is steadily growing, "the absence of a theoretical explanation becomes ever more glaring." More specifically, existing theoretical frameworks have been criticized for being too narrow in their scope and for providing proximate accounts that assume rather than explain why procedural fairness matters (Hibbing & Theiss-Morse, 2008). In reference to the distinction between ultimate ("why") and proximate ("how") explanations (Scott-Phillips, Dickins, & West, 2011), we here seek to provide an ultimate explanation of why human moral judgments about decisions are oriented toward the procedural aspects of the decision. To achieve this, we part with previous work on procedural fairness in social psychology in two ways: First, for reasons developed below, we utilize an evolutionary biological framework and dissect the fitness benefits of

shackelf@oakland.edu

psychological systems for procedural fairness for humans. Second, we provide an explanation that puts politics front and center. The psychology of procedural fairness, we suggest, evolved specifically to deal with key problems related to politics, such as the formation of hierarchies and the delegation of leadership. Part of the failure of previous theoretical accounts, we suggest, emerges because these accounts have not taken their point of departure in political clashes of interest. Notions of procedural fairness are crucial parts of human nature as a political animal, and, to understand the ultimate function of these notions, they need to be analyzed in the context of politics.[1]

The chapter proceeds in four steps. The first section reviews existing theoretical models of procedural fairness and the main points of criticism that have been launched against them. The second section makes the case for an evolutionary approach to procedural fairness, emphasizing the role of politics in human evolution and fleshes out the adaptive problem that procedural fairness evolved to accommodate: exploitative political leaders. The third section specifies how notions of procedural fairness serve two distinct purposes in terms of (1) detecting exploitative tendencies in leaders (i.e., a diagnostic function) and (2) actively counteracting and restricting the potential for exploitation in political processes (i.e., a bargaining function). This section also reviews a range of existing empirical evidence in procedural fairness research from social psychology consistent with this account. The fourth section describes how the psychological system for procedural fairness influences modern political attitudes and behaviors—and why mismatches between the circumstances of ancestral and modern, mass-level politics might cause this influence to be even more powerful today than ancestrally.

Procedural Fairness: Existing Findings and Theory

Over the last four decades, research on procedural fairness in social psychology has provided extensive evidence that people look to certain information about *how decisions come about* when evaluating group authorities and the decisions they implement (Thibaut & Walker, 1975). This literature has uncovered several criteria for what constitutes a widely perceived legitimate or "fair" decision-making process (Leventhal, 1980; Tyler, 1990). The literature has also demonstrated that evaluations of the decision-making process (procedural fairness) are distinct from evaluations of the decision outcomes (e.g., distributive fairness; see Colquitt, 2001) and, hence, require a distinct set of explanations.

Three studies can serve as illustrations of the criteria people used when making judgments of procedural fairness. A seminal illustration comes from Tyler's (1990)

[1] We do not mean to imply that extant (or future) research on procedural fairness conducted in legal, managerial, or educational settings is not important. Rather, we view these contexts as also political in the sense that the decision-making processes manage clashes of interests between opposing parties.

shackelf@oakland.edu

book, *Why People Obey the Law.* Through panel analysis, it is demonstrated how people who had been in recent contact with the legal system were more likely to comply with and accept an unfavorable verdict when the judge adhered to certain procedural criteria such as allowing them to voice their opinions, appeared impartial and unbiased and included all relevant parties in the decision-making process—even after controlling for the outcome of the decision.

In another illustrative study, Cremer and van Knippenberg (2003) had subjects play a public goods game in which subjects were to decide how much of their own endowment they wanted to contribute to a public pot. If the public pot reached a certain size, the pot would be doubled and then divided among all players, creating an incentive to contribute to providing this public good but also an incentive to free-ride and let others make the necessary contributions. Each group of subjects played a total of six contribution rounds and was assigned a group leader deciding throughout the game which subjects deserved a share of the public pot. The authors manipulated whether the group leader allowed subjects to voice their opinions (as opposed to denying voice) before deciding how to divide the public pot and whether the group leader had accurate (as opposed to inaccurate) information about subjects' contribution records to base his or her decision on. As expected, a group leader allowing voice in the decision-making process and basing the decision on accurate information increased subjects' subsequent contribution levels, even after controlling for the payoffs received.

In a final illustrative study, Ramirez (2008) investigated through a survey experiment the effects of media coverage concerning procedural fairness of the Supreme Court on public support for the Court and the individual justices. Specifically, the author had subjects read a news article and manipulated whether the Court was portrayed as allowing all parts to voice their opinions, was representative of the American people, and was objective and impartial in their decision-making. As expected, subjects expressed higher levels of support for the Court and its justices when these procedural criteria were met.

These studies illustrate the findings in the literature on procedural fairness. Other studies have demonstrated how procedural fairness matters beyond the legal context, including in educational (Tyler & Caine, 1981), managerial (Lavelle et al., 2009), and political (Bøggild, 2014; Tyler, 1994) settings. Figure 1 reports six widely considered procedural criteria. The list is not exhaustive and the conceptual terminology sometimes varies from study to study within social psychological literature. Nonetheless, there is general agreement that the listed procedural criteria are of key importance when people make moral evaluations of decisions.

To account for why people pay attention to such features of the decision-making process when evaluating decision-makers and their decisions, research in social psychology has offered three different theoretical accounts. First, early work adopted an *instrumental model* of procedural fairness, holding that people insist on procedural fairness to gain control over and attain the best possible outcome (Thibaut & Walker, 1975). For example, it is argued that people prefer having a voice in decision-making processes because it serves as a means of control over the

shackelf@oakland.edu

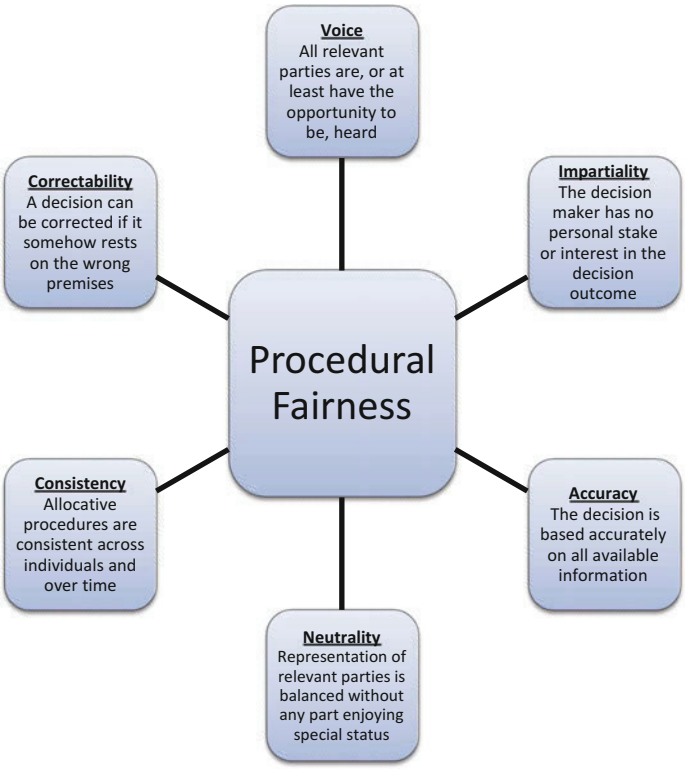

Fig. 1 Procedural fairness criteria. *Note*: The listed criteria reflect the work by Leventhal (1980), Tyler (1990), Colquitt (2001), and Blader and Tyler (2003)

outcome. This model has been criticized for being too narrow in its scope and for disregarding the normative or noninstrumental motivations related to procedural fairness (Tyler, 1990). For example, studies have demonstrated how people care about being granted voice even after a decision is made, which seems beyond the explanatory range of the instrumental model (Lind, Kanfer, & Christopher, 1990).

Second, to accommodate these shortcomings, Tom Tyler and colleagues proposed the *relational model of authority* (Tyler & Lind, 1992), holding that people's attention to procedural fairness stems from a desire to construct and uphold a social identity. From this perspective, the individual is highly vigilant about how decisions come about because it provides him with a clear indication of the extent to which the group and especially its authority figure(s) regard him as an equal and valuable group member (Lind & Tyler, 1988; Tyler & Lind, 1992). When information on procedural fairness signals that authorities do not appreciate the status and standing of the individual (e.g., by not allowing voice or being partial), it responds with decreased levels of trust in authorities and compliance with decisions. For example,

shackelf@oakland.edu

in favor of this relational account, studies have demonstrated how experiencing unfair procedures also lowers self-esteem of the individual (De Cremer, van Knippenberg, van Knippenberg, Mullenders, & Stinglhamber, 2005; Koper, Van Knippenberg, Bouhuijs, Vermunt, & Wilke, 1993). Critics, however, have argued that the relational model does not explain why procedural fairness matters in large-scale settings beyond personal interactions. For example, it seems less straightforward why people would look to information on procedural fairness of national courts or governments to determine whether they are valued group members and construct a social identity on this basis (Leung, Tong, & Allan, 2007). Moreover, scholars have questioned whether the relational model constitutes a theory at all. Simply holding that people have a need for constructing a social identity and that decision-making processes somehow provide relevant information in this regard, it is argued, is merely a statement or at best a description without an underlying theoretical logic (Smith et al., 2007). These critics hold that "Tyler has won the empirical battle but there is no victor in the theoretical war" (Hibbing & Theiss-Morse, 2008, p. 125).

Third, *fairness heuristic theory* (Van den Bos, Lind, Vermunt, & Wilke, 1997) constitutes the most recent theoretical framework in procedural fairness research. According to this framework, people often lack relevant information in evaluating and making sense of the outcomes they receive from group decisions. For example, people are often faced with situations in which they are uncertain about the trustworthiness of decision-makers and whether their own outcome is fair compared to outcomes of others. Fairness heuristic theory holds that under such uncertainty, people compensate by applying information on procedural fairness as a heuristic in evaluations of decision outcomes (Brockner & Wiesenfeld, 1996; Van den Bos, Wilke, & Lind, 1998). In support of this model, Van den Bos et al. (1997) have demonstrated how information on procedural fairness has a larger impact on evaluations of decisions when subjects only know their own outcomes and not the outcomes of others compared to when the outcomes of all subjects are known. Yet, fairness heuristic theory has been subjected to the same criticism leveled against the relational model of authority. This framework focuses only on the cognitive processes related to *how* people apply information on procedural fairness and does not address *why* such information, over other types of information, helps the individual cope with uncertainty (van Prooijen, 2008, p. 36). That is, "what precisely it is about procedural fairness that makes it useful under conditions of uncertainty remains unclear" (Gonzalez & Tyler, 2007, p. 94).

In sum, although existing models have generated important insights on the effects of procedural fairness, they remain contested at the theoretical level. Our view is that each of these theoretical models has important merits and that an integrative account, which is immune to the criticism leveled against the models, can be formulated through further theorizing. Specifically, we argue that criticisms of the previous models can be accommodated by considering where preferences for procedural fairness ultimately stem from. Despite extensive scholarly attention to the concept of procedural fairness, key researchers in this literature have noted that "…little is known about the origins of procedural preferences" (Tyler, 1990, p. 109)

shackelf@oakland.edu

and that what is needed is "a deeper account of the reasons people are so concerned with process variables" (Smith et al., 2007, p. 288). Deep, origin-oriented accounts are important in their own right, but they are also important because we become better able to understand how something works when we know why it exists. That is, by knowing *why* people care about procedures, we inevitably learn more about *how and when* we care about them.

Where to Look for "Why"? Evidence for the Relevance of Biological Evolution

In this chapter, we seek to provide an ultimate explanation of why human moral judgments about decisions are oriented toward the procedural aspects of the decision. In emphasizing the "ultimate" nature of our explanation, we refer to the distinction between ultimate ("why") and proximate ("how") explanations (Scott-Phillips et al., 2011). Ultimate explanations of a trait involve explaining why this given trait (here, moral attention to procedures) rather than alternative traits (e.g., moral attention to outcomes only) exists in a population.

Ultimate explanations are often functional explanations. That is, the trait is explained with reference to the function it serves within a population. Two types of functional explanations are common: explanations of cultural evolution and explanations of biological evolution. Theories of cultural evolution explain a trait with reference to the function it serves within a culturally specific population. In the case of procedural fairness, it could be that such norms help sustain democratic regimes. Theories of biological evolution explain the existence of a trait with reference to the fitness benefits it conferred ancestrally. In determining whether a trait is the effect of cultural evolution, biological evolution (Cosmides & Tooby, 1997), or both (Boyd & Richerson, 1988), one key question relates to how widespread the trait is[2]: what is the appropriate population? Are norms of procedural fairness specific to a few cultures or are they cultural universals? If the former, it lends credence to explanations oriented toward cultural evolution; if the latter, it lends prima facie credence to explanations oriented toward biological evolution. This is certainly not everything that needs to be established to determine the relevance of either explanatory type (Cosmides & Tooby, 1997). But it is an important first step.

The evidence from psychological research is unequivocal. The effects of procedural fairness are pan-cultural, having been replicated in nondemocratic and non-Western countries such as Japan (Takenishi & Takenishi, 1992), Singapore (Khatri, Fern, & Budhwar, 2001), India (Pillai, Williams, & Tan, 2001; Platow et al., 2013), South Africa (VanYperen, Hagedoorn, Zweers, & Postma, 2000), Russia (Giacobbe-Miller, Miller, & Victorov, 1998), and China (Wilking, 2011). Moreover, anthropological

[2] However, see Tooby and Cosmides's (1992) notion of evoked culture that is not culturally universal but still the product of biological evolution.

shackelf@oakland.edu

studies suggest that even simple hunter–gatherer societies entirely excluded from modern civilizations evaluate leaders by means of procedural fairness standards such as voice and impartiality (Boehm, 1993; Meggitt, 1977; see also Van Vugt, Hogan, & Kaiser, 2008). As concluded by Krislov (1991, p. 237), although procedural fairness may be practiced in somewhat different ways across cultures, "the generic desire for it may be much broader, even universal."

These observations provide a prima facie case that intuitions about procedural fairness need to be explained not with reference to culturally specific processes but with reference to processes of biological evolution—that these intuitions constitute an ingrained part of human nature. The conclusion that procedural fairness norms emerge from specialized, deep-seated psychological processes is further sustained by evidence from two disciplines: developmental psychology and affective neuroscience. Developmental studies have shown that children of a very young age possess a sophisticated understanding of procedural fairness and react strongly to perceived unfair procedures (Fry & Corfield, 1983; Hicks & Lawrence, 1993). Gold, Darley, Hilton, and Zanna (1984), for example, find that 6–7-year-old children evaluate adult authorities by means of procedural fairness standards and easily detect and respond to unfair treatment. Studies using techniques from affective neuroscience have furthermore shown that information on procedural fairness yields responses that are highly affective, automatic, and unconscious (Hibbing & Alford, 2004; Weiss, Suckow, & Cropanzano, 1999; see also Ham & van den Bos, 2008), and that these responses rely on distinct parts of the brain that are different from those processing information on distributive fairness (Beugré, 2009; Dulebohn et al., 2009). These observations on children's and adults' reasoning about procedural fairness suggest that intuitions about procedural fairness emerge naturally as part of normal developmental processes and due to the existence of deep-seated, dedicated psychological mechanisms. Both observations reinforce the suggestion that the psychological mechanisms underlying these intuitions could have evolved through biological evolution (see Dulebohn et al., 2009; Hibbing & Alford, 2004). If so, intuitions about procedural fairness should ultimately be explained with reference to their fitness benefits throughout human evolutionary history and, hence, as products of natural selection.

Adaptive Problems in Politics: The Benefits and Costs of Hierarchies

A psychological mechanism can evolve by natural selection to the extent the mechanism helps the organism solve an evolutionarily recurrent problem. An adaptive "problem" is anything that has repercussions for the organism's fitness (Tooby & Cosmides, 1992). In such a case, solving the given problem is the mechanism's biological "function." Identifying the particular problem that a mechanism evolved to solve requires an identification of the context in which the mechanism in question is activated.

shackelf@oakland.edu

In the case of the psychology of procedural fairness, there has been no systematic empirical study of its precise domain of application and this is something that is important to explore empirically in the future. Still, an important observation can be made: as illustrated by the three descriptions of social psychological studies of procedural fairness, considerations of procedural fairness generally only become relevant when decisions are made on behalf of others. When people make decisions for themselves, it can be asked whether the decision process was rational but not whether it was fair. In this sense, the evolution of the psychology of procedural fairness is narrowly and intimately related to the evolution of politics: the struggle for decision-making power and, ultimately, resources in social groups.

Politics as a Set of Adaptive Problems

Over the course of evolutionary history, humans have reaped the benefits of living in groups, providing protection against other groups and predatory enemies, higher reproduction rates, and reduced variation in appropriation of food (Foley, 1995; Kenrick, Li, & Butner, 2003). Group life, however, also brought challenges, most significantly in terms of social coordination issues like group movement, intergroup competition, and internal conflict resolution and peace-keeping (De Waal, 1997; Van Vugt et al., 2008). There are multiple ways to handle such common problems within and between groups, and each way entails a different distribution of costs and benefits across individual members. Group life therefore inevitably involves a negotiation about how the spoils of group life should be divided: who should benefit the most and who should pay the greatest costs? Politics is this negotiation about who gets what.

A crucial part of politics is the formation of hierarchies that establishes shared expectations about resource access, with those in the top of the hierarchy getting more. In most animals, these hierarchies are based on dominance: the alpha male is the one who can dominate conspecifics. For many animal species, this is a matter of sheer physical formidability: the biggest and strongest individual in a group can dominate the rest. In contrast, for humans and a few other animals such as chimpanzees, individual formidability does not translate directly into political power. In highly social animals, coalitions of less formidable individuals can be organized to counter any single individual (Boehm, 1999; De Waal, 2007), and, for humans, political power is mostly a matter of abilities to form and maintain alliances (von Rueden, Gurven, Kaplan, & Stieglitz, 2014; Whiten & Byrne, 1997). As a consequence, for the human political entrepreneur, one strategy for forming a stable base of allies is to seek *prestige* rather than *dominance* (Price & Van Vugt, 2014; Sidanius & Pratto, 2001).[3] Prestige is a status that is granted rather than taken (Henrich &

[3] It should be noted that allies are often also used for dominance. In particular, one group of allies will often seek to dominate other groups of allies (Sidanius & Pratto, 2001). Still, within each of these groups, one individual will often emerge as more prestigious than others.

shackelf@oakland.edu

Gil-White, 2001). It is granted on the basis of the individuals' abilities to address the problems confronting potential allies and in that way provide benefits in exchange for the status these individuals receive (Price & Van Vugt, 2014; von Rueden et al., 2014).

According to a steadily growing literature across the social sciences, a sophisticated psychology of leadership and followership has evolved in response to these selection pressures (Tooby, Cosmides, & Price, 2006; Van Vugt, 2006; Van Vugt, Ahuja, & Van Vugt, 2011). A set of psychological mechanisms—a leadership psychology—evolved to help individuals detect opportunities for prestige generation. Reciprocally, a set of psychological mechanisms—a followership psychology—evolved to follow and bestow prestige on those individuals (i.e., group leaders) displaying competence in solving crucial adaptive problems (Laustsen & Petersen, 2015; Price & Van Vugt, 2014). The psychology of procedural fairness, we suggest, evolved as part of this larger psychology of followership to ensure that leaders return benefits in exchange for the prestige they are granted.

The Costs and Benefits of Followership

The dual existence of a psychology of followership and leadership allows leaders and followers to engage in a reciprocal relationship where the costs of coordinating collective decision-making are delegated to a specific individual, the leader. In accepting these costs, the leader would produce public goods in terms of coordinating group efforts and maintaining intra-group peace. In return, the leader would receive high social status and the associated fitness benefits of increased access to material and reproductive resources of the group (Price & Van Vugt, 2014). In politics, a crucial adaptive problem is therefore to follow and delegate leadership to the right individuals.

A key risk for any follower in a hierarchical, leader–follower relationship is that the delegation of leadership allows the leader to exploit his or her privileged position to either extract benefits for himself or herself (or his or her friends or kin) or impose costs on his or her enemies (Boehm, 1999; von Rueden et al., 2014). As argued above, there are multiple solutions to any problem that confronts the collective, and each of these solutions imposes different benefits and costs on different members of the collective. In this regard, there is often an information asymmetry between followers and leaders.[4] Often only the leader will have an accurate representation of the space of possible solutions and the costs and benefits associated

[4] In the literature on principal–agent problems, the concept of information asymmetry is used to describe the problems that emerge when followers are implementing the decision of the leader. In such situations, followers sometimes have more information available and, hence, can use the resulting degrees of freedom to implement a decision in a way beneficial for the self. Here, we argue that there is also an information asymmetry when decisions are made (and not just when they are implemented), and this information asymmetry is often to the advantage of the leader.

shackelf@oakland.edu

with each. The delegation of leadership involves the risk that this information asymmetry is utilized to the follower's disadvantage.

More formally, from the perspective of the individual follower, the problem is whether his or her welfare receives sufficient consideration in collective decisions. The problem is not so much if this happens once or twice. The real problem is the accumulation of many small fitness losses over multiple decisions. That is, if the individual's welfare is systematically considered less when decisions are made. In the context of leadership and followership, the crucial factor for whether this happens can be referred to as the leader's *welfare tradeoff ratio* (WTR) toward the individual follower. The terminology of WTRs is an attempt to describe the psychological machinery that evolved to regulate social decisions (Petersen, Sell, Tooby, & Cosmides, 2010; Sell, Tooby, & Cosmides, 2009). When social decisions have repercussions for the welfare of others besides the decision-maker, the decision-maker needs to be able to trade-off these repercussions. How much of their own welfare are decision-makers willing to sacrifice for the welfare of another? On the basis of a number of diverse cues such as whether the other is kin, an ally, or an exchange partner, humans seem to compute a psychological summary tally, the WTR, for each person in their social network, and this tally is accessed during social decisions and guides the relevant welfare trade-offs (Petersen, Sell, Tooby, & Cosmides, 2012; Tooby, Cosmides, Sell, Lieberman, & Sznycer, 2008). If the decision-maker is willing to trade-off more of his or her own welfare for the welfare of another person, the decision-maker has a high WTR toward that individual. This psychological tally is relatively stable, and, hence, the magnitude of the WTR of individual A for individual B reliably predicts future resource flows from A to B.[5]

If we apply this logic to the relationship between a follower and a leader, it implies that if a follower faces a leader with a relatively low WTR, this leader would potentially select those particular solutions across time that were to the follower's relative disadvantage. This constitutes a fundamental adaptive problem for followership decisions which could result in serious negative feedback effects on the follower's long-term fitness (Erdal & Whiten, 1996; Haidt, 2013; Price & Van Vugt, 2014).

In Jonathan Haidt's words "Individuals who failed to detect signs of domination and respond to them with righteous and group-unifying anger faced the prospect of reduced access to food, mates, and all the other things that make individuals (and their genes) successful in the Darwinian sense" (Haidt, 2013, p. 173). Hence, humans should not only possess psychological mechanisms designed to evaluate and follow leaders that are *competent* but should also have developed a sophisticated psychology for aligning themselves with leaders that *hold high WTRs* toward the individual and, hence, refrain from exploitation.

[5] The literature distinguishes between *intrinsic* WTRs guiding behavior when no one else oversees the action (determined by kin or friendship relationships) and *monitored* WTRs guiding behavior when actions are subjected to supervision by others (determined by these others' abilities and willingness to confer benefits and costs on the self upon detection of the self's action) (see Petersen et al., 2010). In the main text of this chapter, we use the term WTR exclusively to refer to monitored WTRs. In footnote 7, we return to the difference between monitored and intrinsic WTRs.

shackelf@oakland.edu

The Adaptive Functions of Procedural Fairness

We suggest that the psychology of procedural fairness evolved as part of the psychology for countering a leader who (1) devalues the welfare of the self and (2) is willing to act upon it. Specifically, we argue that attention to the decision-making processes of leaders (and not just the outcomes of these processes) evolved to serve two key interrelated functions. First, attention to information on procedures serves a *diagnostic function* as this information conveys reliable cues to determine the dispositional tendencies and WTRs of group leaders. Second, the moralistic nature of this attention to information on procedures serves a *bargaining function*: when violations of procedural fairness are detected, they are broadcasted, fueling outrage among and mobilizing others who also fear mistreatment. This latter function helps followers not just detect but also to form political coalitions and actively upregulate leaders' WTRs in collective decisions and hence counteract exploitation.

Procedural Fairness as a Diagnostic Tool

Which type of information would have served as reliable indicators of leaders' exploitative tendencies and WTRs over evolutionary history? One possibility is for the individual to evaluate the WTR of a leader toward the self by simply attending to the favorability of decision outcomes generated by the leader. However, evaluating WTRs of a leader by merely attending to the outcomes of decisions adopted by a leader would entail an inadequate and flawed strategy since unfavorable decisions would not necessarily imply a lack of concern for the welfare of the follower (Dubreuil, 2011, pp. 32–37). For example, enforcing decisions which in the short term impose costs on individuals (such as rationing food or, in modern organizations, introducing layoffs or unpopular welfare state retrenchment policies) will at times carry long-term gains. In a similar vein, research on collective action and free-riding has demonstrated that people do not make inferences about the WTRs held by others based on the absolute gains they have received from prior interactions with them but on a different set of indicators that more reliably convey antisocial, cheating dispositions (Delton, Cosmides, Guemo, Robertson, & Tooby, 2012). In short, evaluating the WTR of leaders by mere attention to decision outcomes would terminate potentially viable relationships with pro-social leaders committed to the welfare of the group and its followers.

The key identification problem is this: when the outcome of a decision enforced by a leader imposes relative costs on a follower, how can the follower identify whether this is caused by a low WTR (predicting further cost imposition)? To form reliable assessments about the WTR of a leader under such circumstances, the individual will have to rely on other, non-outcome information or cues. This is where procedural information provides important diagnostic information.

shackelf@oakland.edu

The advantage of attending to information on procedural fairness, in contrast to outcomes, is that it helps reveal the dispositional characteristics of leaders and the considerations giving rise to a negative decision. For example, if a leader enforces a collective decision in which he or she holds a personal, vested interest (i.e., partiality), it sends a reliable signal to the followers concerning the leader's attention to their welfare vis-à-vis his or her own interests in a leadership role. Or, when a follower does not get a voice in collective decisions while others do, it serves as an accurate cue that the leader holds a low WTR toward this individual relative to the welfare of himself and/or other group members. Essentially, when criteria of procedural fairness are violated, it signals that the leader does not adhere to the social contract related to prestige (i.e., service in exchange for status) and the relationship drifts from one of prestige (where benefits are granted) to dominance (where benefits are taken without a reciprocal return of service). Consistent with this, anthropological studies have identified a number of leadership behaviors that generate dissatisfaction with the leader, and their findings draw a conspicuous parallel to the criteria of procedural fairness theory surveyed above (Boehm, 1993, 1999; Meggitt, 1977; Von Furer-Haimendorf, 1967; see also Van Vugt et al., 2008). For example, in an extensive ethnographic study of 48 modern hunter–gatherer foraging societies, Christopher Boehm found that dissatisfaction was triggered among followers by "ineffectiveness, partiality and unresponsiveness in a leadership role" (Boehm, 1993, p. 231).

We suggest that a key evolved function of procedural attention is as a prediction tool for the individual follower. By using procedural features to estimate the WTR of the leader toward them, each individual follower seeks to predict whether resources will flow to them in the future if they continue to follow a specific leader. In emphasizing this function, we also argue that alternative evolved functions cannot adequately explain the psychology of procedural fairness. Most importantly, we suggest (1) that the psychology is not designed to determine which leader is best for the group at large (as some might suggest, see Boehm, 1999) but instead is designed to help individuals determine which leader is best for themselves[6] and (2) that the psychology is not designed to determine the correctness of single decisions but to make predictions about future behavior of the decision-maker. In this way, this function blends the traditional instrumental and relational explanations of procedural fairness: procedural fairness reflects an instrumental attempt to make predictions about the relationship between the leader and a follower. Several features of the structure of procedural fairness judgments are predictable from this evolved function.

First, an extensive literature in social psychology has demonstrated that effects of procedural fairness on evaluations of decision-makers and the decisions they implement depend on whether the outcome benefits or harms the individual (referred to as *outcome favorability*). More specifically, unfair procedures make people lower their evaluations of decision-makers and the decision implemented when the outcome

[6] In terms of selection pressures, the argument that we propose works at the level of individuals and not groups. Moral intuitions are, in other words, strategic in the sense that they serve the interest of the individual and not the group at large (see Petersen, 2013).

shackelf@oakland.edu

generated through such procedures ultimately harms them, while they tend to tolerate or overlook the use of unfair procedures when the outcome is favorable (see Brockner & Wiesenfeld, 1996 for evidence from more than 40 studies). The psychology of procedural fairness is self-centric: it is activated when the self is harmed and used to form representations under these circumstances, in particular. This conclusion is reinforced by studies that show that the effects of procedural fairness also depend on whether the decision affects the individual (i.e., *outcome dependence*). Two studies have demonstrated that the effects of procedural fairness on evaluations of leaders and their decisions are stronger when enacted by leaders representing the individual's group rather than leaders of other groups (Smith, Tyler, Huo, Ortiz, & Allan, 1998; Tyler, Allan Lind, Ohbuchi, Sugawara, & Huo, 1998). Moreover, van Dijke, Marius, and Mayer (2010) have shown that the use of unfair procedures has a larger negative effect on evaluations of the decision-maker and the decision it implements when the decision-maker is powerful and can impose decisions with direct implications on the individual. Finally, Ham and van den Bos (2008) found that people make faster inferences about procedural fairness when reading about fair and unfair events described in the first-person singular compared to descriptions in the third-person singular (see also van Prooijen, 2008). In sum, these findings provide strong evidence that the psychology of procedural fairness is not designed to estimate whether a leader is mistreating group members at large. If so, procedural features should be attended to independently of the outcome of the individual. Rather, it is specifically designed to activate when the specific follower faces a leader with a potentially low WTR toward the self.

Second, consistent with the notion that procedural fairness criteria are used to make representations about WTRs, a range of social psychological studies show that people use information on procedural fairness to evaluate the trustworthiness and perceived legitimacy of the decision-maker (Colquitt, 2001; Tyler & Degoey, 1996). The concept of trust is intimately related to evaluations of WTRs since both relate to the individual's belief that another actor, whose behavior can somehow affect the individual's welfare, will attend to the individual's interests and well-being. Some scholars see the connection between procedural fairness and trust in decision-makers as so self-evident that they treat trustworthiness of the decision-maker as a procedural criterion itself or a proxy for perceived procedural fairness rather than as a dependent variable (e.g., Tyler, 1997). As a result, procedural fairness research has instead often looked to the effects of procedural fairness on perceived legitimacy of the decision-maker. For example, procedural fairness is a key determinant of citizens' perceived legitimacy of legal and political institutions including the Supreme Court and the United States Congress (Tyler, 1990, 1994). Moreover, Dulebohn et al. (2009) have demonstrated how experiencing unfair procedures activates social network regions of the brain associated with social and norm violating behavior, which is in line with the proposed account of procedural fairness as a means to evaluate leaders' WTRs.

Third, the effects of procedural information constitute a two-stage process in which the information serves to determine the trustworthiness of leaders and only then gives meaning to the decision implemented (Bøggild, 2014; Grimes, 2006;

shackelf@oakland.edu

Tyler, 1990). That is, procedural features are first and foremost used to form representations of the dispositions of the *decision-maker* and only secondly to form representations of the *decision*. Research based on fairness heuristic theory has provided further evidence for this notion demonstrating that information on procedural fairness only affects group members' satisfaction with a given decision when the trustworthiness of the decision-maker is ambiguous while such information has little effect when trustworthiness of the decision-maker is known and certain (Van den Bos, Wilke, & Lind, 1998). In a similar study, Van Dijke and Verboon (2010) demonstrated how information about procedural fairness has no effects when trust in the decision-maker is already high. In sum, these studies indicate that people attend to information on procedural fairness to gain a representation of the cooperative dispositions of leaders when these dispositions are ambiguous. Procedural information is not so much attended to in order to evaluate a single decision but is attended to in order to predict future behavior of the leader by evaluating his or her stable dispositions.

Fourth, if a follower uses procedural information to estimate the WTR of a leader toward the self, this information should also influence a broader range of intrapsychic factors such as self-esteem and perceived standing. Conceptually, we can think of self-esteem as a monitoring device of the fitness-related success of the individual in terms of maintaining viable social relationships (having social allies/friends, attracting mates, etc.) (Leary & Baumeister, 2000; Leary & Downs, 1995). According to Hill and Buss (2006, p. 350), self-esteem functions as "an internal gauge designed to monitor individuals' successes in interpersonal relationships and, particularly the degree to which they are being included or excluded from social groups, and to motivate corrective actions when one's level of social inclusion gets dangerously low." From this perspective, we should expect information on procedural fairness to serve as relevant inputs to this system by cueing whether the individual engages in a viable, profitable relationship with the group leader. Cross-sectional and experimental studies have demonstrated that individuals experiencing unfair relative to fair procedures (e.g., not being allowed to voice their opinions in group decision) report lower levels of self-esteem (De Cremer et al., 2005; Koper et al., 1993). Additionally, procedural fairness influences self-reported well-being (Schmitt & Dörfel, 1999) and mental health of individuals (Beijersbergen, Dirkzwager, Eichelsheim, van der Laan, & Nieuwbeerta, 2014).

Fifth, if procedural information is used as a measuring stick to gauge the WTR of a leader, it should also serve to regulate individuals' willingness to sacrifice their own welfare for the group. If unfair procedures signal that a leader is unlikely to let future group resources flow in the direction of the individual, it would be an inferior strategy in fitness terms to remain committed to and invest in group interests and efforts in the face of such cues. There is substantial evidence that leader behavior that violates procedural fairness criteria lowers the efforts that the individual puts into communal efforts. In particular, studies have demonstrated that procedural fairness is an important component in inducing "organizational citizenship" (Blader & Tyler, 2003; Folger & Konovsky, 1989; Sweeney & McFarlin, 1993; Tyler & Degoey, 1995). In the workplace, for example, people experiencing fair procedures

shackelf@oakland.edu

are more committed to the company and its goals (Lavelle et al., 2009), report higher job satisfaction (Pillai, Schriesheim, & Williams, 1999), and are less inclined to switch jobs (Daileyl & Kirk, 1992). Moreover, scholars have argued that procedural fairness is a key aspect of transformational leadership because fair procedures make group members committed to the goals and success of the group (Pillai et al., 1999). In sum, information on procedural fairness provides the individual with reliable information on the WTR of leaders and their actions and motivates withdrawal from communal efforts when such information indicates that future resources are unlikely to flow in the direction of the individual.

Procedural Fairness as a Bargaining Tool

When a follower identifies a leader with a low WTR toward the self, this signals an important adaptive problem. Under ancestral circumstances, in which the individual was crucially dependent on social support, it would entail significant fitness benefits if this problem could be solved. A strategy would be to leave the group, but, again due to the importance of social support, this strategy would often be a last resort. In many situations, the fitness benefits of staying in a group with an exploitative leader would outweigh the benefits of voluntary ostracism. Another more viable strategy, which we argue is crucial for the understanding of the psychology of procedural fairness, would be to seek to replace the leader with a leader with a higher WTR or, through the threat of replacement, upregulate the current leader's WTR.[7] Essentially, these outcomes entail a political renegotiation of the hierarchy within the collective.

As argued above, the crucial tool for hierarchy negotiations is the formation and mobilization of alliances. While an exploitative group leader should easily be able to ignore a single individual's efforts to counteract exploitation, such reactions should be harder to bypass when supported by a coalition of followers. Anthropological studies have shown how followers engage in political alliances of

[7] As described in footnote 5, it is possible to distinguish between intrinsic and monitored WTRs. In principle, it is adaptive for a follower to seek to increase both WTRs of a leader. However, because intrinsic WTRs are set by factors related to kinship and deep friendship, it is difficult to recalibrate these to a significant extent. Hence, the bargaining strategies we describe here refer to attempts to upregulate monitored WTRs. In relation to intrinsic WTRs, the task of a follower is instead to actively disengage the intrinsic WTRs of a leader from the leaders' decision-making process, that is, to ensure that the private affiliations of leaders do not shape their decisions. This will be the case for decision-making processes that conform to procedural fairness criteria. In fact, within the conceptual framework of WTRs, one can define a leader's decision as impartial to the extent it is uncorrelated with the leader's WTRs. It should also be noted that impartiality is often not in itself an adaptive target for the individual follower. Each follower holds an interest in increasing the WTRs of the leader toward the self to the largest extent possible. However, since upregulating a leader's WTR requires coalitional coordination, the only arrangement equally powerful followers can coordinate on is to opt for equal treatment of individuals (i.e., disengage intrinsic WTRs all together and upregulate monitored WTRs to the same average for all followers).

shackelf@oakland.edu

moral communities to counteract dominant behavior through leveling mechanisms of gossip, criticism, ridicule, disobedience, and, in the last instance, banishment or assassination of out-of-line leaders (Boehm, 1993, 1999). Moreover, computer simulations suggest that such mobilization of coalitions against exploitative leaders could be adaptive under ancestral conditions (Gavrilets, 2012; Gavrilets, Duenez-Guzman, & Vose, 2008).

We argue that the transition from the problem of identifying leaders with a low WTR toward the self to the problem of mobilizing others to deal with that problem has shaped the psychology of procedural fairness in important ways. It constitutes a key, separate evolved function of procedural fairness. In particular, we argue that it is impossible to understand the *moralistic* nature of attention to procedural features without understanding the need for mobilization.

Let us consider a follower, named A, that suffers costs because a leader, X, has a low WTR toward A. Only under two conditions would it make adaptive sense for a set of other followers—B, C, and D—to aid A against X. The first condition is if B, C, and D had a high WTR toward A, that is, if they valued A's welfare because of A's value as a reciprocation partner or kin. In such a case, the costs imposed on A would yield indirect costs on B, C, and D, and, hence, X's exploitation would constitute a shared problem. Under ancestral conditions with dense social networks, this condition might indeed have occurred with some frequency. At the same time, it should be noted that only a fraction of the costs imposed on A are experienced by B, C, and D (specifically, the costs imposed on A weighted by how much they each value A's welfare) and, hence, B, C, and D will be less motivated than ideally seen from the perspective of A. The second condition, however, would provide stronger motivations to aid A. If X did not just hold low WTRs toward A but also low WTRs toward B, C, and D, these other followers should be as motivated to aid A against X as A would be to get X away from the top position. From a mobilization perspective, if A detects that X holds a low WTR, it is therefore in the interest of A to broadcast information that can convince other followers that they themselves are in the same situation: facing an exploitative leader. Because of this, the selection pressures that have structured our psychology of procedural fairness should have implemented a motivation to scan for those precise procedural features that would signify that not only the self but also others are being exploited and, upon detection, broadcast these widely.

This, we argue, has structured the content of intuitions of procedural fairness. Intuitions of procedural fairness tag processes that lead to negative outcomes for the self as immoral and unfair to the extent these lack impartiality and that not all relevant parties are involved in a balanced way and are able to voice their opinion. Attending to and moralizing these particular features is, seen from the perspective of an individual follower, crucial because they signify that the leader holds low WTRs toward a larger segment of the collective and helps that individual mobilize others on his or her behalf. In this way, the psychology of procedural fairness directs attention toward the specific information that will make it most likely that others will come to the aid of the individual.

Intuitions about procedural fairness serve as an effective coordination device for group members with coinciding interests to mobilize their efforts and form political

shackelf@oakland.edu

coalitions with the power to counteract exploitation. By insisting that everyone gets a voice in authoritative decision-making and that the group leader has no personal interest in the decision (along with the other criteria displayed in Fig. 1), followers hold an effective means for keeping exploitative leaders in check. Specifically, by enforcing procedural fairness criteria in authoritative decision-making, each follower can make sure that they are not receiving disproportionally worse treatment than others. Of course, for each individual follower, it would be better to receive disproportionally *better* treatment, but, given the constraints emerging from the need for coalitional support, impartiality provides a feasible, second-best option.

These observations do not imply that a follower does not react negatively to negative outcomes that stems from processes that live up to procedural fairness criteria (people also care about outcome favorability and distributive justice, see, e.g., Törnblom & Vermunt, 2007). The reaction, however, only becomes moralistic when others can be expected to mobilize on behalf of the follower. As argued forcefully by DeScioli and Kurzban (2013), morality is specifically designed for coordinating collective responses and our moral psychology has a suite of design features that enable it to fulfill this function (Petersen, 2013). These include that (1) the detection of behavior that violates moral rules elicit feelings of moral outrage (Tetlock, 2003); (2) outrage motivates the imposition of costs (social and/or physical) on those committing them and, hence, serve to incentivize moral behavior (Jensen & Petersen, 2011); (3) outrage also creates a motivation to spread information about the perpetrator and the offense (e.g., through gossip), hence helping spread the feeling of outrage; (4) the motivation to spread information is sustained by a belief that moral intuitions are universally shared and, hence, that others will respond in kind (Turiel, 1983); and (5) the content of moral intuitions and rules are explained with reference to a collective good that makes them (appear) in the interest of the majority and, hence, further facilitating the spread of outrage. Intuitions of procedural fairness match this template.

We emphasize that attention to procedural fairness is strategic in nature. It evolved to increase fitness of the individual and not to help the group at large. As described in detail in the preceding section, advocating for procedural fairness and morally condemning exploitative group leaders is exercised by group members who themselves are subject to exploitation, while those unaffected or benefitting from exploitative or nepotistic leaders should be less inclined to condemn or counteract such behavior. These findings demonstrate that people, at least in part, broadcast notions of procedural fairness strategically when they have an incentive to upregulate WTRs of leaders in group decision-making. Put differently, there is an instrumental "brake" on moral outrage against leaders breaking procedural fairness criteria in situations where such outrage would not serve to increase the fitness of the individual. At the same time, it should be noted that the only rules that everybody can agree upon are rules that provide people with equal status in the decision-making process. If power is equally dispersed in a group, the de facto outcome of everybody's strategic moralization (i.e., trying to mobilize others for the rules that make the self best off) will therefore be a rule that makes everybody equal. Something that looks as if it was implemented for the benefit of everybody can in

shackelf@oakland.edu

this way emerge from the collective coordination of self-interest. In sum, intuitions about procedural fairness serve as a coordination device for group members to form political coalitions and impose criteria in group decision-making that serve to upregulate WTRs of leaders toward the self.

Evolutionary Mismatches: The Evolved Functions of Procedural Fairness and Modern Politics

The ultimate explanation of procedural fairness, we have argued, relates to the politics of ancestral groups. Our psychology of procedural fairness is designed to solve the complex and multilayered adaptive problems that arise when political power is negotiated. Specifically, intuitions of procedural fairness help us to safely delegate political power to others in return for services by providing a cognitive and motivational machinery that constantly monitors whether these leaders continue to be favorably disposed toward us and orchestrate collective outrage against those who are not.

The empirical evidence that we have reviewed above has mainly been conducted in settings that are not political in the everyday sense of this term: small-scale, interpersonal settings such as in legal, managerial, and educational settings. At the same time, these smaller social settings are in many ways more parallel to the political settings of our ancestors than are mass politics settings. Humans have lived in small-scale, simple hunter–gatherer societies of no more than 50–150 individuals for nearly 2.5 million years. State-liked societies with potential for mass politics only emerged extremely recently in evolutionary terms (Diamond, 1999; Wade, 2007).

This is important because evolution is a slow and protracted process. Psychological mechanisms shaped by natural selection—such as those underlying notions of procedural fairness—should therefore be adapted to function within the ancestral, small-scale environment in which they evolved rather than the present environments (Tooby & Cosmides, 1990). A consequence is that differences between ancestral and modern environments (sometimes referred to as mismatches) imply that the operations of psychological mechanisms might not always be fitness enhancing under modern conditions (Hagen & Hammerstein, 2006). One prominent example concerns the innate human fear of spiders and snakes, which still, although not posing a significant threat in a modern context, is perceived as more dangerous than contemporary dangers like electric outlets or guns (Ohman & Mineka, 2001). Such a mismatch also manifests itself in modern political attitude formation (Petersen & Aarøe, 2012) and in evaluations of modern-day political leaders (Van Vugt et al., 2008).

In this final section, we review evidence that suggest that a mismatch between ancestral and modern conditions is also shaping the way intuitions about procedural fairness enter judgments and behaviors in the context of modern politics. As anthropologists Erdal and Whiten (1996, p. 158) also note, ancestral psychological mechanisms designed to counteract exploitative group leaders "need not in principle lead

shackelf@oakland.edu

to behavior which is functionally effective" in a modern context since the "correlation between the proximate cause of the behavior and the genetic function of the behavior may be broken in a radically different environment." Specifically, we suggest that a number of cues that draw attention away from violations of procedural fairness in small-scale settings are lacking or are less vivid in the context of large-scale politics. In other words, for reasons provided below we should expect there to be no instrumental "brake" on moral outrage following a breach of procedural fairness criteria in a political context. Accordingly, modern politics come to be shaped by procedural fairness even more strongly than many other contexts.

Evidence from political science indicates that notions of procedural fairness also play an important role in politics today. Recent public opinion research has demonstrated that notions of procedural fairness influence trust in a variety of modern political institutions—such as government, national courts, administrative units, political parties, and political representatives—such that institutions that conform to our intuitions of procedural fairness receive more trust (Bøggild, 2014; Grimes, 2006; Ramirez, 2008; Tyler, 1994). Political policies that are passed after processes that conform to our intuitions of procedural fairness are also supported to a higher extent (Bøggild, 2014). Furthermore, a number of studies on political communication suggest that media audiences desire information about the strategic and self-interested motivations of politicians and that news editors happily provide such coverage. For example, a recent study has demonstrated a "news-selection bias" in which people seek strategic news covering self-interested motivations and actions of politicians over standard issue-related coverage of politics (Trussler & Soroka, 2014). At the same time, such coverage is easy and inexpensive to produce and is therefore willingly supplied by political analysts and news editors (Farnsworth & Lichter, 2010; Patterson, 1994). In addition, negative campaigns from politicians vigorously question the moral quality and dispositional tendencies of their political opponents (Lau & Rovner, 2009). In short, information about exploitative and self-serving tendencies of modern political leaders is both high in supply and demand.

Procedural fairness is also of obvious importance in small-scale social settings, but, as reviewed in the preceding sections, it is a standard finding that individuals pay less attention to procedural information when the relevant decisions do not impose costs on these individuals. According to three recent studies, this instrumental "brake" on procedural fairness is not operating in mass politics. Two studies have considered whether attention to procedural information interacts with the favorability of the outcome in a political context. One study (Bøggild, 2014) investigated the effects of impartiality on trust in political decision-makers and support for public policies. More specifically, the study showed through three survey experiments that trust in a political decision-maker and support for its political decision was in part a function of whether or not the decision-maker had a personal, vote-maximizing, and ultimately reelection motive in introducing the decision. However, contrary to the extensive literature in social psychology, these effects were not moderated by outcome favorability. Across three different studies with different operationalizations of outcome favorability, the effects of impartial decision-making were constant and substantial, independently of whether outcomes were favorable to the individual or

shackelf@oakland.edu

not. A second study (Kumlin, 2004) demonstrated a significant correlation between citizens' perceived voice opportunities in political decision-making processes and general trust in politicians. This effect, however, did not differ across individuals expressing different levels of satisfaction with the services they received from the state. This study concluded that in a political context there is "very little support for the hypothesis that the impact of voice grows with service dissatisfaction" (Kumlin, 2004, p. 263). A third study considered whether attention to procedural information interacted with whether the individual was influenced by an outcome. Grimes (2005) used panel data to show that trust in a political, administrative unit and support for its political decision depended on two procedural criteria in terms of perceived voice opportunities and the decision-maker's willingness to justify its decision publicly. However, this study found little support for the expectation that these effects were larger among those affected by the decision. The effects of public justification on both trust in the administrative unit and support for its decision were constant across those affected and unaffected by the decision. The effect of voice opportunities on trust in the administrative unit was also constant across these segments, while there was a tendency for the effect of voice opportunities on support for the political decision to be marginally larger among those affected by the decision. That is, out of four interaction terms, only one was marginally significant, which testifies to a limited, if any, moderating effect of outcome dependence in a large-scale political setting. In the context of mass politics, people assign weight to procedural fairness concerns independently of whether they were affected by the political decision or not (Grimes, 2005, p. 131).

The finding that personal self-interest is less effective in regulating psychological responses in mass politics than in other contexts is consistent with both classical research in political science (Sears & Funk, 1991) and research from an evolutionary perspective (Petersen & Aarøe, 2012). Given the contextual differences between modern politics and ancestral environments, this attenuated role of self-interest is not unexpected. Multiple differences add to this outcome: First, the novel and technical nature of many political issues makes it difficult for our psychology to compute the costs and benefits of various policies for our welfare on the basis of evolutionarily recurrent information. Second, the length of the causal link between a policy decision and its personal consequences implies that many people primarily ascribe negative outcomes to personal rather than political choices (see Feldman, 1982). Third, our evolved decision-making psychology—designed to operate on the basis of the vivid, social cues that inhere in face-to-face interaction—has less clear cues available in the abstract context of mass politics: most people won't know and will never meet most of those affected by policy decisions (Petersen & Aarøe, 2013). All these processes add up and make people perceive politics as abstract, remote, and disconnected from their everyday lives (Lane, 1962; Lippmann, 1922; Schumpeter, 1943). Because of this, politics can easily be psychologically represented as an arena for cheap signals of one's moral commitments, an arena where one can send a signal without paying the costs of forfeiting one's self-interest. Rather than self-interest, symbols such as those related to procedural fairness

shackelf@oakland.edu

become a key factor when people form political opinions (Herek, 1986; Inglehart, 2008; Sears & Funk, 1991).

In modern politics notions of procedural fairness are important. In fact, because of evolutionarily mismatches between ancestral and modern environments, these notions potentially gain more importance today than ancestrally. Because of the difficulties in connecting self-interest and politics, there is no instrumental "brake" on moral outrage against group leaders in modern, mass contexts. Aided by mass and social media, moralization can "run free" without a direct connection to the instrumental, fitness-related goals that these moral intuitions originally evolved to fulfill. In mass politics, as suggested by the recent empirical studies, procedural fairness intuitions might not merely manifest themselves strategically in situations where they serve a clear fitness-related purpose but affect opinions in a more universal and unconditional manner. Thus, in a political context in particular "the means *is* the end" (Hibbing & Theiss-Morse, 2001, p. 243) and constitutes important "procedural goods" to citizens of democratic societies (Lane, 1988).

Conclusion

In this chapter, we argued that intuitions of procedural fairness reflect a psychological machinery that evolved by natural selection to help our ancestors solve adaptive problems related to politics. This evolutionary perspective on procedural fairness contributes to the existing literature in three important ways.

First, it offers new theoretical insights by arguing that notions of procedural fairness serve both (1) a diagnostic function in evaluating WTRs and exploitative tendencies of group leaders and (2) a bargaining function in terms of mobilizing support against exploitative leaders and bargain for better outcomes. As stressed throughout the chapter, an evolutionary perspective does not discredit existing models offered by social psychologists but rather allows for a synthesis between these more proximate models. From an evolutionary perspective, information on procedural fairness does serve as a means for the individual to evaluate its relationship with the group authority and status within the group (in line with *the relational model of authority*). However, attention to these cues does not ultimately stem from a need to construct a social identity but rather serves to reduce uncertainties as to whether an authority can be trusted or will exploit followers for self-interested reasons (in line with *fairness heuristic theory*). Moreover, notions of procedural fairness serve a direct instrumental purpose as an effective means for bargaining for better outcomes (in line with *the instrumental model*). Hence, these theoretical models have provided and will continue to provide useful insights on the effects and psychological function of procedural fairness. However, an evolutionary perspective, we argue, provides an important organizing principle and metatheoretical paradigm for synthesizing existing theoretical models and empirical findings and for further hypothesis generation. The existing literature has mainly (but not exclusively) focused on procedural fairness as a diagnostic tool, and future work could

shackelf@oakland.edu

benefit from considering how notions of procedural fairness are used as a strategic bargaining tool and as a coordination device for followers with coinciding interests to engage in moral communities and enforce moral constraints on the discretionary power of group leaders.

Second, an evolutionary perspective offers potential conceptual advancements to the procedural fairness literature. In spite of the extensive empirical support for the effects of procedural fairness, conceptual work in this literature remains "in its early stages" (Cloutier & Vilhuber, 2008, p. 714). The criteria displayed in Fig. 1 are most often referred back to Leventhal (1980), who provides no account for how or from where these criteria are deduced. As a result, although most would probably concur with the fairness of the criteria displayed in Fig. 1, we are still left unaware of what "fair" actually means. An evolutionary perspective, however, provides an ultimate account of procedural fairness and hence allows for a systematic deduction of relevant procedural criteria or principles. An evolutionary perspective entails a functionalist approach to procedural fairness, analyzing the concept from the starting point of the fundamental need or problem it was designed to accommodate. Hence, *any non-outcome information related to authoritative decision-making that would somehow serve as a reliable cue for WTRs of leaders should be relevant components in procedural fairness evaluations.* Arguably, all of the existing procedural criteria displayed in Fig. 1 should fit under this simple conceptualization. For example, nature should have selected against individuals who did not respond with moral outrage and anger when group leaders refused them a say in group decisions (voice) or had vested interests in the decisions introduced (impartiality) (for similar naturalistic conceptualizations of *distributive* fairness, see Binmore, 2011; Krebs, 2008).

Third, an evolutionary perspective provides a possible account of recent divergent findings on procedural fairness in social psychological and political settings and stresses important differences between social and political cognition more generally. In a political context—in contrast to small-scale, interpersonal settings such as in the courtroom or in the workplace—people often fail to see the connection between their opinions or behaviors and their immediate self-interest. Hence, in a political context, moralization and fairness intuitions "run free" and become detached from the instrumental function they originally evolved to fulfill. Similar findings have been demonstrated on other political issues as well such as on public opinion toward immigration. Specifically, scholars have demonstrated that the evolved psychological system of disgust, which originally evolved to help the individual avoid contact with pathogens, is an important factor in explaining attitudes toward immigration. Even though catching an infection from interacting with individuals belonging to a different ethnic group in a modern-day context would be extremely low, priming people with disgusting images or smells still drives up anti-immigration attitudes significantly (Faulkner, Schaller, Park, & Duncan, 2004). Hence, evolved disease-avoidance mechanisms play an important role in public opinion formation although they may not serve the survival-based function they originally evolved to fulfill. Future work should consider more systematically how evolved cognitive systems might "run free" and become detached from their survival-based function in a large-scale modern political context.

shackelf@oakland.edu

In closing, an evolutionary perspective suggests that we should also be cautious about the potential downsides of the effects of procedural fairness. Humans might not only be equipped with notions of procedural fairness in order to defend against exploitation but also with counter-adaptations designed to circumvent such defenses in others (Buss & Duntley, 2008; Whiten & Byrne, 1997). For example, leaders might attempt to *appear* procedurally fair (e.g., by allowing group members voice) to avoid being held accountable for risky or poor decision-making (DeScioli & Bokemper, 2014; MacCoun, 2005). This implies that procedural fairness could also be applied in a manipulative sense to conceal incompetence or self-serving dispositions in a leadership role.

References

Aalberg, T., Strömbäck, J., & de Vreese, C. H. (2011). The framing of politics as strategy and game: A review of concepts, operationalizations and key findings. *Journalism, 13*(2), 1464884911427799. doi:10.1177/1464884911427799.

Beijersbergen, K. A., Dirkzwager, A. J. E., Eichelsheim, V. I., van der Laan, P. H., & Nieuwbeerta, P. (2014). Procedural justice and prisoners' mental health problems: A longitudinal study. *Criminal Behaviour and Mental Health, 24*(2), 100–112. doi:10.1002/cbm.1881.

Beugré, C. D. (2009). Exploring the neural basis of fairness: A model of neuro-organizational justice. *Organizational Behavior and Human Decision Processes, 110*(2), 129–139.

Binmore, K. (2011). *Natural justice*. Oxford, UK: Oxford University Press.

Blader, S. L., & Tyler, T. R. (2003). A four-component model of procedural justice: Defining the meaning of a 'Fair' process. *Personality & Social Psychology Bulletin, 29*(6), 747–758. doi:10.1177/0146167203029006007.

Boehm, C. (1993). Egalitarian behavior and reverse dominance hierarchy. *Current Anthropology, 34*(3), 227. doi:10.1086/204166.

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

Bøggild, T. (2014). *Do the means justify the ends? Procedural fairness, politicians' reelection motives, and support for public policy*. Paper presented at the annual meeting of the Midwest political science association, Chicago, 2014.

Boyd, R., & Richerson, P. J. (1988). *Culture and the evolutionary process*. Chicago: University of Chicago Press.

Brockner, J., & Wiesenfeld, B. M. (1996). An integrative framework for explaining reactions to decisions: Interactive effects of outcomes and procedures. *Psychological Bulletin, 120*(2), 189–208.

Buss, D. M., & Duntley, J. D. (2008). Adaptations for exploitation. *Group Dynamics: Theory, Research, and Practice, 12*(1), 53–62. doi:10.1037/1089-2699.12.1.53.

Cappella, J. N., & Jamieson, K. H. (1996). News frames, political cynicism, and media cynicism. *The Annals of the American Academy of Political and Social Science, 546*(1), 71–84. doi:10.1177/0002716296546001007.

Charness, G., & Levine, D. (2007). Intention and stochastic outcomes: An experimental study. *Economic Journal, 117*(522), 1051–1072. doi:10.2139/ssrn.437870.

Cloutier, J., & Vilhuber, L. (2008). Procedural justice criteria in salary determination. *Journal of Managerial Psychology, 23*(6), 713–740. doi:10.1108/02683940810894765.

Colquitt, J. A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology, 86*(3), 386–400. doi:10.1037/0021-9010.86.3.386.

shackelf@oakland.edu

Cosmides, L., & Tooby, J. (1997). Dissecting the computational architecture of social inference mechanisms. In *Characterizing human psychological adaptations (Ciba Foundation Symposium #208)* (pp. 132–156). Chichester, UK: Wiley.

Cremer, D. D., & van Knippenberg, D. (2003). Cooperation with leaders in social dilemmas: On the effects of procedural fairness and outcome favorability in structural cooperation. *Organizational Behavior and Human Decision Processes, 91*(1), 1–11. doi:10.1016/S0749-5978(02)00539-3.

Daileyl, R. C., & Kirk, D. J. (1992). Distributive and procedural justice as antecedents of job dissatisfaction and intent to turnover. *Human Relations, 45*(3), 305–317. doi:10.1177/001872679204500306.

De Cremer, D., van Knippenberg, B., van Knippenberg, D., Mullenders, D., & Stinglhamber, F. (2005). Rewarding leadership and fair procedures as determinants of self-esteem. *The Journal of Applied Psychology, 90*(1), 3–12. doi:10.1037/0021-9010.90.1.3.

Delton, A. W., Cosmides, L., Guemo, M., Robertson, T. E., & Tooby, J. (2012). The psychosemantics of free riding: Dissecting the architecture of a moral concept. *Journal of Personality and Social Psychology, 102*(6), 1252–1270. doi:10.1037/a0027026.

DeScioli, P., & Bokemper, S. (2014). Voting as a counter-strategy in the blame game. *Psychological Inquiry, 25*(2), 206–214. doi:10.1080/1047840X.2014.901130.

DeScioli, P., & Kurzban, R. (2013). A solution to the mysteries of morality. *Psychological Bulletin, 139*(2), 477–496. doi:10.1037/a0029065.

De Waal, F. B. (1997). *Good natured the origins of right and wrong in humans and other animals*. Cambridge, MA: Harvard University Press.

De Waal, F. B. (2007). *Chimpanzee politics: Power and sex among apes*. Baltimore: JHU Press.

Diamond, J. M. (1999). *Guns, germs and steel: The fates of human societies*. New York: Norton.

Dubreuil, B. (2011). *Human evolution and the origins of hierarchies: The state of nature*. New York: Cambridge University Press.

Dulebohn, J. H., Conlon, D., Sarinopoulos, I., Davison, R., McNamara G. (2009). The biological bases of unfairness: Neuroimaging evidence for the distinctiveness of procedural and distributive justice. *Organizational Behavior and Human Decision Processes, 110*(2), 140–151

Easton, D. (1953). *An inquiry into the state of political sciences*. New York: Knopf.

Erdal, D., & Whiten, A. (1996). Egalitarian and Machiavellian intelligence in human evolution. In P. Mellars & K. Gibson (Eds.), *Modelling the human mind*. Cambridge, UK: McDonald Institute Monographs.

Falk, A., Fehr, E., & Fischbacher, U. (2008). Testing theories of fairness—Intentions matter. *Games and Economic Behavior, 62*(1), 287–303. doi:10.1016/j.geb.2007.06.001.

Farnsworth, S. J., & Lichter, R. S. (2010). *The nightly news nightmare: Media coverage of U.S. Presidential Elections, 1988–2008* (3rd ed.). Lanham, MD: Rowman & Littlefield Publishers.

Faulkner, J., Schaller, M., Park, J. H., & Duncan, L. A. (2004). Evolved disease-avoidance mechanisms and contemporary xenophobic attitudes. *Group Processes & Intergroup Relations, 7*(4), 333–353. doi:10.1177/1368430204046142.

Feldman, S. (1982). Economic self-interest and political behavior. *American Journal of Political Science, 26*(3), 446–466. doi:10.2307/2110937.

Foley, R. (1995). The adaptive legacy of human evolution: A search for the environment of evolutionary adaptedness. *Evolutionary Anthropology: Issues, News, and Reviews, 4*(6), 194–203. doi:10.1002/evan.1360040603.

Folger, R., & Konovsky, M. A. (1989). Effects of procedural and distributive justice on reactions to pay raise decisions. *The Academy of Management Journal, 32*(1), 115–130.

Fry, P. S., & Corfield, V. K. (1983). Children's judgments of authority figures with respect to outcome and procedural fairness. *The Journal of Genetic Psychology, 143*(2), 241–250. doi:10.1080/00221325.1983.10533557.

Gavrilets, S. (2012). On the evolutionary origins of the egalitarian syndrome. *Proceedings of the National Academy of Sciences*, August, 201201718. doi:10.1073/pnas.1201718109.

shackelf@oakland.edu

Gavrilets, S., Duenez-Guzman, E. A., & Vose, M. D. (2008). Dynamics of alliance formation and the egalitarian revolution. *PLoS One, 3*(10), e3293. doi:10.1371/journal.pone.0003293.

Giacobbe-Miller, J. K., Miller, D. J., & Victorov, V. I. (1998). A comparison of Russian and U.S. pay allocation decisions, distributive justice judgments, and productivity under different payment conditions. *Personnel Psychology, 51*(1), 137–163. doi:10.1111/j.1744-6570.1998.tb00719.x.

Gold, L. J., Darley, J. M., Hilton, J. L., & Zanna, M. P. (1984). Children's perceptions of procedural justice. *Child Development, 55*(5), 1752–1759. doi:10.2307/1129922.

Gonzalez, C. M., & Tyler, T. R. (2007). Why do people care about procedural fairness? The importance of membership monitoring. In *Distributive and procedural justice* (2007th edn., pp. 91–111). Abingdon, UK: Ashgate.

Grimes, M. (2005). *Democracy's infrastructure: The role of procedural fairness in fostering consent* (Doctoral thesis). http://gup.ub.gu.se/publication/31280-democracys-infrastructure-the-role-of-procedural-fairness-in-fostering-consent.

Grimes, M. (2006). Organizing consent: The role of procedural fairness in political trust and compliance. *European Journal of Political Research, 45*(2), 285–315. doi:10.1111/j.1475-6765.2006.00299.x.

Hagen, E. H., & Hammerstein, P. (2006). Game theory and human evolution: A critique of some recent interpretations of experimental games. *Theoretical Population Biology, 69*(3), 339–348. doi:10.1016/j.tpb.2005.09.005.

Haidt, J. (2013). *The righteous mind: Why good people are divided by politics and religion*. New York: Vintage Books.

Ham, J., & van den Bos, K. (2008). Not fair for me! The influence of personal relevance on social justice inferences. *Journal of Experimental Social Psychology, 44*, 699–705.

Henrich, J., & Gil-White, F. J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior, 22*(3), 165–196. doi:10.1016/S1090-5138(00)00071-4.

Herek, G. M. (1986). The instrumentality of attitudes: Toward a neofunctional theory. *Journal of Social Issues, 42*(2), 99–114. doi:10.1111/j.1540-4560.1986.tb00227.x.

Hibbing, J. R., & Alford, J. R. (2004). Accepting authoritative decisions: Humans as wary cooperators. *American Journal of Political Science, 48*(1), 62–76. doi:10.1111/j.0092-5853.2004.00056.x.

Hibbing, J. R., & Theiss-Morse, E. (2001). *What is it about government that Americans dislike?* Cambridge, UK: Cambridge University Press.

Hibbing, J. R., & Theiss-Morse, E. (2008). Voice, validation, and legitimacy. In *Cooperation: The political psychology of effective human interaction* (pp. 123–142). Malden, MA: Blackwell.

Hicks, A. J., & Lawrence, J. A. (1993). Children's criteria for procedural justice: Developing a young people's procedural justice scale. *Social Justice Research, 6*(2), 163–182. doi:10.1007/BF01048475.

Hill, S. E., & Buss, D. M. (2006). The evolution of self-esteem. In M. H. Kernis (Ed.), *Self-esteem: Issues and answers*. New York: Psychology Press.

Inglehart, R. F. (2008). Changing values among Western publics from 1970 to 2006. *West European Politics, 31*(1–2), 130–146. doi:10.1080/01402380701834747.

Jensen, N., & Petersen, M. (2011). To defer or to stand up? How offender formidability affects third party moral outrage. *Evolutionary Psychology, 9*(1), 118–136.

Kenrick, D. T., Li, N. P., & Butner, J. (2003). Dynamical evolutionary psychology: Individual decision rules and emergent social norms. *Psychological Review, 110*(1), 3–28.

Khatri, N., Fern, C. T., & Budhwar, P. (2001). Explaining employee turnover in an Asian context. *Human Resource Management Journal, 11*(1), 54–74. doi:10.1111/j.1748-8583.2001.tb00032.x.

Koper, G., Van Knippenberg, D., Bouhuijs, F., Vermunt, R., & Wilke, H. (1993). Procedural fairness and self-esteem. *European Journal of Social Psychology, 23*(3), 313–325. doi:10.1002/ejsp.2420230307.

Krebs, D. L. (2008). The evolution of a sense of justice. In J. Duntley & T. K. Shackelford (Eds.), *Evolutionary forensic*. New York: Oxford University Press.

shackelf@oakland.edu

Krislov, S. (1991). Why do people obey the law? *Journal of Public Administration Research and Theory, 1*(2), 235–239.
Kumlin, S. (2004). *The Personal and the Political: How Personal Welfare State Experiences Affect Political Trust and Ideology*. : New York: Palgrave Macmillan.
Lane, R. E. (1962). *Political ideology: Why the American common man believes what he does*. New York: Free Press of Glencoe.
Lane, R. E. (1988). Procedural goods in a democracy: How one is treated versus what one gets. *Social Justice Research, 2*(3), 177–192. doi:10.1007/BF01054555.
Lasswell, H. (1950). *Politics: Who gets what, when, how*. New York: P. Smith.
Lau, R., & Rovner, I. (2009). Negative campaigning. *Annual Review of Political Science, 12*, 285–306.
Laustsen, L., & Petersen, M. B. (2015). Does a competent leader make a good friend? Conflict, ideology and the psychologies of friendship and followership. *Evolution and Human Behavior, 36*(4), 286–293 doi:10.1016/j.evolhumbehav.2015.01.001.
Lavelle, J. J., Brockner, J., Konovsky, M. A., Price, K. H., Henley, A. B., Taneja, A., et al. (2009). Commitment, procedural fairness, and organizational citizenship behavior: A multifoci analysis. *Journal of Organizational Behavior, 30*(3), 337–357. doi:10.1002/job.518.
Leary, M. R., & Baumeister, R. F. (2000). The nature and function of self-esteem: Sociometer theory. In M. Zanna (Ed.), *Advances in experimental social psychology*. San Diego, CA: Academic.
Leary, M. R., & Downs, D. L. (1995). Interpersonal functions of the self-esteem motive: The self-esteem system as a sociometer. In *Efficacy, agency, and self-esteem.* New York: Plenum.
Leung, K., Tong, K.-K., & Allan, E. (2007). Realpolitik versus fair process: Moderating effects of group identification on acceptance of political decisions. *Journal of Personality and Social Psychology, 92*(3), 476–489. doi:10.1037/0022-3514.92.3.476.
Leventhal, G. S. (1980). What should be done with equity theory? In *Social exchange: Advances in theory and research*. New York: Plenum.
Lind, E. A., Kanfer, R., & Christopher, P. (1990). Voice, control, and procedural justice: instrumental and noninstrumental concerns in fairness judgments. *Journal of Personality and Social Psychology, 59*(5), 952–959. doi:10.1037/0022-3514.59.5.952.
Lind, E. A., & Tyler, T. R. (1988). *The social psychology of procedural justice*. New York: Springer.
Lippmann, W. (1922). *Public opinion*. New York: Harcourt, Brace.
MacCoun, R. J. (2005). Voice, control, and belonging: The double-edged sword of procedural fairness. *Annual Review of Law and Social Science, 1*(1), 171–201. doi:10.1146/annurev.lawsocsci.1.041604.115958.
Meggitt, M. (1977). Blood is their argument. Palo Alto, CA: Mayfield.
Ohman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review, 108*(3), 483–522.
Ong, Q., Riyanto, Y. E., & Sheffrin, S. M. (2012). How does voice matter? Evidence from the ultimatum game. *Experimental Economics, 15*(4), 604–621.
Patterson, T. E. (1994). *Out of order*. New York: Vintage Books.
Petersen, M. (2015). Evolutionary political psychology. In D. Buss (Ed.), *Handbook of evolutionary psychology*. Hoboken, NJ: Wiley.
Petersen, M. B. (2013). Moralization as protection against exploitation: Do individuals without allies moralize more? *Evolution and Human Behavior, 34*(2), 78–85. doi:10.1016/j.evolhumbehav.2012.09.006.
Petersen, M. B., & Aarøe, L. (2012). Is the political animal politically ignorant? Applying evolutionary psychology to the study of political attitudes. *Evolutionary Psychology: An International Journal of Evolutionary Approaches to Psychology and Behavior, 10*(5), 802–817.
Petersen, M. B., & Aarøe, L. (2013). Politics in the mind's eye: Imagination as a link between social and political cognition. *American Political Science Review, 107*(2), 275–293. doi:10.1017/S0003055413000026.
Petersen, M., Sell, A., Tooby, J., & Cosmides, L. (2010). Evolutionary psychology and criminal justice: A recalibrational theory of punishment and reconciliation. In H. Høegh-Olesen (Ed.),

shackelf@oakland.edu

Human morality & sociality: Evolutionary & comparative perspectives (pp. 72–131). New York: Palgrave Macmillan.

Petersen, M. B., Sell, A., Tooby, J., & Cosmides, L. (2012). To punish or repair? Evolutionary psychology and lay intuitions about modern criminal justice. *Evolution and Human Behavior, 33*(6), 682–695. doi:10.1016/j.evolhumbehav.2012.05.003.

Pillai, R., Schriesheim, C. A., & Williams, E. S. (1999). Fairness perceptions and trust as mediators for transformational and transactional leadership: A two-sample study. *Journal of Management, 25*(6), 897–933. doi:10.1177/014920639902500606.

Pillai, R., Williams, E. S., & Tan, J. J. (2001). Are the scales tipped in favor of procedural or distributive justice? An investigation of the U.S., India, Germany, and Hong Kong (China). *International Journal of Conflict Management, 12*(4), 312–332. doi:10.1108/eb022861.

Platow, M. J., Eggins, R. A., Chattopadhyay, R., Brewer, G., Hardwick, L., Milsom, L., et al. (2013). Two experimental tests of relational models of procedural justice: Non-instrumental voice and authority group membership. *The British Journal of Social Psychology/The British Psychological Society, 52*(2), 361–376. doi:10.1111/j.2044-8309.2011.02083.x.

Popkin, S. L. (1991). *The reasoning voter: Communication and persuasion in presidential campaigns*. Chicago: University of Chicago Press.

Price, M. E., & Van Vugt, M. (2014). The evolution of leader-follower reciprocity: The theory of service-for-prestige. *Frontiers in Human Neuroscience, 8*, 363. doi:10.3389/fnhum.2014.00363.

Ramirez, M. D. (2008). Procedural perceptions and support for the U.S. Supreme Court. *Political Psychology, 29*(5), 675–698. doi:10.1111/j.1467-9221.2008.00660.x.

Schmitt, M., & Dörfel, M. (1999). Procedural injustice at work, justice sensitivity, job satisfaction and psychosomatic well-being. *European Journal of Social Psychology, 29*(4), 443–453. doi:10.1002/(SICI)1099-0992(199906)29:4<443::AID-EJSP935>3.0.CO;2-C.

Schumpeter, J. A. (1943). *Capitalism, socialism and democracy*. New York: Routledge.

Scott-Phillips, T. C., Dickins, T. E., & West, S. A. (2011). Evolutionary theory and the ultimate–proximate distinction in the human behavioral sciences. *Perspectives on Psychological Science, 6*(1), 38–47. doi:10.1177/1745691610393528.

Sears, D. O., & Funk, C. L. (1991). The role of self-interest in social and political attitudes. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 24, pp. 1–91). New York: Academic.

Sell, A., Tooby, J., & Cosmides, L. (2009). Formidability and the logic of human anger. *Proceedings of the National Academy of Sciences, 106*(35), 15073–15078. doi:10.1073/pnas.0904312106.

Sidanius, J., & Pratto, F. (2001). *Social dominance: An intergroup theory of social hierarchy and oppression*. Cambridge, UK: Cambridge University Press.

Smith, H. J., Tyler, T. R., Huo, Y. J., Ortiz, D. J., & Allan, E. (1998). The self-relevant implications of the group-value model: Group membership, self-worth, and treatment quality. *Journal of Experimental Social Psychology, 34*(5), 470–493. doi:10.1006/jesp.1998.1360.

Smith, K. B., Larimer, C. W., Littvay, L., & Hibbing, J. R. (2007). Evolutionary theory and political leadership: Why certain people do not trust decision makers. *The Journal of Politics, 69*(02), 285–299. doi:10.1111/j.1468-2508.2007.00532.x.

Sweeney, P. D., & McFarlin, D. B. (1993). Workers' evaluations of the 'Ends' and the 'Means': An examination of four models of distributive and procedural justice. *Organizational Behavior and Human Decision Processes, 55*(1), 23–40. doi:10.1006/obhd.1993.1022.

Takenishi, A., & Takenishi, M. (1992). Does commitment affect the meaning of fairness? Commonality and stability of fairness criteria in a political setting. *Social Justice Research, 5*(4), 415–429. doi:10.1007/BF01050757.

Tetlock, P. E. (2003). Thinking the unthinkable: Sacred values and taboo cognitions. *Trends in Cognitive Sciences, 7*(7), 320–324.

Thibaut, J. W., & Walker, L. (1975). *Procedural justice: A psychological analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates, Incorporated.

Tooby, J., & Cosmides, L. (1990). The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology and Sociobiology, 11*, 375–424.

shackelf@oakland.edu

Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 19–136). New York: Oxford University Press.

Tooby, J., Cosmides, L., & Price, M. E. (2006). Cognitive adaptations for N-Person exchange: The evolutionary roots of organizational behavior. *Managerial and Decision Economics, 27*(2-3), 103–129.

Tooby, J., Cosmides, L., Sell, A., Lieberman, D., & Sznycer, D. (2008). Internal regulatory variables and the design of human motivation: A computational and evolutionary approach. In A. Elliot (Ed.), *Handbook of approach and avoidance motivation* (pp. 251–271). Mahwah, NJ: Lawrence Erlbaum Associates.

Törnblom, K. Y., & Vermunt, R. (2007). *Distributive and procedural justice: Research and social applications*. Aldershot, UK: Ashgate Publishing Ltd.

Trussler, M., & Soroka, S. (2014). Consumer demand for cynical and negative news frames. *The International Journal of Press/Politics, 19*(3), 360–379. doi:10.1177/1940161214524832.

Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge, UK: Cambridge University Press.

Tyler, T. R. (1990). *Why people obey the law*. New Haven, CT/London: Yale University Press.

Tyler, T. R. (1994). Governing amid diversity: The effect of fair decisionmaking procedures on the legitimacy of government. *Law & Society Review, 28*(4), 809. doi:10.2307/3053998.

Tyler, T. R. (1997). The psychology of legitimacy: A relational perspective on voluntary deference to authorities. *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc, 1*(4), 323–345. doi:10.1207/s15327957pspr0104_4.

Tyler, T., Allan Lind, E., Ohbuchi, K.-I., Sugawara, I., & Huo, Y. J. (1998). Conflict with outsiders: Disputing within and across cultural boundaries. *Personality and Social Psychology Bulletin, 24*(2), 137–146. doi:10.1177/0146167298242003.

Tyler, T. R., & Caine, A. (1981). The influence of outcomes and procedures on satisfaction with formal leaders. *Journal of Personality and Social Psychology, 41*(4), 642–655. doi:10.1037/0022-3514.41.4.642.

Tyler, T. R., & Degoey, P. (1995). Collective restraint in social dilemmas: Procedural justice and social identification effects on support for authorities. *Journal of Personality and Social Psychology, 69*(3), 482–497. doi:10.1037/0022-3514.69.3.482.

Tyler, T. R., & Degoey, P. (1996). Trust in organizational authorities: The influence of motive attributions on willingness to accept decisions. In *Trust in organizations: Frontiers of theory and research* (1996th edn., Vol. 1996, pp. 331–357). Thousand Oaks, CA: Sage.

Tyler, T. R., & Lind, E. A. (1992). A relational model of authority in groups. *Advances in Experimental social psychology, 25*, 115–192.

Ulbig, S. G. (2002). Policies, procedures, and people: Sources of support for government? *Social Science Quarterly, 83*(3), 789–809. doi:10.1111/1540-6237.00115.

Ulbig, S. G. (2008). Voice is not enough the importance of influence in political trust and policy assessments. *Public Opinion Quarterly, 72*(3), 523–539. doi:10.1093/poq/nfn030.

Van den Bos, K., Lind, E. A., Vermunt, R., & Wilke, H. A. (1997). How do I judge my outcome when I do not know the outcome of others? The psychology of the fair process effect. *Journal of Personality and Social Psychology, 72*(5), 1034–1046.

Van den Bos, K., Wilke, H. A. M., & Lind, E. (1998). When do we need procedural fairness? The role of trust in authority. *Journal of Personality and Social Psychology, 75*(6), 1449–1458. doi:10.1037/0022-3514.75.6.1449.

Van den Bos, K., Wilke, H. A. M., Lind, E. A., & Vermunt, R. (1998). Evaluating outcomes by means of the fair process effect: Evidence for different processes in fairness and satisfaction judgments. *Journal of Personality and Social Psychology, 74*(6), 1493–1503. doi:10.1037/0022-3514.74.6.1493.

shackelf@oakland.edu

Van Dijke, M., De Cremer, D., & Mayer, D. M. (2010). The role of authority power in explaining procedural fairness effects. *Journal of Applied Psychology, 95*(3), 488–502. doi:10.1037/a0018921.

Van Dijke, M., & Verboon, P. (2010). Trust in authorities as a boundary condition to procedural fairness effects on tax compliance. *Journal of Economic Psychology, 31*(1), 80–91. doi:10.1016/j.joep.2009.10.005.

van Prooijen, J. -W. (2008). Egocentrism in procedural justice effects. *Advances in Group Processes, 25*(December), 29–54. doi:10.1016/S0882-6145(08)25002-2.

Van Vugt, M. (2006). Evolutionary origins of leadership and followership. *Personality and Social Psychology Review, 10*(4), 354–71. doi:10.1207/s15327957pspr1004_5.

Van Vugt, M., Ahuja, A., & Van Vugt, M. (2011). *Naturally selected: The evolutionary science of leadership*. New York: Harper Business.

Van Vugt, M., Hogan, R., & Kaiser, R. B. (2008). Leadership, followership, and evolution: Some lessons from the past. *The American Psychologist, 63*(3), 182–96. doi:10.1037/0003-066X.63.3.182.

VanYperen, N. W., Hagedoorn, M., Zweers, M., & Postma, S. (2000). Injustice and employees' destructive responses: The mediating role of state negative affect. *Social Justice Research, 13*(3), 291–312. doi:10.1023/A:1026411523466.

Von Furer-Haimendorf, C. (1967). *Morals and merit. A study in values and social controls in South Asian societies* (1st ed.). Chicago: University of Chicago Press.

von Rueden, C., Gurven, M., Kaplan, H., & Stieglitz, J. (2014). Leadership in an egalitarian society. *Human Nature, 25*(4), 538–566. doi:10.1007/s12110-014-9213-4.

Wade, N. (2007). *Before the dawn: Recovering the lost history of our ancestors*. New York: Penguin.

Weiss, H. M., Suckow, K., & Cropanzano, R. (1999). Effects of justice conditions on discrete emotions. *Journal of Applied Psychology, 84*(5), 786–794. doi:10.1037/0021-9010.84.5.786.

Whiten, A., & Byrne, R. W. (Eds.). (1997). *Machiavellian intelligence II: Extensions and evaluations* (2nd ed.). New York: Cambridge University Press.

Wilking, J. R. (2011). The portability of electoral procedural fairness: Evidence from experimental studies in China and the United States. *Political Behavior, 33*(1), 139–159. doi:10.1007/s11109-010-9119-8.

shackelf@oakland.edu

Property Law Reflections of a Sense of Right and Wrong

Jeffrey Evans Stake

This chapter is about evolution and the law. It applies an evolutionary lens to a few particular legal rules, rules of property law. The application of evolutionary science may help us understand some rules of law by showing that it makes sense for those rules to seem right and just. The application of evolutionary science may help us critique other rules of law by showing why the rules might seem wrong and unjust. Rules that conform to our notions of justice will often continue in effect, mostly without being noticed or questioned. Rules that seem unjust will undermine public support for the law and its agents. Examples of both sorts are presented below.

The first examples to be presented suggest that property law may be based on an evolved sense of right and wrong. It might seem obvious that property law can be understood, in part, as an expression of an evolved morality, but that is not the way property has been approached, traditionally, by many of those in the profession. To appreciate that an evolutionary viewpoint adds something to property theory, it may help to review a traditional legal perspective on property.

A Traditional View of Property

Property law is often seen by lawyers as solely an invention of the state. This view dates back at least to the writings of Jeremy Bentham, who wrote, "Property and law are born together, and die together. Before laws were made there was no property; take away laws and property ceases." (Bentham, 1864, p. 112). Bentham could have meant to make only the trivial point that "property" is a branch of law, and so if there is no law at all, then there is no subset called property. Or Bentham could

J.E. Stake (✉)
Indiana University Maurer School of Law,
211 South Indiana Avenue, Bloomington, IN 47405, USA
e-mail: stake@indiana.edu

© Springer International Publishing Switzerland 2016

T.K. Shackelford, R.D. Hansen (eds.), *The Evolution of Morality*,
Evolutionary Psychology, DOI 10.1007/978-3-319-19671-8_13

shackelf@oakland.edu

have meant only to set out the definition of property: rights protected by the state. But let us assume that Bentham did not mean to be trivial or merely to define the word "property." On these assumptions, Bentham seems to say that there is no protection of possessions without the state.

We can see a modern variation on Bentham's theme in one of the classic textbooks for property students, Cribbet and Johnson's (1989) *Principles of the Law of Property*. They wrote, "It could be argued, *although not very plausibly*, that even animals have an institution of property." [p. 4, italics added]. Rather, they thought, a dog's rights in a bone "are not protected by law. They are protected solely by brute force."

Many property theorists would agree that one essential element for "property" is state protection of the individual's assertion of a claim to a thing. But that does not mean that the state invented social respect for an individual's connection to things or that a dog must rely solely on brute force to retain possession. As we shall see, this concept of property rooted in human exceptionalism does not fit easily with evolutionary theory.

Bentham (1864) and Cribbet and Johnson (1989) have suggested that nothing protects one's belongings other than the law. Is that right? Is there nothing to property but law? Are there any other ways our things might be protected from challengers? Aside from the state, or one's own brute force and ingenuity, is there any other source of protection for one's possessions?

Respect for Possession in the Law

Let us start with a case from England three centuries ago. In *Armory v. Delamirie* (King's Bench, 1722), "a chimney sweeper's boy" sued one of the greatest English silversmiths of all time, Paul de Lamerie. (Chimney sweeps employed small boys because they could fit into the chimneys.) The sweep's boy found a "jewel," took it to the silversmith's shop, and delivered it to an apprentice. Under the pretense of weighing it, the apprentice removed the stones from the setting and told his master that it came to three halfpence. The master offered that to the boy, who refused and insisted on a return of the jewel. After the apprentice returned the empty socket, the sweep's boy sued the silversmith for the value of the stones. The silversmith defended by saying that the boy had no claim because he was not the owner. The boy won. He won even though he was poor, even though he was not the owner, and even though the Court cited no statute or prior case requiring that he win. The decision for the sweep's boy established the principle that a prior possessor has enough of an interest in a thing to win a suit for the value of it.

Respect for Possession in Nature: The Bourgeois Strategy

Why might this result make sense? Charles Darwin teaches us to seek enlightenment in the struggle for survival or the competition for sex. Survival will help us in this case. Survival requires resources and, as population increases, resources become

shackelf@oakland.edu

scarce. Scarcity creates competition, and competition can be costly, even deadly, for the competitors. Conspecifics can reduce their costs and increase their chances of survival by finding a way to settle disputes without fighting. Competitors can save energy and reduce injuries if they can find a nonviolent way to allocate contested resources.

To settle disputes without a fight, conspecifics in the resource allocation game need to play by a shared strategy that will lead to different behaviors for the players in each dispute. If the strategy leads to both parties playing "hawk" in a dispute, the strategy fails to avoid a costly fight and results in reduced chances of survival. To generate different behaviors, the strategy needs to focus on some characteristic that will differentiate between the parties in most cases. If conspecifics share a strategy for settling resource disputes that does hinge on such an asymmetry, it may qualify as an evolutionarily stable strategy (ESS) (Maynard Smith & Parker, 1976). In other words, if the population of conspecifics plays by the strategy, there is nothing to be gained by playing a different strategy. The ESS, once established, will not be displaced by another strategy unless there is a change in the environment.

There are many sorts of asymmetries that competitors could use to solve disputes over resources. The strategy could focus upon which competitor is heavier. For example, when funnel-web spiders vie for a web, the heavier often wins. The spiders apparently can assess weight by the frequency of web vibration when a spider moves. Riechert (reported in Maynard Smith, 1982, p. 116) confirmed that weight matters by gluing a lead weight to the abdomen of one spider intruding into the web of another spider of approximately the same size. Artificially doubled in weight, the intruder usually won. The riskiest contests were the ones where the web was highly valuable and the resident was slightly smaller than the intruder.

An asymmetry of weight is correlated with what it takes for a spider to win a physical fight, so the strategy based on weight is a correlated strategy. For another example of a correlated strategy, consider male stalk-eyed flies. Egge, Brandt, and Swallow (2011) observed that the flies often compare the length of their eye stalks by lining up face to face and sometimes settle disputes without injury according to which of the contestants has eyes furthest apart. This is a correlated strategy because eye-stalk length is correlated with size, and size is correlated with resource holding potential.

But an asymmetry need not be correlated to be the basis of an ESS. The "bourgeois" strategy is another way to settle resource disputes without fighting (Maynard Smith, 1982, p. 22). This strategy is not based on the characteristics of the two contestants but rather on their relationships to a resource. The asymmetry revolves around possession. The bourgeois strategy provides the following instructions to the contestant: If I am the possessor of the resource, fight to retain possession. If my opponent is the possessor of the resource, defer to his possession.

There is evidence that some species play a bourgeois strategy. Stevens (1988) studied feral horses on islands off of North Carolina. After a rain, bands of horses stop to drink fresh water that has collected in small pools. There are frequent encounters between bands of horses competing for water at these temporary pools. Stevens observed 233 contests, of which the resident band won 80 percent. This result would be unlikely as a matter of chance. The horses without possession seem

shackelf@oakland.edu

to have an inclination to defer to those in possession. This inclination is not always paramount; the horses did not always play the bourgeois strategy, respecting prior possession. But in nearly all the cases where the intruders ousted the residents, the intruders had a numerical advantage. Perhaps feral horses also play a numerical advantage strategy, which would be correlated with ability to win a physical contest, but turn to the bourgeois strategy when numbers are close to even. And of course, if thirst were strong enough, it might trump either strategy.

Sigg and Falett (1985) tested baboons for the possibility of playing a bourgeois strategy. They gave food to a subordinate baboon, waited 5 min, then released a more dominant baboon into the same cage, and watched to see whether the dominant baboon took the food within 30 min. They found that males did not take from other males, indicating some respect for prior possession, although both males and females would take from females about half of the time.

An experiment by Kummer (reported in Maynard Smith, 1989, p. 213) suggests that potential disputes over food are not the only baboon disputes within the scope of the bourgeois strategy. One male hamadryas baboon showed respect for prior possession of a female baboon by staying well away when let into a cage where another male had already taken possession of the female, a female unknown to both before the experiment. Kummer reversed the males with a new female and the result reversed.

Similarly, Packer (reported in Maynard Smith, 1989, p. 213) found that, although male ownership of female anubis baboons did not last more than a few days, within each day the first to possess a female was respected by other males, even ones that were more dominant as measured by food-access priority. When John Mellencamp sings "I saw you first" in Key West Intermezzo, he suggests that some humans would like to apply a bourgeois strategy to resolve conflicts over mates, lending some support to the feminist critique of patriarchy. What *feels* right does not determine what *is* right. Is is not ought.

Gintis (2007) has formalized the biological advantages of a respect for possession. He shows, however, that if the resource has high enough value, the respect for possession may break down. For example, an animal that ordinarily will not fight over food possessed by a conspecific might switch to a different strategy and begin a fight if he is starving. The presumption that the possessor wins may be overcome by other considerations.

Given that the math supports it and that a bourgeois strategy may have evolved in other animals, it is possible that the bourgeois strategy is one of the strategies harbored in human brains. If so, it might seem morally wrong to us when other humans deviate from the bourgeois strategy, that is to say when they fail to respect possession. In *Armory v. Delamirie*, the sweep's boy had prior possession, and the silversmith's apprentice failed to respect that possession. The Court's decision in *Armory* may be a manifestation of a natural respect for possession, a bourgeois strategy embedded deep in human psychology. Perhaps there is more to property than brute force and the protection of the state. Perhaps we have a property instinct.

shackelf@oakland.edu

Two Elements of Possession: Physical Control and Intent

What does it take to trigger this sense of property? What is required for possession? The common law requires two things for a person to possess property. One is physical control and the other is intent to maintain physical control. Both of these make sense as a matter of biology. Starting with the latter, if you do not intend to assert control over a resource, I can ignore your physical control and take control myself without any danger of a fight. For me to respect your apparent control, I need to see some evidence that you are willing to fight to retain control. My sense of respect for your things is counterproductive, wasteful of opportunities, if it extends to things you do not care about.

But a strategy based on intent to control alone would not be very helpful. On one hand, it would too often fail to prevent fights. In many situations, two persons could simultaneously intend to control a resource. The failure of asymmetry would lead to frequent physical contests, costly to both parties. On the other hand, resources might go unused. Two parties with the opportunity to gain physical control might both refrain from doing so if the strategy says to defer when the other party appears to have the intent to gain control. It would be wasteful if your genes and mine are telling us to refrain from acquiring resources that are merely intended to be controlled by the other because it might be the case that the other does not have the capacity to carry out that intent. Thus, it makes sense for the bourgeois strategy to be triggered by a combination of certain physical control and intent to maintain dominion and control.

Elements of Possession in the Law

We can see an application of the physical control element in the early American case of *Pierson v. Post* [Supreme Court of New York, 1805]. In that case, two fox hunters fought for ownership of a fox pelt. The first hunter, Post, had flushed a fox and was chasing it on Long Island, land considered at the time to be "waste" land. Pierson saw the fox and killed it. Post sued for the fox pelt, and the dispute cost so much that one family had to sell their house to the other family. In the end, Post lost.

To resolve the case, the court had to decide whether a high likelihood of apprehension should be good enough to establish possession. The majority decided that possession required certain control, essentially physical occupation, not merely a high likelihood of control. In terms of the bourgeois strategy, this makes sense. If the strategy could be triggered by a good chance of control, there would be many cases where two parties would have both intent and a good chance of control, and the players, playing by the rules, would end up in a physical fight, with harm to both. By requiring certain control, the court reduced the number of cases in which both parties would qualify as possessors. In other words, the essential asymmetry needed by an ESS is not created as often by a requirement of a good chance of control as it is by

shackelf@oakland.edu

a requirement of certain control. In sum, the property law doctrine for assigning initial rights in things can be understood as a human expression of the bourgeois strategy.

The requirements for possession also fit with results from studies on infants. Experiments done by Woodward (1998, 1999) show that humans at 9 months of age recognize that a grasping hand has an intent motivating it. Infants watched a human arm reach from behind a curtain to grasp toy A and not toy B. After habituation, the positions of the toys were switched. If the hand grasped toy A again, infants did not look long. But if the hand grasped toy B, the infants stared longer, probably because that violated their expectation of an intent to control toy A. A hand that dropped on top of toy A did not generate the same expectation, at least as indicated by the fact that there was no difference in looking time. Even infants know that grasping hands reveal goals, intent. This study indicates that the law's requirement of physical control and intent to control may be rooted in human psychology. Imagining "no possessions" might not be as easy as John Lennon thought it could be.

Title by Prescription and Changes in Value with Duration of Possession

Evolutionary psychology can help us to understand another, and perhaps more puzzling, set of rules from traditional English (and now American) property law, the doctrine of adverse possession. This doctrine reallocates title to land from an owner to a possessor if the owner waits too long to eject the possessor from the land. If A possesses O's land long enough, A becomes the owner and O's title is extinguished. Why should this be so? There have been many theories offered to justify this doctrine. Some of them made sense many years ago, and some have superficial appeal today, but few if any stand up under careful examination (Stake, 2001).

We can make some sense of the doctrine with help from studies of bird psychology. In a study that has been performed in a number of ways on a number of different species of birds, an experimenter removes bird 1 from its territory, allows bird 2 to occupy the territory, releases bird 1 back into the territory, and observes the lengths and outcomes of the resulting fights. The experimenters vary the time that the original bird 1 is out of possession and the time that the replacement bird 2 is in possession. Krebs (1982), for example, found that the value of territory to great tits changes with possession and absence. Birds are willing to fight longer as the length of their possession increases, and they are less willing to fight as their length of absence increases. As possession lengthens, value increases. As absence lengthens, value decreases. If it is true for birds, it might also be true for humans.

In an adverse possession case, a judge is faced with conflicting claims to territory; both parties think it is theirs. The court must decide which of the two parties will lose title. The claim of one party, the title holder or owner, nowadays is usually based on a recorded deed. The claim of the other party, the adverse possessor, is based on possession. Faced with the choice of depriving one party or another, it might be efficient for the court to deprive the party that places a lower value on the

shackelf@oakland.edu

land, the party that would feel less pain from the deprivation, the party that would fight less hard for possession. If territory value for humans changes as it does for birds, eventually the value to the current possessor exceeds the value to the previous possessor or title holder. After that point in time, a court will do less harm by depriving the owner of title than by depriving the possessor of possession.

The elements of the adverse possession doctrine help to support the conclusion that the adverse possessor places a higher value on the land. One element of the doctrine is "actual possession"; the adverse possessor must be in actual possession, meaning occupancy of the territory. A second element is that the adverse possessor must have a "claim of right", and in many jurisdictions this must be a good faith claim of right. These two elements help to establish that the adverse possessor's attachment is growing and his value is on the rise. Another element is "exclusive possession"; the adverse possessor must be in exclusive possession, meaning that the record owner is not in occupancy. This element gives some assurance that either the title holder never had possession or that the title holder's attachment is waning, that the value to the title holder is declining. Two other elements are "continuous possession" and the statutory period of time. These elements help to assure us that the adverse possessor's climbing valuation and the title holder's falling valuation have been continuing for a long period of time, long enough that the lines have crossed and the former now exceeds the latter.

Thus, in an adverse possession case, the law mimics the result that would obtain if the claimants fought it out in a state of nature, but without the need for a fight. The doctrine could also be seen as reaching the result that would obtain from market transactions if both parties were endowed with enough wealth to bargain effectively over the land. By avoiding the incentive for a fight, perhaps the law reaches a result that fits with common notions of justice and fairness.

Davies (1976) has observed that in some conditions pied wagtails frequently leave their territories but never leave them unattended for too long. They return to their territories periodically to evict intruders before the intruders become hard to evict. The doctrine of adverse possession asks owners to behave like pied wagtails. Owners who fail to do so perhaps do not deserve the law's protection.

The Law as Extended Human Phenotype

Once again, maybe there is more to property than brute force and the protection of the state. The rules of property may be based in part on expectations of human behavior, an evolved human sense of right and wrong. Beaver genes interact with their environment to produce the beaver and then the beaver's dam (Dawkins, 1982). Just as the dam is part of the extended beaver phenotype and the web is part of the extended spider phenotype, some laws of property may be part of the extended human phenotype. Perhaps, like the common law, our sense of morality is highly contextual, a built-up set of right or wrong actions in various situations.

shackelf@oakland.edu

The Sense of Justice in Conflict with the Law of Eminent Domain

I do not claim that all rules of property law fit with our sense of justice. Many do not. These rules may be rules that are amoral or may be contrary to our moral sense. In the latter cases, they might continue to exist for reasons of efficiency, or they might not have had enough time to evolve to rules that fit better with our sense of justice. Evolutionary analysis can help to explain and identify, and perhaps even reform, some of the rules that do not conform to our sense of justice.

For an example, consider a relatively recent case from the United States Supreme Court, *Kelo v. City of New London* (2005). New London, Connecticut, was in bad shape. It came up with a plan to redevelop the downtown harbor area, with new housing, public recreation, and business uses. To implement the plan, the government commissioned the New London Development Corporation (NLDC). The NLDC tried to buy the 100 parcels needed for the development, but some of the owners refused to sell, including Susette Kelo and Wilhelmina Dery. Governments around the world have had for centuries the power of eminent domain to deal with these situations, where holdouts might otherwise prevent efficient development of land. The NLDC used the power of eminent domain to condemn the properties of Kelo and Dery and a few others. Kelo and others brought suit to enjoin the condemnation proceedings on the ground that the taking of their land would not be "for public use" as required by Fifth Amendment. In a decision that is consistent with more than a century of Supreme Court precedent, the Court held that taking land for purposes that benefit the public qualifies as taking "for public use" and thus is permitted by the Constitution as long as just compensation is paid. The reaction to this was widespread public disapproval. More than 40 states changed their laws to make condemnations more difficult, some of them by increasing the compensation that the government must pay in some cases.

The Fifth Amendment says, "… nor shall private property be taken for public use without just compensation." Thus, the words of the Constitution establish that payment to the aggrieved owner must be "just." The Supreme Court has said that the standard for just compensation is fair market value. In my view, part of the great dissatisfaction with the power of eminent domain stems from the fact that the law does not require the government to pay what people think is adequate compensation.

As we saw in our discussion of the bourgeois strategy above, possession might trigger certain emotions in humans, emotions leading to greater willingness to fight to retain possession than existed before possession. This proposition is supported by psychological experiments (e.g., Kahneman, Knetsch, & Thaler, 1991) in which it has been found that granting someone possession of an item triggers an increase in its value to that person or increases the pain of losing the item. In this context, this is often called the "endowment effect," which is closely connected to the bourgeois strategy. And, as we saw in the bird studies discussed above, the value of territory

shackelf@oakland.edu

changes with possession and absence, increasing as possession lengthens and decreasing as absence continues. If, again, humans react as birds do to possession and absence and if other humans recognize that human trait, people will have more subtle notions of just compensation than the Supreme Court has recognized.

Combining those studies gives us insights into this idea of just compensation and hence into one of the reasons that the *Kelo* decision touched a nerve. First, if possession triggers an endowment effect, then people like Ms. Kelo who are being dispossessed by a governmental taking of property ought to get more compensation for their loss than owners losing similar land but who are not losing possession. Second, if attachment grows with length of possession, those who have been in longer possession at the time of the taking should be given greater compensation than those who have been in possession for less time. Ms. Dery, who was born in her house in 1918 and had lived there for her whole life, deserved more compensation for her attachment than Ms. Kelo. Third, once an owner has been out of possession for a period of time, the bonus for possession ought to start to erode away, just as the attachment to territory appears to wither away for birds. Fourth, it seems reasonably certain that the birds under observation had no concept of title or ownership protected by the state or collective or group. Attachment does not require a sense of legal title. If that is true, then not only do owners feel losses when they are dispossessed by governmental expropriation, others who live on the property also feel a loss. The children and other family members of owners deserve compensation when they lose their home even though their names do not appear on the deed.

Do people react as the bird studies indicate they might? Fiery Cushman and I surveyed participants on the Moral Sense Test web site (http://mprlab327.webfactional.com/mst/) [Stake & Cushman, in progress]. The participants were asked to specify the compensation that they thought would be just compensation in a number of different situations. Their responses confirmed most of what was predicted above. The participants indicated that the compensation that is just when the government takes land is about $8000 more than market value, is about $20,000 higher when the owner is in possession (which did not increase for a doubly valuable home), is about $5000 for each additional 10 years of possession, and is about $5000 greater for each child displaced by the taking. The participants did not, however, feel that justice required extra compensation to be awarded when a married couple was displaced instead of a single owner.

The constitution does not require extra compensation for possession. It does not recognize that longer possession might increase attachment to land or that absence might diminish that attachment. And the law does not recognize that people might feel substantial losses from governmental takings even though they do not own the land taken. For all these reasons, it should not surprise us when the public reacts with alarm to well-publicized news stories involving the exercise of the power of eminent domain. But the power of eminent domain could be reformed, and an evolutionary perspective can help us to move the law of just compensation toward greater compatibility with our sense of justice.

shackelf@oakland.edu

Conclusions

Speaking in what he called a "spectral voice" of Justice Oliver Wendell Holmes, Jr., Kaplan (1983, p. 1843) said, "in order to attract general obedience to law, surely a condition of stable government, the law must remain within a reasonable range of popular conceptions of what is right." In a number of ways, the law of property conforms to the sense of justice that one might predict from biological and psychological studies of value. The law of possession fits nicely with the theory of an ESS, with evidence for the bourgeois strategy in nonhuman animals, and with evidence of loss aversion in humans. The law of adverse possession and its details fit with the findings in animal studies that attachment to territory increases with time in possession and decreases with absence.

Of course, there are many areas of property law that do not trigger moral sentiments one way or another. And on the other end of the spectrum, there exist rules of American property law that deviate from what would be predicted to be our sense of moral justice. If the rules of just compensation are not changed, they will continue to work injustice, undermine respect for government, and detract from general obedience to the law.

These different examples create some tension in this chapter. Does the law reflect a human sense of morality or not? Both. Some rules of law do reflect our sense of right and wrong, and some run counter to it. From the existence of the latter, we can see that the law does not determine our sense of justice, however much it shapes it. From the former, we can see that adaptive moral senses extend their influence into our system of rules governing tangible resources. Some rules of property are part of the extended human phenotype.

Cases

Armory v. Delamirie (King's Bench, 1722) 1 Strange 505, 1 Sess. Cas. K.B. 505

Kelo v. City of New London (Supreme Court of the United States, 2005) 545 U.S. 469

Pierson v. Post (Supreme Court of New York, 1805) 3 Caines 175–182, 2 Am. Dec. 264

References

Bentham, J. (1864). *The theory of legislation, principles of the civil code, Part I*. Dumont ed., Hildreth Trans.

Cribbet, J. E., & Johnson, C. E. (1989). *Principles of the law of property* (3rd ed.).

Davies, N. B. (1976). Food, flocking and territorial behaviour of the pied wagtail (*Motacilla alba yarrellii* Gould). *Journal of Animal Ecology, 45*, 235–253.

Dawkins, R. (1982). *The extended phenotype; The long reach of the gene*.

Egge, A. R., Brandt, Y., & Swallow, J. G. (2011). Sequential analysis of aggressive interactions in the stalk-eyed fly *Teleopsis dalmanni. Behavioral Ecology and Sociobiology, 65*, 369–379.

shackelf@oakland.edu

Gintis, H. (2007). The evolution of private property. *Journal of Economic Behavior & Organization, 64*, 1–16. doi:10.1016/j.jebo.2006.02.002.
Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives, 5*, 193–206.
Kaplan, B. (1983). Encounters with O.W. Holmes, Jr. *Harvard Law Review, 96*, 1828–1852.
Krebs, J. R. (1982). Territorial defence in the great tit (*Parus major*): Do residents always win? *Behavioral Ecology and Sociobiology, 11*, 185–194.
Maynard Smith, J. (1982). *Evolution and the theory of games*.
Maynard Smith, J. (1989). *Did Darwin get it right?*
Maynard Smith, J., & Parker, G. A. (1976). The logic of asymmetric contests. *Animal Behavior, 24*, 159–175.
Mellencamp, J., & Green, G. *Key West Intermezzo (I Saw You First)*.
Sigg, H., & Falett, J. (1985). Experiments on respect of possession and property in hamadryas baboons (*Papio hamadryas*). *Animal Behaviour, 33*, 978–984.
Stake, J. E. (2001). The uneasy case for adverse possession. *Georgetown Law Journal, 89*, 2419–2474.
Stevens, E. F. (1988). Contests between bands of feral horses for access to freshwater: The resident wins. *Animal Behaviour, 36*, 1851–1853.
Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*, 1–34.
Woodward, A. L. (1999). Infants' ability to distinguish between purposeful and non-purposeful behaviors. *Infant Behavior and Development, 22*, 145–160.

shackelf@oakland.edu

Game Theory and Morality

Moshe Hoffman, Erez Yoeli, and Carlos David Navarrete

Introduction

Consider the following puzzling aspects of our morality:

1. Many of us share the view that one should not use people, even if it benefits them to be used, as Kant intoned in his second formulation of the categorical imperative: "Act in such a way that you treat humanity, whether in your own person or in the person of any other, never merely as a means to an end, but always at the same time as an end" (Kant, 1997). Consider dwarf tossing, where dwarfs wearing protective padding are thrown for amusement, usually at a party or pub. It is viewed as a violation of dwarfs' basic dignity to use them as a means for amusement, even though dwarves willingly engage in the activity for economic gain. Many jurisdictions ban dwarf tossing on the grounds that the activity violates dwarfs' basic human rights, and these laws have withstood lawsuits raised by dwarfs suing over the loss of employment (!).
2. Charitable giving is considered virtuous, but little attention is paid to how just the cause or efficient the charity. For example, Jewish and Christian traditions advocate giving 10 % of one's income to charity, but make no mention of the importance of evaluating the cause or avoiding wasteful charities. The intuition that giving to charity is a moral good regardless of efficacy results in the persistence of numerous inefficient and corrupt charities. For example, the Wishing Well Foundation has, for nearly a decade, ranked as one of CharityNavigator.

M. Hoffman (✉) • E. Yoeli
Program for Evolutionary Dynamics, Harvard University,
One Brattle Square, Suite 6, Cambridge, MA 02138, USA
e-mail: moshehoffman@fas.harvard.edu

C.D. Navarrete
Department of Psychology, and, the Ecology, Evolutionary Biology and Behavior Program,
Michigan State University, East Lansing, MI, USA

© Springer International Publishing Switzerland 2016

T.K. Shackelford, R.D. Hansen (eds.), *The Evolution of Morality*,
Evolutionary Psychology, DOI 10.1007/978-3-319-19671-8_14

shackelf@oakland.edu

com's most inefficient charities. Yet its mission of fulfilling wishes by children with terminal illnesses is identical to that of the more efficient Make-A-Wish Foundation. Worse yet, scams masquerading as charities persist. One man operating as The US Navy Veteran's Association collected over 100 million dollars—over 7 years!—before anyone bothered to investigate the charity.

3. In every culture and age, injunctions against murder have existed. If there is one thing much of humanity seems to agree on, it's that ending the life of another without just cause which is among the worst of moral violations. Yet cultures don't consider the loss of useful life years in their definition, even though it is relevant to the measure of harm done by the murder. Why is our morality so much more sensitive to *whether* a life was lost than to how much life was lost?

There are numerous other examples of how our moral intuitions appear to be rife with logical inconsistencies. In this chapter, we use game theory to provide insight on a range of moral puzzles similar to the puzzles described above.

What Is Game Theory and Why Is It Relevant?

In this section, we review the definition of a game, and of a Nash equilibrium, then discuss how evolution and learning processes would yield moral intuitions consistent with Nash equilibria.

Game theory is a tool for the analysis of social interactions. In a game, the *payoff* to each *player* depends on their actions, as well as the actions of others. Consider the Prisoner's Dilemma (Chammah & Rapoport, 1965; see Fig. 1), a model that captures the paradox of cooperation. Each of two players chooses whether to cooperate or to defect. Cooperating reduces a player's payoff by $c > 0$ while increasing the other's payoffs by $b > c$. Players could be vampire bats with the option of sharing blood, or firms with the option of letting each other use their databases, or premed students deciding whether to take the time to help one another to study. The payoffs, *b* and *c*, may represent likelihood of surviving and leaving offspring, profits, or chance of getting into a good medical school.

Solutions to such games are analyzed using the concept of a Nash equilibrium[1]—a specification of each player's action such that no player can increase his payoff by deviating unilaterally. In the Prisoner's Dilemma, the only Nash equilibrium is for neither player to cooperate, since regardless of what the other player does, cooperation reduces one's own payoff.

[1] Note that we focus on the concept of Nash equilibrium in this chapter and not evolutionary stable strategy (ESS), a refinement of Nash that might be more familiar to an evolutionary audience. ESS are the Nash equilibria that are most relevant in evolutionary contexts. However, ESS is not well defined in many of our games, so we will focus on the insights garnered from Nash and directly discuss evolutionary dynamics when appropriate.

shackelf@oakland.edu

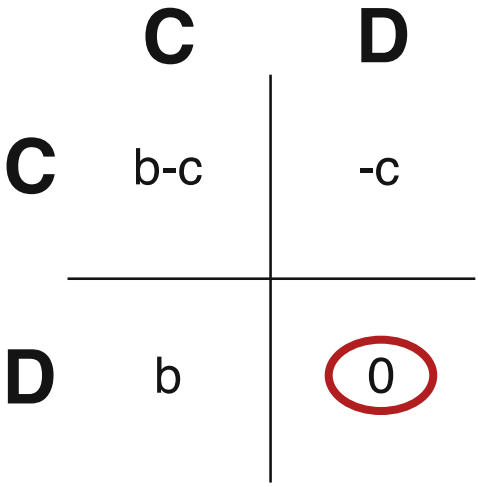

Fig. 1 The Prisoner's Dilemma. Player 1's available strategies (C and D, which stand for cooperate and defect, respectively) are represented as *rows*. Player 2's available strategies (also C and D) are represented as *columns*. Player 1's payoffs are represented at the intersection of each row and column. For example, if player 1 plays D and player 2 plays C, player 1's payoff is b. The Nash equilibrium of the game is (D, D). It is indicated with a *circle*

Game theory has traditionally been applied in situations where players are rational decision makers who deliberately maximize their payoffs, such as pricing decisions of firms (Tirole, 1988) or bidding in auctions (Milgrom & Weber, 1982). In these contexts, behavior is expected to be consistent with a Nash equilibrium, otherwise one of the agents—who are actively deliberating about what to do—would realize she could benefit from deviating from the prescribed strategy.

However, game theory also applies to evolutionary and learning processes, where agents do not deliberately choose their behavior in the game, but play according to strategies with which they are born, imitate, or otherwise learn. Agents play a game and then "reproduce" based on their payoffs, where reproduction represents offspring, imitation, or learning. The new generation then play the game, and so on. In such settings, if a mutant does better (mutation can be genetic or can happen when agents experiment), then she is more likely to reproduce or her behavior imitated or reinforced, causing the behavior to spread. This intuition is formalized using models of evolutionary dynamics (e.g., Nowak, 2006).

The key result for evolutionary dynamic models is that, except under extreme conditions, behavior converges to Nash equilibria. This result rests on one simple, noncontroversial assumption shared by all evolutionary dynamics: Behaviors that are relatively successful will increase in frequency. Based on this logic, game theory models have been fruitfully applied in biological contexts to explain phenomena such as animal sex ratios (Fisher, 1958), territoriality (Smith & Price, 1973), cooperation (Trivers, 1971), sexual displays (Zahavi, 1975), and parent–offspring conflict (Trivers, 1974). More recently, evolutionary dynamic models have been applied in human contexts where conscious deliberation is believed to not play an important role, such as in the adoption of religious rituals (Sosis & Alcorta, 2003), in the expression and experience of emotion (Frank, 1988; Winter, 2014), and in the use of indirect speech (Pinker, Nowak, & Lee, 2008).

shackelf@oakland.edu

Crucially for this chapter, because our behaviors are mediated by moral intuitions and ideologies, if our moral behaviors converge to Nash, so must the intuitions and ideologies that motivate them. The resulting intuitions and ideologies will bear the signature of their game theoretic origins, and this signature will lend clarity on the puzzling, counterintuitive, and otherwise hard-to-explain features of our moral intuitions, as exemplified by our motivating examples.

In order for game theory to be relevant to understanding our moral intuitions and ideologies, we need only the following simple assumption: *Moral intuitions and ideologies that lead to higher payoffs become more frequent*. This assumption can be met if moral intuitions that yield higher payoffs are held more tenaciously, are more likely to be imitated, or are genetically encoded. For example, if every time you transgress by commission you are punished, but every time you transgress by omission you are not, you will start to intuit that commission is worse than omission.

Rights and the Hawk–Dove Game

In this section we will argue that just as the Hawk–Dove model explains animal territoriality (Maynard Smith & Price, 1973, *to be reviewed shortly*), *the Hawk–Dove model sheds light onto our sense of rights* (Descioli & Karpoff, 2014; Gintis, 2007; Myerson, 2004).

Let us begin by asking the following question (Myerson, 2004): "Why [does] a passenger pay a taxi driver after getting out of the cab in a city where she is visiting for one day, not expecting to return?" If the cabby complains to the authorities, the passenger could plausibly claim that she had paid in cash. The answer, of course, is that the cabby would feel that the money the passenger withheld was his—that he had a right to be paid for his service—and get angry, perhaps making a scene or even starting a fight. Likewise, if the passenger did in fact pay, but the cabby demanded money a second time, the passenger would similarly be infuriated. This example illustrates that people have powerful intuitions regarding rightful ownership. In this section, we explore what the Hawk–Dove game can teach us about our sense of property rights.

The reader is likely familiar with the Hawk–Dove game, a model of disputes over contested resources. In the Hawk–Dove game, each player decides whether to fight over a resource or to acquiesce (i.e. play Hawk or Dove). If one fights and the other does not, the fighter gets the resource, worth v. If both fight, each pays a cost c and split the resource. That is, each gets $v/2-c$. If neither fights, they split the resource and get $v/2$. As long as $v/2 < c$, then in any stable Nash equilibrium, one player fights and the other acquiesces. That is, if one player expects the other to fight, she is better off acquiescing, and vice versa (see Fig. 2).

Crucially, it is not just a Nash equilibrium for one player to always play Hawk and the other to always play Dove. It is also an equilibrium for both players to condition whether they play Hawk on an *uncorrelated asymmetry*—a cue or event that

shackelf@oakland.edu

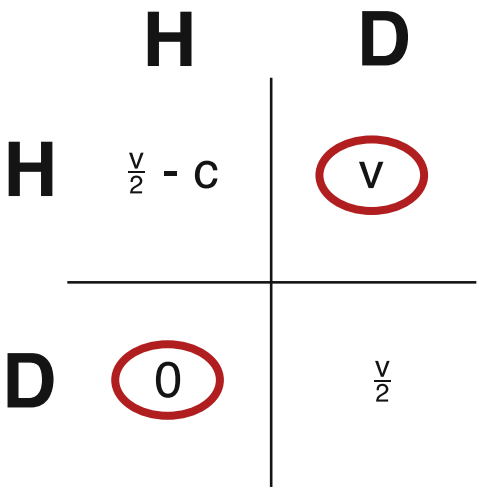

Fig. 2 The Hawk–Dove game. The Nash equilibria of the game are circled

does not necessarily affect the payoffs, but does distinguish between the players, such as who arrived at the territory first or who built the object. If one conditions on the event (say, plays Hawk when she arrives first), then it is optimal for the other to condition on the event (to play Dove when the other arrives first).

As our reader is likely aware, this was the logic provided by Maynard Smith to explain animal territoriality—why animals behave aggressively to defend territory that they have arrived at first, even if incumbency does not provide a defensive advantage and even when facing a more formidable intruder. Over the years, evidence has amassed to support Maynard Smith's explanation, such as experimental manipulation of which animal arrives first (Davies, 1978; Sigg & Falett, 1985).

Like other animals, we condition how aggressively we defend a resource on whether we arrive first. Because our behaviors are motivated by beliefs, we are also more likely to believe that the resource is "ours" when we arrive first. Studies have shown these effects with children's judgments of ownership, in ethnographies of prelegal societies, and in computer games. In one such illustration, DeScioli and Wilson (2011) had research subjects play a computer game in which they contested a berry patch. Subjects who ended up keeping control of the patch usually arrived first, and this determined the outcome more often than differences in fighting ability in the game.

This sense of ownership is codified in our legal systems, as illustrated by the quip "possession is 9/10ths of the law," and in a study involving famous legal property cases conducted by Descioli and Karpoff (2014). In a survey, these researchers asked participants to identify the rightful owner of a lost item, after reading vignettes based on famous property rights legal cases. Participants consistently identified the possessor of the found item as its rightful owner (as the judges had at the time of the case). This sense of ownership is also codified in our philosophical tradition, e.g., in Locke (1988), who found property rights in initial possession. Note that, as has also been found in animals, possession extends to objects on one's land: In DeScioli and

shackelf@oakland.edu

Karpoff's survey, another dictate of participants' (and the judges') property rights intuitions was who owned the land on which the lost item was found.

Also like animals, our sense of property rights is influenced by who created or invested in the resource, another uncorrelated asymmetry. In locales that sometimes grant property rights to squatters—individuals who occupy lands others have purchased—a key determinant of whether the squatters are granted the land is whether they have invested in it (Cone vs. West Virginia Pulp & Paper Co., 1947; Neuwirth, 2005). Locke also intuited that investment in land is part of what makes it ours: In *Second Treatise on Civil Government* (1689), Locke wrote, "everyman has a property in his person; this nobody has a right to but himself. The labor of his body and the work of his hand, we may say, are properly his."

If the Hawk–Dove model underlies our sense of property rights, we would expect to see psychological mechanisms that motivate us to feel entitled to an object when we possess it or have invested in it. Here are three such mechanisms, which can be seen by reinterpreting some well-documented "biases" in the behavioral economics literature. The first such bias is the *endowment effect*: We value items more if we are in possession of them. The endowment effect has been documented in dozens of experiments, where subjects are randomly given an item (mug, pen, etc.) and subsequently state that they are willing to sell the mug for much more than those who were not given the mug are willing to pay (Kahneman, Knetsch, & Thaler, 1990). In the behavioral economics literature, the endowment effect has sometimes been explained by loss aversion, which is when we are harmed more by a loss than we benefit from an equivalent gain. However, the source of loss aversion is not questioned or explained. When it is, loss aversion is also readily explained by the Hawk–Dove game (Gintis, 2007).

A second bias that also fits the Hawk–Dove model is the *IKEA effect*: Our valuation of an object is influenced by whether we have developed or built the resource. The IKEA effect has been documented by asking people how much they would pay for items like Lego structures or IKEA furniture after randomly being assigned to build them or receive them pre-built. Subjects are willing to pay more for items they build themselves.

A third such bias that fits the Hawk–Dove model is the *sunk cost fallacy* (Mankiw, 2007; Thaler, 1980), which leads us to "throw good money after bad" when we invest in ventures simply because we have already put so much effort into them, arguably because our prior efforts lead us to value those ventures more.

Possession and past investment are not the only uncorrelated asymmetries that can dictate rights. Rights can be dictated by a history of agreements, as happens when one party sells another deed to a house or car, or, as in our taxicab example, by whether a service was provided. There are also countless examples in which rights were determined by perhaps unfair or arbitrary characteristics such as race and sex: Black Americans were expected to give up their seat for Whites in the Jim Crow South and women to hand over their earnings or property to their husbands throughout the ages.

Hawk–Dove is not just a post hoc explanation for our sense of rights; it also leads to the following novel insight: We can formally characterize the properties that

shackelf@oakland.edu

uncorrelated asymmetries must have. This requires a bit more game theory to illustrate; the logic is detailed in the section on categorical distinctions but the implications are straightforward: Uncorrelated asymmetries must be discrete (as in who arrived first or whether someone has African ancestry) and cannot be continuous (who is stronger, whether someone has darker skin). Indeed, we challenge the reader to identify a case where our sense of rights depends on surpassing a threshold in a continuous variable (stronger than? darker than?). More generally, an asymmetry must have the characteristic that, when it occurs, every observer believes it occurred with a sufficiently high probability, where the exact level of confidence is determined by the payoffs of the game. This is true of public, explicit speech and handshakes, but not innuendos or rumors. (Formally, explicit speech and handshakes induce what game theorists term common *p*-beliefs.)

The Hawk–Dove explanation of our sense of rights also gives useful clarity on when there will be conflict. Conflict will arise if both players receive opposing signals regarding the uncorrelated asymmetry, such as two individuals each believing they arrived first, or when there are two uncorrelated asymmetries that point in conflicting directions, such as when one person invested more and the other arrived first. The former source of conflict appears to be the case in the Israeli–Palestinian conflict. Indeed, both sides pour great resources into demonstrating their early possession, especially Israel, through investments in and public displays of archeology and history. The latter source of conflict appears to be the case in many of the contested legal disputes in the study by DeScioli and Karpoff (2014) mentioned above. An example is one person finds an object on another's land. Indeed, this turns out to be a source of many legal conflicts over property rights, and a rich legal tradition has developed to assign precedence to one uncorrelated asymmetry over another (Descioli & Karpoff, 2014). As usual, we see similar behavior in animals in studies that provide empirical support for Maynard Smith's model for animal territoriality: When two animals are each given the impression they arrived first by, for example, clever use of mirrors, a fight ensues (Davies, 1978).

Authentic Altruism, Motives, and the Envelope Game

In this section, we present a simple extension of the Repeated Prisoner's Dilemma to explain why morality depends not just on what people do but also what they think or consider.

In the Repeated Prisoner's Dilemma and other models of cooperation, players judge others by their actions—whether they cooperate or defect. However, we not only care about whether others cooperate but also about their decision-making process: We place more trust in cooperators who never even considered defecting. To quote Kant, "In law a man is guilty when he violates the rights of others. In ethics he is guilty if he only thinks of doing so."

The Envelope Game (Fig. 3) models why we care about thoughts and considerations and not just actions (Hoffman, Yoeli, & Nowak, 2015). The Envelope Game

shackelf@oakland.edu

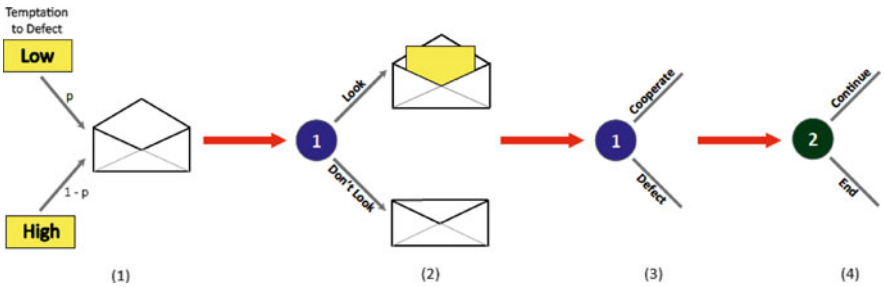

Fig. 3 A single stage of the Envelope Game

is a repeated game with two players. In each round, player 1 receives a sealed envelope, which contains a card stating the costs of cooperation (high temptation to defect vs. low temptation to defect). The temptation is assigned randomly and is usually low. Player 1 can choose to look inside the envelope and thus find out the magnitude of the temptation or choose not to look. Then player 1 decides to cooperate or to defect. Subsequently, player 2 can either continue to the next round or end the game. As in the Repeated Prisoner's Dilemma, the interaction repeats with a given likelihood, and if it does, an envelope is stuffed with a new card and presented to player 1, etc.

In this model, as long as temptations are rare, large, and harmful to player 2, it is a Nash equilibrium for player 1 to "cooperate without looking" in the envelope and for player 2 to continue if and only if player 1 has cooperated and not looked. We refer to this as the *cooperate without looking* (CWOL) equilibrium.[2] This equilibrium emerges in agent-based simulations of evolution and learning processes.[3] Notice that if player 1 could not avoid looking inside the envelope, or player 2 could not observe whether player 1 looked, there would not be a cooperative equilibrium since player 1 would benefit by deviating to defection in the face of large temptations. Not looking permits cooperative equilibria in the face of large temptations.

The Envelope Game is meant to capture the essential features of many interesting aspects of our morality, as described next.

Authentic Altruism. Many have asked whether "[doing good is] always and exclusively motivated by the prospect of some benefit for ourselves, however subtle" (Batson, 2014), for example, the conscious anticipation of feeling good (Andreoni,

[2] Technically, the conditions under which we expect players to avoid looking and attend to looking are $c_h > a/(1-w) > c_l p + c_h(1-p)$ and $bp + d(1-p) < 0$), where c_h and c_l are the magnitudes of the high and low temptations, respectively; p is the likelihood of the low temptation; $a/(1-w)$ is the value of a repeated, cooperative interaction to player 1; and $bp + d(1-p)$ is the expected payoff to player 2 if player 1 only cooperates when the temptation is low.

[3] The simulations employ numerical estimation of the replicator dynamics for a limited strategy space: cooperate without looking, cooperate with looking, look and cooperate only when the temptation is low, and always defect for player 1, and end if player 1 looks, end if player 1 defects, and always end for player 2.

shackelf@oakland.edu

1990), avoidance of guilt (Cain, Dana, & Newman, 2014; Dana, Cain, & Dawes, 2006; DellaVigna, List, & Malmendier, 2012), anticipation of reputational benefits or reciprocity (as Plato's Glaucon suggests, when he proffers that even a pious man would do evil if given a ring that makes him invisible; Trivers, 1971). At the extreme, this amounts to asking if saintly individuals such as Gandhi or Mother Teresa were motivated thus, or if they were "authentic" altruists who did good without anticipating any reward and would be altruistic even in the absence of such rewards. Certainly, religions advocate doing good for the "right" reasons. In the Gospel of Matthew, Chapter 6, Jesus advocates, "Be careful not to practice your righteousness in front of others to be seen by them. If you do, you will have no reward from your Father in heaven," after which he adds, "But when you give to the needy, do not let your left hand know what your right hand is doing, so that your giving may be in secret. Then your Father, who sees what is done in secret, will reward you."

The Envelope Game suggests authentic altruism is indeed possible: By focusing entirely on the benefits to others and ignoring the benefits to themselves, authentic altruists are trusted more, and the benefits from this trust outweigh the risk of, for example, dying a martyr's death. Moreover, this model helps explain why we think so highly of authentic altruists, as compared to others who do good, but with an ulterior motive (consider, as an example, the mockery Sean Penn has faced for showing up at disaster sites such as Haiti and Katrina with a photographer in tow).

Principles. Why do we like people who are "principled" and not those who are "strategic"? For example, we trust candidates for political office whose policies are the result of their convictions and are consistent over time and distrust those whose policies are carefully constructed in consultation with their pollsters and who "flip-flop" in response to public opinion (as caricatured by the infamous 2004 Republican presidential campaign television ad showing John Kerry windsurfing and tacking from one direction to another). CWOL offers the following potential explanation. Someone who is strategic considers the costs and benefits to themselves of every decision and will defect when faced with a large temptation, whereas someone who is guided by principles is less sensitive to the costs and benefits are to themselves and thus less likely to defect. Imagine our flip-flopping politician was once against gay marriage but supports it now that it is popular. This indicates the politician is unlikely to fight for the cause if it later becomes unpopular with constituents or risks losing a big donor. Moreover, this model may help explain why ideologues that are wholly devoted to a cause (e.g., Hitler, Martin Luther King, and Gandhi) are able to attract so many followers.

Don't Use People. Recall Kant's second formulation of the categorical imperative: "Act in such a way that you always treat humanity, whether in your own person or in the person of any other, never simply as a means but always at the same time as an end." In thinking this through, let's again consider dwarf tossing. Many see it as a violation of dwarfs' basic dignity to use them as a means for amusement, even though they willingly engage in the activity for economic gain. Our aversion to using people may explain many important aspects of our moral intuitions, such as

shackelf@oakland.edu

why we judge torture as worse than imprisonment or punishment (torture is harming someone as a means to obtaining information) and perhaps one of the (many) reasons we oppose prostitution (prostitution is having sex with someone as a means to obtaining money). The Envelope Game clarifies the function of adhering to this maxim. Whereas those who treat someone well as means to an end would also mistreat them if expedient, those who treat someone well as an end can be trusted not to mistreat them when expedient.

Attention to Motives. The previous two applications are examples of a more general phenomenon: that we judge the moral worth of an action based on the motivation of the actor, as argued by deontological ethicists, but contested by consequentialists. The deontological argument is famously invoked by Kant: "Action from duty has its moral worth not in the purpose to be attained by it but in the maxim in accordance with which it is decided upon, and therefore does not depend upon the realization of the object of the action but merely upon the principle of volition in accordance with which the action is done without regard for any object of the faculty of desire" (Kant, 1997). These applications illustrate that we attend to motives because they provide valuable information on whether the actor can be trusted to treat others well even when it is not in her interest.

Altruism Without Prospect of Reciprocation. CWOL also helps explain why people cooperate in contexts where there is no possibility of reciprocation, such as in one-shot anonymous laboratory experiments like the dictator game (Fehr & Fischbacher, 2003), as well as when performing heroic and dangerous acts. Consider soldiers who throw themselves on a grenade to save their compatriots or stories like that of Liviu Librescu, a professor at the University of Virginia and a Holocaust survivor, who saved his students during a school shooting. When he heard the shooter coming toward his classroom, Librescu stood behind the door to his classroom, expecting that when the shooter tried to shoot through the door, it would kill him and his dead body would block the door. Mr. Librescu, clearly, did not expect this act to be reciprocated. Such examples have been used as evidence for group selection (Wilson, 2006), but can be explained by individuals "not looking" at the chance of future reciprocation. Consistent with this interpretation, cooperation during extreme acts of altruism is more likely to be intuitive than deliberative (Rand & Epstein, 2014), and those who cooperate without considering the prospect of reciprocation are more trusted (Critcher, Inbar, & Pizarro, 2013). We also predict that people are more likely to cooperate intuitively when they know they are being observed.

The Omission–Commission Distinction and Higher-Order Beliefs

We explain the omission–commission distinction and the means–by-product distinction by arguing that these moral intuitions evolved in contexts where punishment is coordinated. Then, even when intentions are clear to one witness for omissions and by-products, a witness will think intentions are less clear to the other witnesses.

shackelf@oakland.edu

Why don't we consider it murder to let someone die that we could have easily saved? For example, we sometimes treat ourselves to a nice meal at a fancy restaurant rather than donating the cost of that meal to a charity that fights deadly diseases. This extreme example illustrates a general phenomenon: that people have a tendency to assess harmful commissions (actions such as killing someone) as worse, or more morally reprehensible, than equally harmful omissions (inactions such as letting someone die). Examples of this distinction abound, in ethics (we assess withholding the truth as less wrong than lying (Spranca, Minsk, & Baron, 1991)), in law (it is legal to turn off a patient's life support and let the patient die, as long as one has the consent of the patient's family; however, it is illegal to assist the patient in committing suicide even with the family's consent), and in international relations. For example, consider the Struma, a boat carrying Jewish refugees fleeing Nazi persecution in 1942. En route to Palestine, the ship's engine failed, and it was towed to a nearby port in Turkey. At the behest of the British authorities then in control of Palestine, passengers were denied permission to disembark and find their way to Palestine by land. For weeks, the ship sat at port. Passengers were brought only minimal supplies, and their requests for safe haven were repeatedly denied by the British and others. Finally, the ship was towed to known hostile waters in the Black Sea, where it was torpedoed by a Russian submarine almost immediately, killing 791 of 792 passengers. Crucially, though, the British did not torpedo the ship themselves or otherwise execute passengers—an act of commission that they and their superiors would undoubtedly have found morally reprehensible.

Why do we distinguish between transgressions of omission and commission? To address this question, we present a simple game theory model based on the insight by DeScioli, Bruening, and Kurzban (2011). The intuition can be summarized in four steps:

1. We note that moral condemnation motivates us to punish transgressors. Such punishment is potentially costly, e.g., due to the risk of retaliation. We expect people to learn or evolve to morally condemn only when such costs are worth paying.
2. Moral condemnation can be less costly when others also condemn, perhaps because the risk of retaliation is diffused, because some sanctions do not work unless universally enforced or, worse, because others may sanction individuals they believe wrongly sanctioned. This can be modeled using any game with multiple Nash equilibria, including the Repeated Prisoner's Dilemma and the Side-Taking Game. The Coordination Game is the simplest game with multiple equilibria, so we present this game to convey the basic intuition. In the Coordination Game, there are two players who each simultaneously choose between two actions, say punish and don't punish. The key assumption is that each player prefers to do what she expects the other to do, which can be captured by assuming each receives a if they both punish, d if neither punish, $b < d$ if one punishes and the other does not, and $c < a$ if one does not punish while the other does (Fig. 4).
3. Transgressions of omission that are intended are difficult to distinguish from unintended transgression, as is the case when perpetrators are simply not paying

shackelf@oakland.edu

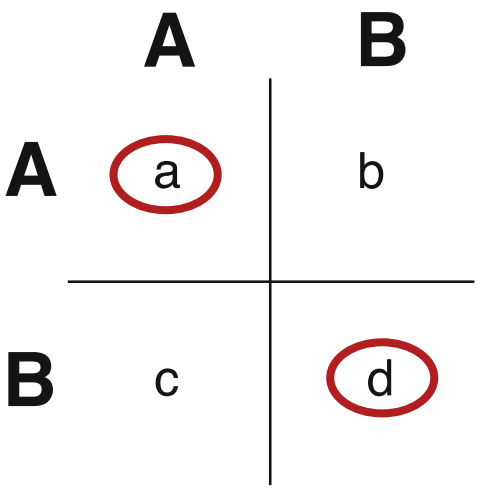

Fig. 4 The Coordination Game. In our applications, A stands for punish, and B stands for don't punish

attention or do not have enough time to react with better judgment (DeScioli et al., 2011). Relative to the example of the tennis player with the allergy described above, it is usually hard to distinguish between a competitor who does not notice his opponent orders the dish with the allergen versus one who notices but does not care. In contrast, transgressions of commission must be intended almost by definition.

4. Suppose the witness knows an omission was intentional: In the above example, the tennis player's opponent's allergy is widely known, and the witness saw the player watch his opponent order the offending dish, had time to react, thought about it, but did not to say anything. The witness suspects that others do not know the competitor was aware his opponent ordered the dish, but believes the tennis player should be condemned for purposely withholding information from his competitor. However, since the witness does not wish to be the sole condemner, she is unlikely to condemn. In contrast, when a witness observes a transgression of commission (e.g., the player recommends the dish), the witness is relatively confident that others present interpret the transgression as purposely harmful, since his recommendation reveals that the player was obviously paying attention and therefore intended to harm his opponent. So, if all other individuals present condemn the tennis player when they observe the commission, each does not anticipate being the sole condemner.

For the above result to hold, all that is needed is the following: (1) The more the costs of punishment decrease, the more others punish and (2) omissions are usually unintended (Dalkiran, Hoffman, Paturi, Ricketts, & Vattani, 2012; Hoffman et al., 2015).[4]

[4] In fact, even if one knows that others know that the transgression was intended, omission will still be judged as less wrong, since the transgression still won't create what game theorists call common *p*-belief, which is required for an event to influence behavior in a game with multiple equilibria.

shackelf@oakland.edu

This explanation for the omission–commission distinction leads to two novel predictions: First, for judgments and emotions not evolved to motivate witnesses to punishment but to, say, motivate witnesses to avoid dangerous partners (such as the emotion of fear; in contrast to anger or moral disgust), the omission–commission distinction is expected to be weaker or disappear altogether. Second, for transgressions of omission that, without any private information, can be presumed intentional (such as a mother who allows her child to go hungry or a person who does not give to a charity after being explicitly asked), we would not expect much of an omission–commission distinction in moral condemnation.

As with the all models discussed in this chapter, the game theoretic explanation for the omission–commission distinction does not rest on rational, conscious, strategic calculation. In fact, in this particular case, all reasonable evolutionary dynamic models lead away from punishing omissions. The fact that the above results do not rest on rational, strategic thinking is particularly important in this setting since there is evidence that the distinction between omissions and commissions is not determined deliberately but rather intuitively (Cushman, Young, & Hauser, 2006) and appears to be evolved (DeScioli et al., 2011) and that consciously considering what others believe is an onerous process (Camerer, 2003; Epley, Keysar, Van Boven, & Gilovich, 2004; Hedden & Zhang, 2002).

This same model can explain several other puzzling aspects of our morality. The first is the *means–by-product* distinction. This distinction has been documented in studies that ask respondents to judge the following variants of the classic "trolley" problem. In the standard trolley "switch" case (Foot, 1967), a runaway trolley is hurtling toward a group of five people. To prevent their deaths, the trolley must be switched onto a side track where it will kill an innocent bystander. In studies using this case, the vast majority of subjects choose the utilitarian option, judging it permissible to cause the death of one to save five (e.g., Cushman et al., 2006; Mikhail, 2007). In the "footbridge" variant (Thomson, 1976), the trolley is hurtling toward the group of five people, but the switch to divert it is inoperable. The only way to save the five is to push a man who is wearing a heavy backpack off a bridge onto the track, thereby slowing the trolley enough so the five can escape, but killing the man. In contrast to the standard switch version, where causing the death of one person is but a by-product of the action necessary to save five, most subjects in the footbridge case find it morally impermissible to force the man with the backpack onto the tracks (Cushman et al., 2006; DeScioli, Gilbert, & Kurzban, 2012)—that is, when the man is used as a means to saving the five—even though the consequences are the same, and the decision to act was made knowingly and deliberately in both cases.

Such effects are found in less contrived situations, as well. Consider the real-life distinction between terrorism, in which civilian casualties are used a means to a political goal, and anticipated collateral damage, which is a by-product of war, even when the same number of civilians are knowingly killed and the same political ends are desired (say increased bargaining power in a subsequent negotiation).

The explanation again uses "higher-order beliefs" and is based on the key insight in DeScioli et al. (2011) and formalized in Dalkiran et al. (2012) and Hoffman et al. (2015): When the harm is done as a by-product, the harm is not usually anticipated.

shackelf@oakland.edu

So even when a witness knows that the perpetrator anticipated the harm, the witness believes other witnesses will not be aware of this and will presume the harm was not anticipated by the perpetrators. For instance, suppose we observe Israel killing civilians as a by-product of a strategic raid on Hamas militants. Even if we knew Israel had intelligence that confirmed the presence of civilians, we might not be sure others were privy to this information. On the other hand, when the harm is done as a means, the harm must be anticipated, since otherwise the perpetrator would have no motive to commit the act. Why would Hamas fire rockets at civilian towns with no military presence if Hamas does not anticipate a chance of civilian casualties? Consequently, it is Nash equilibrium to punish harm done as a means but not harm done as a by-product.

Similar arguments can be made for why we find direct physical transgressions worse than indirect ones, a moral distinction relevant to, for instance, the United States' current drone policy. Cushman et al. (2006) found that subjects condemn pushing a man off a bridge (to stop a train heading toward five others) more harshly than flipping a switch that leads the man to fall through a trap door. Pushing the victim with a stick is viewed as intermediate in terms of moral wrongness. Such moral wrongness judgments are consistent with considerations of higher-order beliefs: When a man is physically pushed, any witness knows the pushing was intended, but when a man is pushed with a stick some might not realize this, and even those who realize it might suspect others will not. Even more so when a button is pressed that releases a trap door.

It is worth noting that the above argument does not depend on a specific model of punishment, as in DeScioli and Kurzban's (2009) Side-Taking Game. The above model also makes the two novel predictions enumerated above, but nevertheless captures the same basic insight. It is also worth noting the contrast between the above argument and that of Cushman et al. (2006) and Greene et al. (2009), whose models rest on ease of learning or ease of mentally simulating a situation. It is not obvious to us how those models would explain that the omission–commission and means–by-product distinctions seem to depend on priors or be unique to settings of coordinated punishment.

Why Morality Depends on Categorical Distinctions

We explain why our moral intuitions depends so much more strongly on whether a transgression occurred than on how much damage was caused. Our argument again uses coordinated punishment and higher-order beliefs: *When a categorical distinction is violated, you know others know it was violated*, *but this is not always true for continuous variables*.

Consider the longstanding norm against the use of chemical weapons. This norm recently made headlines when Bashar al-Assad was alleged to have used chemical weapons to kill about a thousand Syrian civilians, outraging world leaders who had

shackelf@oakland.edu

been silent over his use of conventional weapons to kill over 100,000 Syrian civilians. A Reuters/Ipsos poll at the time found that only 9 % of Americans favored intervention in Syria, but 25 % supported intervention if the Syrian government forces used chemical weapons against civilians (Wroughton, 2013). In the past, the United States has abided by the norm against the use of chemical weapons even at the expense of American lives: In WWII, Franklin D. Roosevelt chose to eschew chemical weapons in Iwo Jima even though, as his advisors argued at the time, their use would have saved thousands of American lives. It might even have been more humane than the flame-throwers that were ultimately used against the Japanese ("History of Chemical Weapons," 2013). We say that the norm against chemical weapons is a categorical norm because those who abide by it consider whether a transgression was committed (did Assad use chemical weapons?), rather than focusing entirely on how much harm was done (how many civilians did Assad kill?). Other norms are similarly categorical. For instance, in the introduction to this chapter, we noted that across cultures and throughout history, the norm against murder has always been categorical: We consider whether a life was terminated, not the loss of useful life years. Likewise, discrimination (e.g., during Jim Crow) is typically based on categorical definitions of race (the "one drop rule") and not, say, the darkness of skin tone. Human rights are also categorical. A human rights violation occurs if someone is tortured or imprisoned without trial, regardless of whether it was done once or many times and regardless of whether the violation was helpful in, say, gaining crucial information about a dangerous enemy or an upcoming terror attack. We even assign rights in a categorical way to all *Homo sapiens* and not based on intelligence, sentience, ability to feel pain, etc.

Why is it that we attend to such categorical distinctions instead of paying more attention to the underlying continuous variable? We use game theory to explain this phenomenon as follows: Suppose that two players (say, the United States and France) are playing a Coordination Game in which they decide whether to punish Syria, and each wants to sanction only if the other sanctions. We assume the United States does not want to levy sanctions unless it is confident France will as well, which corresponds to an assumption on the payoffs of the game (if we reverse this assumption, it changes one line in the proof, but not the result).

We model the underlying measure of harm as a continuous variable (in our example, it is the number of civilians killed). For simplicity, we assume this variable is uniformly distributed, which means Assad is equally likely to kill any number of people. This assumption is, again, not crucial, and we will point out the line in the proof that it affects. Importantly, we assume that players do not directly observe the continuous variable, but instead receive some imperfect signal (e.g., the United States observes the body count by its surveyors).

Imagine a norm that dictates that witnesses punish if their estimate of the harm from a transgression is above some threshold (e.g., levy sanctions against Syria if the number of civilians killed is estimated to be greater than 100,000). As it turns out, this is not a Nash equilibrium. To see why, consider what happens when the United States gets a signal right at the threshold. The United States thinks there is a

shackelf@oakland.edu

50 % chance that France's estimates are lower than its own[5] and, thus, that there is a 50 % chance that France's estimates are lower than the threshold. This further implies that the United States assesses only a 50 % chance that France levies sanctions, so the United States is not sufficiently confident that France will sanction, to make it in the United States's interest to sanction.

What we have shown so far is that for a threshold of 100,000, it is in the interest of the United States to deviate from the strategy dictated by the threshold norm when it gets a signal at the threshold. This means that 100,000 is not a viable threshold, and (since 100,000 was chosen arbitrarily) there is no Nash equilibrium in which witnesses punish if their estimate of the harm from a transgression is above some arbitrary threshold.

It should be noted that this result only requires that there are sufficiently many possibilities, not that there is in fact a continuum. Neither does it require that the distribution is uniform nor that the Coordination Game is not affected by the behavior of Assad. The only crucial assumptions are that the distribution is not too skewed and that the payoffs are not too dependent on the behavior of Assad (for details, see Dalkiran et al., 2012; Hoffman, Yoeli, & Dalkiran, 2015).

What happens if such norms are learned or evolved and subject to selection? Suppose there is a norm to attack whenever more than 100,000 civilians are killed. Players will soon realize that they should not attack unless, say, 100,100 civilians are killed. Then, players will learn not to attack when they estimate 100,200 civilians are killed and so on, indefinitely. Thus, every threshold will eventually "unravel," and no one will ever attack.[6]

Now let's consider a categorical norm, for example, the use of chemical weapons. We again model this as a random variable, though this time, the random variable can only take on two values (0 and 1), each with some probability. Again, players do not know with certainty whether the transgression occurred, but instead get a noisy signal. In our example, the signal represents France or the United States's assessment of whether Assad used chemical weapons, and there is some likelihood the assessors make mistakes: They might not detect chemical weapons when they had been used or might think they have detected chemical weapons when none had been used.

Unlike with the threshold norm, provided the likelihood of a mistaken signal is not too high, there is a Nash equilibrium where both players punish when they receive a signal that the transgression occurred. That is, the United States and France each levy sanctions if their assessors detect chemical weapons. This is because when the United States detects chemical weapons, the United States believes France

[5] This is where the assumption of a uniform distribution comes in. Had we instead assumed, for instance, that the continuous variable is normally distributed, then it would not be exactly 50–50 but would deviate slightly depending on the standard deviation and the location of the threshold. Nevertheless, the upcoming logic will still go through for most Coordination Games, i.e. any Coordination Game with risk dominance not too close to .5.

[6] As with omission, this follows from iterative elimination of strictly dominated strategies (see Hoffman et al., 2015, for details).

shackelf@oakland.edu

likely detected them and will likely levy sanctions. So the United States's best response is to levy sanctions. Similarly, if the United States does not detect chemical weapons, it expects France did not and will not levy sanctions, so the United States is better off not levying them.

This result is useful for evaluating whether it is worthwhile to uphold a norm. The Obama administration was harshly criticized for threatening to go to war after the Assad regime used chemical weapons but not earlier, although the regime had already killed tens of thousands of civilians. The model clarifies that Obama's position was not as inconsistent as his critics had charged: The norm against chemical weapons may be worth enforcing since it is sustainable, whereas norms against civilian casualties are harder to sustain and hence might not be worthwhile to enforce.

Let's return to some more of our motivating examples. Our model can explain why we define murder categorically: It is not possible to punish differently for different amount of quality life years taken, but it is possible to punish differentially for a life taken. As with omission–commission, however, we do expect sadness or grief to depend greatly on life years lost, even if the punishment or moralistic outrage will be less sensitive. This is a prediction of the model that, as far as we know, has yet to be tested.

Similarly, the "one-drop" rule is a categorical norm, so it can be socially enforced in an apartheid society. In contrast, consider a rule that advocates giving up one's seat for someone with lighter skin. Since this is based on a threshold in a continuous variable, while it might be enforceable by a unilateral authority, it cannot be enforced by "mob rule." Other forms of discrimination, such as discriminating against the less attractive, or the less tall, or the elderly, all being continuous variables, cannot be socially enforced via coordinated punishment, and hence, we expect such discrimination to be of a different form. In particular, it will not be based not on punishing violators. For example, male CEOs might still prefer young attractive female secretaries, and taller men are more likely to be hired as CEOs, not because of coordinated rewards or punishment but because those who hire the CEOs or secretaries are likely to be satisfying their own preferences or doing what they expect will lead to higher profits.

Likewise, the number of victims tortured by a regime or the number of lives saved by torturing is continuous. Thus, a regime cannot be punished by a coordinated attack by other countries or by a coordinated rebellion by its citizens based on the number of people tortured or the paucity of reasons for such torture. But, a regime can be attacked or overthrown depending on whether a physical harm was inflicted on a citizen by the state. Hence, human rights are treated as inalienable, even in the absence of an a priori justification for this nonutilitarian norm. And why are human rights ascribed to all living *Homo sapiens*? Perhaps not because of a good logical a priori argument, but simply because violations of human rights are enforceable by coordinated punishment, but no regime can be punished for harming any "person" of less than a certain degree of consciousness.

Finally, here is one last application. The model might also explain why revolutions are often caused by categorical events, such as a new tea tax or a single, widely

shackelf@oakland.edu

publicized self-immolation, and not a breach of a threshold in, say, the quality of life of citizens or the level of corruption. This explanation requires simply that we recognize revolutions as a coordination problem (as argued in Morris & Shin, 2002; Chwe, 2013), where each revolutionary chooses whether to revolt, and each is better off revolting only if sufficiently many others revolt.

Quirks of Altruism and the Repeated Prisoner's Dilemma with Incomplete Information

The Repeated Prisoner's Dilemma has famously been used as an explanation for the evolution of cooperation among non-kin (Axelrod & Hamilton, 1981; Dawkins, 2006; Pinker, 2003; Trivers, 1971). *In this section, we show how the same basic model can be used to explain many of the quirky features of our pro-social preferences and ideologies.*

Recall that in the Prisoner's Dilemma, each of two players simultaneously chooses whether to cooperate. Cooperation reduces a player's own payoffs by $c>0$ while increasing the other's payoffs by $b>c$. The only Nash equilibrium is for neither player to cooperate. In the Repeated Prisoner's Dilemma, the players play a string of Prisoner's Dilemmas. That is, after the players play a Prisoner's Dilemma, they learn what their opponent did and play another Prisoner's Dilemma against the same opponent with probability δ (and the game ends with probability $1-\delta$). As is well known in the evolutionary literature, there are equilibria in which players end up cooperating, provided $\delta>c/b$. In all such equilibria, cooperation is sustained because any defection by one player causes the other player to defect. This is called reciprocity. As the reader is surely familiar, there is ample evidence for the Repeated Prisoner's Dilemma as a basis for cooperation from computer simulations (e.g., Axelrod, 1984) and animal behavior (e.g., Wilkinson, 1984). The model can be extended to explain contributions to public goods if, after deciding whether to contribute to a public good, players play a Repeated Prisoner's Dilemma (see, e.g., Panchanathan & Boyd, 2004) (Fig. 5).

The key to understanding these quirks is that players often have incomplete information. For example:

1. Players do not always observe contributions. It is intuitive that, for cooperation to occur in equilibrium, contributions need to be observed with sufficiently high probability.
2. Others cannot always tell whether a player had an opportunity to contribute. For defection to be penalized, it must be the case that others can tell that a player had the opportunity to cooperate and did not (i.e. the player should not be able to hide the fact that there was an opportunity to cooperate).
3. Sometimes, there are two ways to cooperate, and one has a higher benefit, b. Then, the only way this more effective type of cooperation can be sustained in equilibrium is if others know which cooperative act is more effective.

shackelf@oakland.edu

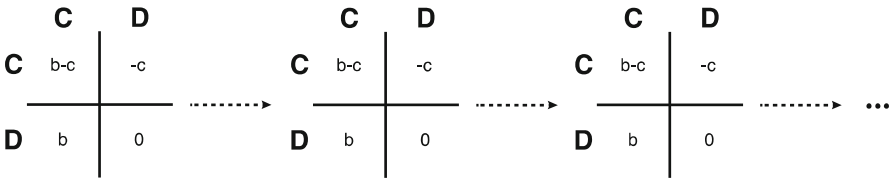

Fig. 5 The Repeated Prisoner's Dilemma. Two players play a Prisoner's Dilemma. They each observe the other's action, then, with probability δ, play another Prisoner's Dilemma against the same opponent, etc.

Technically, for the second and third point, what is needed is common knowledge that a player had an opportunity to cooperate or of the more effective means of cooperation. If observers were to know one purposely chose to defect or chose the less cooperative act, but they do not know that others know this, then observers think others will think punishment is not warranted, and observers will not punish. The argument is analogous to the discussion of higher-order beliefs in the omission–commission subsection and formalized in Dalkiran et al. (2012) and Hoffman et al. (2015).

Interpreting the Quirks of Altruism

Below we discuss some of the quirky features of altruism identified by economists and psychologists. In each case, we will argue that these features might be puzzling, but not when viewed through the lens of the above model:

Insensitivity to Effectiveness. We are surprisingly insensitive to the impact of our charitable contributions. We vote because we "want to be a part of the democratic system," or we "want to make a difference," despite the fact that our likelihood of swinging an election (even in a swing state) is smaller than our likelihood of being struck by lightning (Gelman, Silver, & Edlin, 2012). Why is our desire to "make a difference" or "be a part of the system" immune to the actual difference we are making? Our charitable contributions or volunteer efforts suffer from the same insensitivity. Why does anyone give money or volunteer time to Habitat for Humanity? The agency flies high earners who have never held a hammer halfway across the world to build houses that would be substantially more cheaply built by local experts funded by the high earners. Experimental evidence demonstrates our insensitivity: Experimental subjects are willing to pay the same amount to save 2000, 20,000, or 200,000 birds (Desvousges et al., 2010). Likewise, when donors are told their donations will be matched, tripled, or quadrupled, they donated identical amounts (Karlan & List, 2006). Why do we give so much, but do not ensure our gifts have a large impact?

The explanation follows directly from the above model: It is often the case that observers do not know which acts are effective and which are not and, certainly, this

shackelf@oakland.edu

usually is not commonly known. Thus, they will not reward or punish based on effectiveness, and we ourselves will not attend to effectiveness in equilibrium. This explanation suggests that if we want to increase efficacy of giving, we ought to focus on making sure donors' friends and colleagues are aware of the efficacy of different options. In fact, this is perhaps more important than informing the donor of efficacy, since the donor will be motivated to uncover efficacy herself.

Magnitude of the Problem. We are surprisingly unaware of and unaffected by the magnitudes of the problems we contribute to solving. How many of those who participated in the recent ALS Ice Bucket Challenge have even the vaguest sense of the number of ALS victims? (Answer: about 1/100th the victims of heart disease.) How much happier would these individuals have been if the number of ALS victims were cut in half? Multiplied by 100? The same questions could be asked about AIDS or cleft lips. If we were actually motivated by our desire to rid the world of such afflictions as we often proclaim, then we would be happier if there were fewer afflicted individuals and less happy if there were more. But we are not even aware of these numbers, let alone affected by them. This suggests an alternative motivation than the one we proclaim.

On the other hand, if we give in order to gain social rewards, it does not matter whether the problem is large or small, provided others recognize it as a problem and the social norm is to give. If our learned or evolved preferences were drastically impacted by the magnitude of the crises, we would be sensitive to whether the problem was solved, perhaps motivating us to ensure that others solve it, which we would not get credit for, or perhaps motivating us to devote too much of our resources to solving it, beyond what we would actually get rewarded for.

Observability. There is overwhelming evidence that people give more when their gifts are observed. Much of this evidence comes from the lab, where it has been demonstrated a myriad of ways (e.g., Andreoni & Petrie, 2004; Bolton, Katok, & Ockenfels, 2005; List, Berrens, Bohara, & Kerkvliet, 2004). For instance, when participants play a public goods game in the laboratory for money, their contributions are higher when they are warned that one subject will have to announce to the room of other participants how much they contributed (List et al., 2004). However, evidence also comes from real-world settings, which find large effects in settings as diverse as blood donation (Lacetera & Macis, 2010), blackout prevention (Yoeli, Hoffman, Rand, & Nowak, 2013), and support for national parks (Alpizar, Carlsson, & Johansson-Stenman, 2008). In Switzerland, voting rates fell in small communities when voters were given the option to vote by mail (Funk, 2006), which makes it harder to tell who did not vote, even though it also makes it easier to vote. In fact, our willingness to give more when observed extends to subtle, subconscious cues of being observed: People give twice as much in dictator games when there are markings on the computer screen that vaguely represent eyes (Haley & Fessler, 2005), and they are more likely to pay for bagels in their office when the payment box has a picture of eyes above it (Bateson, Nettle, & Roberts, 2006).

These results should not surprise anyone who believes our pro-social tendencies are influenced by reputational concerns (though the magnitudes are surprisingly large).

shackelf@oakland.edu

The effectiveness of subconscious cues of observability points to a primary role for reputations in our learned or evolved proclivities toward pro-social behavior. The large impact of subtle cues of observability, however, calls into question alternative explanations not based on reputations.

Explicit Requests. When we are asked directly for donations, we give more than if we are not asked, even though no new information is conveyed by the request. In a study of supermarket shoppers around Christmas time, researchers found that passersby were more likely to give to the Salvation Army if volunteers not only rang their bell but explicitly asked for a donation (Andreoni, Rao, & Trachtman, 2011). If our motive is to actually do good, or perhaps proximally to feel good by the act of giving, we should not be impacted by an explicit request.

However, if we evolved or learned to give in order to gain rewards or avoid punishment as described above, then we ought to be more likely to give when, if we did not give, it would be common knowledge that we had the option to give and chose not to. The explicit request makes the denial common knowledge.

It is worth emphasizing that our evolved intuition to respond to explicit asks may be (mis)applied to individual settings that lack social rewards. Imagine you are approached by a Salvation Army volunteer in front of a store in a city where you are visiting for one day only. A literal reading of the model would suggest that you should be no more likely to respond to an explicit request. But it is more realistic to expect that if your pro-social preferences were learned or evolved in repeated interactions then applied to this new setting, you would respond in a way that is not optimal for this particular setting and nonetheless give more when explicitly asked (just as our preferences for sweet and fatty foods, which evolved in an environment where food was scarce, lead us to overeat now that food is abundant).

Avoiding Situations in Which We Are Expected to Give. In the same supermarket study, researchers discovered that shoppers were going out of their way to exit the store through a side door, to avoid being asked for a contribution by the Salvation Army volunteers. In another field experiment, those who were warned in advance that a solicitor would come to the door asking for charitable donations were more likely to not be home. The researchers estimated that among those who gave, 50 % would have avoided being home if warned in advance of the solicitor's time of arrival (DellaVigna et al., 2012). In a laboratory analog, subjects who would have otherwise given money in a $10 dictator game were willing to pay a dollar to keep the remaining nine dollars and prevent the recipient from knowing that a dictator game could have been played (Dana et al. 2006). If our motive were to have an impact, we would not pay to avoid putting ourselves in a situation where we could have such an impact. Likewise, if our motive were to feel good by giving, we would not pay to avoid this feeling.

In contrast, if we evolved or learned to give in order to gain rewards or avoid punishment, then we would pay to avoid situations where we are expected to give. Again, this would be true even if, in this particular setting, we were unlikely to actually be punished.

shackelf@oakland.edu

Norms. People are typically *conditionally cooperative*, meaning that they are willing to cooperate more when they believe others contribute more. For example, students asked to donate to a university charity gave 2.3 percentage points more when told that others had given at a rate of 64 % than when they were told giving rates were 46 % (Frey & Meier, 2004). Hotel patrons were 26 % more likely to reuse their towels when informed most others had done the same (Goldstein, Cialdini, & Griskevicius, 2008). Households have been shown to meaningfully reduce electricity consumption when told neighbors are consuming less, both in the United States (Ayres, Raseman, & Shih, 2012) and in India (Sudarshan, 2014).

Such conditional cooperation is easily explained by the game theory model: When others give, one can infer that one is expected to give and may be socially sanctioned if one does not.

Strategic Ignorance. Those at high risk of contracting a sexually transmitted disease (STD) often go untested, presumably because if they knew they had the STD, they would feel morally obliged to refrain from otherwise desirable activity that risks spreading the STD. Why is it more reproachable to knowingly put a sexual partner at risk when one knows one has the STD than to knowingly put a sexual partner at risk by not getting tested? There is evidence that we sometimes pursue *strategic ignorance* and avoid information about the negative consequences of our decisions to others. When subjects are shown two options, one that is better for themselves but worse for their partners and one that is worse for themselves but better for their partners, many choose the option that is better for their partners. But, when subjects must first press a button (at no cost) to reveal which option is better for their partners, they choose to remain ignorant and simply select the option that is best for themselves (Dana, Weber, & Kuang, 2007).

This quirk of our moral system is again easy to explain with the above model. Typically, information about how one's actions affect others is hard to obtain, so people cannot be blamed for not having such information. When one can get such information easily, others may not know that it is easy to obtain and will not punish anyone who does not have the information. For example, although it is trivially easy to look up charities' financial ratings on websites like charitynavigator.org, few people know this and *could* negatively judge those that donate without first checking such websites. And even when others know that one can get this information easily, they might suspect that others do not know this, and so avoid punishing, since others won't expect punishment. To summarize, strategic ignorance prevents common knowledge of a violation and so is likely to go unpunished. We again emphasize that we will be lenient of strategic ignorance, even when punishment is not literally an option.

Norm of Reciprocity. We feel compelled to reciprocate favors, even if we know that the favors were done merely to elicit reciprocation and even if the favor asked in return is larger than the initial one granted (Cialdini 2001). For instance, members of Hare Krishna successfully collect donations by handing out flowers to disembarking passengers at airports, even though passengers want nothing to do with the flowers: They walk just a few feet before discarding them in the nearest bin.

shackelf@oakland.edu

Psychologists and economists sometimes take this "norm" as given, without asking where it comes from, and a naive reading of Trivers would lead one to think that we should be sensitive to the magnitude of the initial favor and whether it is manipulative.

However, according to the above model, reciprocity is the Nash equilibrium, even if the favors are not evenly matched or manipulative, since, in equilibrium, we are neither sensitive to such quantitative distinctions nor to whether the initial reciprocity was manipulative, unless these facts are commonly known.

Self-Image Concerns. People sometimes play mental tricks in order to appear *to themselves* as pro-social. For example, in an experiment, subjects will voluntarily take on a boring task to save another subject from doing it, but if given the option of privately flipping a coin to determine who gets the task, they often flip—and flip, and flip again—until the "coin" assigns the task to the other subject (Batson, Kobrynowicz, Dinnerstein, Kampf, & Wilson, 1997). Why would we be able to fool ourselves and not, say, recognize that we are gaming the coin flip? Why do we care what we think of ourselves at all? Are there any constraints on how we will deceive ourselves?

Such self-image considerations can be explained by noting that our self-image can act as a simple proxy, albeit an imperfect one, for what others think of us, and also that we are more convincing to others when we believe something ourselves (Kurzban, 2012; Trivers, 2011). This explanation suggests that the ways we deceive ourselves correspond to quirks described throughout this section—for example, we will absolve ourselves of remaining strategically ignorant even when it is easy not to, or be convinced that we have done good by voting, even if we cannot swing an election.

Framing Effects. Whether we contribute is highly dependent on the details of the experiment, such as the choice set (List, 2007) and the labels for the different choices (Ross & Ward, 1996; Roth, 1995). Such findings are often taken as evidence that social preferences cannot be properly measured in the lab (Levitt & List, 2007).

We believe a more fruitful interpretation is simply that the frame influences whether the laboratory experiment "turns on" our pro-social preferences, perhaps by simulating a situation where cooperation is expected (Levitt & List, 2007).

One-Shot Anonymous Giving: We give in anonymous, one-shot settings, such as dictator games. We also sacrifice for others in the real world when there is no chance of reciprocation: Heroes jump on grenades to save their fellow soldiers or block the door to a classroom with their bodies to prevent a school shooter from entering (Rand & Epstein, 2014). This is often seen as evidence for a role of group selection (Fehr & Fischbacher, 2003).

However, an alternate explanation is that we do not consider the likelihood of reciprocation (Hoffman et al., 2015), as described above. To explain the laboratory evidence, there are two more possibilities. First, subjects may believe there is some chance their identity will be revealed and feel the costs of being revealed as selfish are greater than the gains from the experiment (Delton, Krasnow, Cosmides, & Tooby, 2011). Second, we again emphasize that learned or evolved preferences and ideologies are expected to be applied even in novel settings to which they are not optimized.

shackelf@oakland.edu

Conclusion

In this chapter we have showed that a single approach–game theory, with the help of evolution and learning–can explain many of our moral intuitions and ideologies. We now discuss two implications.

Group Selection. Our chapter relates to the debate on group selection, whereby group level competition and reproduction is supposed to occasionally cause individuals to evolve to sacrifice their own payoffs to benefit the group (e.g., Wilson, 2006). One of the primary pieces of evidence cited in support of group selection is the existence of human cooperation and morality (Fehr & Fischbacher, 2003; Fehr, Fischbacher, & Gächter, 2002; Gintis, Bowles, Boyd, & Fehr, 2003; Haidt, 2012; Wilson, 2010, 2012), in particular: giving in one-shot anonymous laboratory experiments, intuitively sacrificing one's life for the group (jumping on the grenade), and contributions to public goods or charity. However, we have reviewed an alternative explanation for these phenomena that does not rest on group selection. It also yields predictions about these phenomena that group selection does not, such as that people are more likely to cooperate when they are being observed and there is variance in the cost of cooperation. The approach described here also explains other phenomena, such as categorical norms and ineffective altruism. These lead to social welfare losses, which is suboptimal from the group's perspective. The categorical norm against murder, for example, leads to enormous waste when keeping alive, sometimes for years, those who have virtually no chance of a future productive life.

Admittedly, despite their inefficiencies, these moral intuitions do not rule out group selection, since group selection can be weak relative to individual selection. But it does provide a powerful argument that group selection is unnecessary for explaining many interesting aspects of human morality. It also suggests that group selection is, indeed, at most, weak. One example that makes this especially clear is discrete norms. Recall that we argued that continuous norms are not sustainable because individuals benefit by deviating around the threshold. Notice that this benefit is small, since the likelihood that signals are right around the threshold is low. Group selection could easily overwhelm the benefit one would get from deviating from this Nash equilibrium, suggesting group selection is weak (i.e. there are few group-level reproductive events, high migration rates, high rates of "mutation" in the form of experimentation among individuals, etc.).

Logical Justification of Moral Intuitions. In each of the applications above, we explained moral intuitions without referring to existing a priori logical justifications by philosophers or others. Our explanation for our sense of rights does not rely on Locke's "state of nature." No argument we gave rests on God as an orderly designer, on Platonic ideals, on Kant's concepts of autonomy and humanity, etc. What does this mean for these a priori justifications? It suggests that they are not the source of our morality and are, instead, post hoc justifications of our intuitions (Haidt, 2012).

To see what we mean, consider the following analogy. One might wonder why we find paintings and sculptures of voluptuous women beautiful. Before the

shackelf@oakland.edu

development of sexual selection theory, one might have argued that perfect spheres are some kind of Platonic solid, and inherently desirable, or that curvy hips yield golden ratios. But with our current understanding of sexual selection, we recognize that our sense of beauty has evolved and that there is no platonic sense of beauty outside of that shaped by sexual selection. Any argument about perfect spheres is unparsimonious and likely flawed. Without the help of evolution and game theory, did philosophers conjure the moral equivalents of perfect spheres and golden ratios? The state of nature, the orderly designer, Platonic ideals, autonomy, and humanity, etc.—perhaps these arguments are also unfounded and unnecessary.

References

Alpizar, F., Carlsson, F., & Johansson-Stenman, O. (2008). Anonymity, reciprocity, and conformity: Evidence from voluntary contributions to a national park in Costa Rica. *Journal of Public Economics, 92*(5), 1047–1060.

Andreoni, J., & Petrie, R. (2004). Public goods experiments without confidentiality: A glimpse into fund-raising. *Journal of Public Economics, 88*(7), 1605–1623.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal, 100*, 464–477.

Andreoni, J., Rao, J. M., & Trachtman, H. (2011). *Avoiding the ask: A field experiment on altruism, empathy, and charitable giving*. Technical report, National Bureau of Economic Research.

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

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science, 211*(4489), 1390–1396.

Ayres, I., Raseman, S., & Shih, A. (2012). Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. *Journal of Law, Economics, and Organization*, 2–20.

Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters, 2*(3), 412–414.

Batson, C. D. (2014). *The altruism question: Toward a social-psychological answer*. Hillside, NJ: Psychology Press.

Batson, C. D., Kobrynowicz, D., Dinnerstein, J. L., Kampf, H. C., & Wilson, A. D. (1997). In a very different voice: Unmasking moral hypocrisy. *Journal of Personality and Social Psychology, 72*(6), 1335.

Bolton, G. E., Katok, E., & Ockenfels, A. (2005). Cooperation among strangers with limited information about reputation. *Journal of Public Economics, 89*(8), 1457–1468.

Cain, D., Dana, J., & Newman, G. (2014). *Giving vs. giving in.* Yale University Working Paper.

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton, NJ: Princeton University Press.

Chammah, A. M., & Rapoport, A. (1965). *Prisoner's dilemma; a study in conflict and cooperation*. Ann Arbor, MI: University of Michigan.

Chwe, M. (2013). *Rational ritual: Culture, coordination, and common knowledge.* Princeton, NJ: Princeton University Press.

Cone v. West Virginia Pulp & Paper Co., 330 U.S. 212, 67 S. Ct. 752, 91 L. Ed. 849 (1947).

Critcher, C. R., Inbar, Y., & Pizarro, D. A. (2013). How quick decisions illuminate moral character. *Social Psychological and Personality Science, 4*(3), 308–315.

Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment testing three principles of harm. *Psychological Science, 17*(12), 1082–1089.

shackelf@oakland.edu

Dalkiran, N. A., Hoffman, M., Paturi, R., Ricketts, D., & Vattani, A. (2012). Common knowledge and state-dependent equilibria. In *Algorithmic game theory* (pp. 84–95). New York: Springer.
Dana, J., Cain, D. M., & Dawes, R. M. (2006). What you don't know won't hurt me: Costly (but quiet) exit in dictator games. *Organizational Behavior and Human Decision Processes, 100*(2), 193–201.
Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Economic Theory, 33*(1), 67–80.
Davies, N. B. (1978). Territorial defense in the speckled wood butterfly (pararge aegeria): The resident always wins. *Animal Behaviour, 26*, 138–147.
Dawkins, R. (2006). *The selfish gene*. Oxford, UK: Oxford University Press.
DellaVigna, S., List, J. A., & Malmendier, U. (2012). Testing for altruism and social pressure in charitable giving. *The Quarterly Journal of Economics, 127*(1), 1–56.
Delton, A. W., Krasnow, M. M., Cosmides, L., & Tooby, J. (2011). Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proceedings of the National Academy of Sciences, 108*(32), 13335–13340.
DeScioli, P., Bruening, R., & Kurzban, R. (2011). The omission effect in moral cognition: Toward a functional explanation. *Evolution and Human Behavior, 32*(3), 204–215.
DeScioli, P., Gilbert, S. S., & Kurzban, R. (2012). Indelible victims and persistent punishers in moral cognition. *Psychological Inquiry, 23*(2), 143–149.
DeScioli, P., & Wilson, B. J. (2011). The territorial foundations of human property. *Evolution and Human Behavior, 32*(5), 297–304.
DeScioli, P., & Kurzban, R. (2009). Mysteries of morality. *Cognition, 112*(2), 281–299.
Descioli, P., & Karpoff, R. (2014). *People's judgments about classic property law cases*. Brandeis University Working Paper.
Desvousges, W. H., Johnson, F. R., Dunford, R. W., Boyle, K. J., Hudson, S. P., Wilson, K. N., et al. (2010). *Measuring nonuse damages using contingent valuation: An experimental evaluation of accuracy*. Research Triangle Park, NC: RTI Press.
Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology, 87*(3), 327.
Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature, 425*(6960), 785–791.
Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature, 13*(1), 1–25.
Fisher, R. A. (1958). *The genetic theory of natural selection*. Mineola, NY: Dover.
Frank, R. H. (1988). *Passions within reason: The strategic role of the emotions*. New York: WW Norton & Co.
Frey, B. S., & Meier, S. (2004). Social comparisons and pro-social behavior: Testing "conditional cooperation" in a field experiment. *American Economic Review, 94*, 1717–1722.
Funk, P. (2006). Modern voting tools, social incentives and voter turnout: Theory and evidence. In *Annual Meeting of the American Economic Association*, Chicago, IL.
Gelman, A., Silver, N., & Edlin, A. (2012). What is the probability your vote will make a difference? *Economic Inquiry, 50*(2), 321–326.
Gintis, H. (2007). The evolution of private property. *Journal of Economic Behavior & Organization, 64*(1), 1–16.
Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior, 24*(3), 153–172.
Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A room with a view-point: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research, 35*(3), 472–482.
Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition, 111*(3), 364–371.
Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York: Vintage.
Haley, K. J., & Fessler, D. M. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior, 26*.

shackelf@oakland.edu

Hedden, T., & Zhang, J. (2002). What do you think I think you think? Strategic reasoning in matrix games. *Cognition, 85*(1), 1–36.
Hoffman, M., Yoeli, E., & Nowak, M. A. (2015). Cooperate without looking: Why we care what people think and not just what they do. *Proceedings of the National Academy of Sciences, 112*(6), 1727–1732.
Hoffman, M., Yoeli, E., & Dalkiran, A. (2015). *Social applications of common knowledge*. Harvard University Working Paper.
Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1990). Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy, 98*, 1325–1348.
Kant, I. (1997/1787). Groundwork of the metaphysics of morals (1785). *Practical Philosophy, 108*.
Karlan, D., & List, J. A. (2006). *Does price matter in charitable giving? Evidence from a large-scale natural field experiment*. Technical report, National Bureau of Economic Research.
Kurzban, R. (2012). *Why everyone (else) is a hypocrite: Evolution and the modular mind.* Princeton, NJ: Princeton University Press.
Lacetera, N., & Macis, M. (2010). Social image concerns and prosocial behavior: Field evidence from a nonlinear incentive scheme. *Journal of Economic Behavior & Organization, 76*(2), 225–237.
Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *The Journal of Economic Perspectives, 21*, 153–174.
List, J. A. (2007). On the interpretation of giving in dictator games. *Journal of Political Economy, 115*(3), 482–493.
List, J. A., Berrens, R. P., Bohara, A. K., & Kerkvliet, J. (2004). Examining the role of social isolation on stated preferences. *American Economic Review, 94*, 741–752.
Locke, J. (1988/1688). *Locke: Two treatises of government student edition*. Cambridge, UK: Cambridge University Press.
Mankiw, N. G. (2007). *Principles of economics*. Thomson Learning.
Milgrom, P., & Weber, R. J. (1982). The value of information in a sealed-bid auction. *Journal of Mathematical Economics, 10*(1), 105–114.
Morris, S., & Shin, H. S. (2002). Social value of public information. *The American Economic Review, 92*(5), 1521–1534.
Myerson, R. B. (2004). Justice, institutions, and multiple equilibria. *Chicago Journal of International Law, 5*, 91.
Neuwirth, R. (2005). *Shadow cities: A billion squatters, a new urban world*. New York: Routledge.
Nowak, M. A. (2006). *Evolutionary dynamics: Exploring the equations of life*. Cambridge, MA: Harvard University Press.
Panchanathan, K., & Boyd, R. (2004). Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature, 432*(7016), 499–502.
Pinker, S. (2003). *The blank slate: The modern denial of human nature*. UK: Penguin.
Pinker, S., Nowak, M. A., & Lee, J. J. (2008). The logic of indirect speech. *Proceedings of the National Academy of Sciences, 105*(3), 833–838.
Rand, D. G., & Epstein, Z. G. (2014). *Risking your life without a second thought: Intuitive decision-making and extreme altruism*. Available at SSRN 2424036.
Ross, L., & Ward, A. (1996). Naive realism: Implications for social conflict and misunderstanding. In T. Brown, E. Reed, & E. Turiel (Eds.), *Values and knowledge* (pp. 103–135).
Roth, A. E. (1995). Bargaining experiments. In J. H. Kagel & E. R. Alvin (Eds.), *The handbook of experimental economics* (pp. 253–342). Princeton, NJ: Princeton University Press.
Sigg, H., & Falett, J. (1985). Experiments on respect of possession and property in hamadryas baboons (*Papio hamadryas*). *Animal Behaviour, 33*(3), 978–984.
Smith, J. M., & Price, G. (1973). The logic of animal conflict. *Nature, 246*, 15.
Sosis, R., & Alcorta, C. (2003). Signaling, solidarity, and the sacred: The evolution of religious behavior. *Evolutionary Anthropology: Issues, News, and Reviews, 12*(6), 264–274.
Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice. *Journal of Experimental Social Psychology, 27*(1), 76–105.

shackelf@oakland.edu

Sudarshan, A. (2014). *Nudges in the marketplace: Using peer comparisons and incentives to reduce household electricity consumption*. Technical report, Harvard University Working Paper.
Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization, 1*(1), 39–60.
Thomson, J. J. (1976). Killing, letting die, and the trolley problem. *The Monist, 59*(2), 204–217.
Tirole, J. (1988). *The theory of industrial organization*. Cambridge, MA: MIT Press.
Trivers, R. (2011). *The folly of fools: The logic of deceit and self-deception in human life.* New York: Basic Books.
Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46*, 35–57.
Trivers, R. L. (1974). Parent-offspring conflict. *American Zoologist, 14*(1), 249–264.
Wilkinson, G. S. (1984). Reciprocal food sharing in the vampire bat. *Nature, 308*(5955), 181–184.
Wilson, D. S. (2010). *Darwin's cathedral: Evolution, religion, and the nature of society*. Chicago: University of Chicago Press.
Wilson, E. O. (2012). *The social conquest of earth*. New York: WW Norton & Company.
Wilson, D. S. (2006). Human groups as adaptive units: Toward a permanent consensus. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind: Culture and cognition*. Oxford, UK: Oxford University Press.
Winter, E. (2014). *Feeling smart: Why our emotions are more rational than we think.* Public Books.
Wroughton, L. (2013). As Syria War Escalates, Americans Cool to U.S. Intervention: Reuter/Ipsos Poll. *Reuters*, August 24, 2013. Accessed at http://www.reuters.com/article/2013/08/25/us-syria-crisis-usa-poll-idUSBRE97O00E20130825 on 12 March 2015.
Yoeli, E., Hoffman, M., Rand, D. G., & Nowak, M. A. (2013). Powering up with indirect reciprocity in a large-scale field experiment. *Proceedings of the National Academy of Sciences, 110*(Suppl. 2), 10424–10429.
Zahavi, A. (1975). Mate selection: A selection for a handicap. *Journal of Theoretical Biology, 53*(1), 205–214.

shackelf@oakland.edu

Index

© Springer International Publishing Switzerland 2016

T.K. Shackelford, R.D. Hansen (eds.), *The Evolution of Morality*,
Evolutionary Psychology, DOI 10.1007/978-3-319-19671-8

shackelf@oakland.edu

shackelf@oakland.edu

shackelf@oakland.edu

shackelf@oakland.edu

shackelf@oakland.edu

shackelf@oakland.edu

shackelf@oakland.edu
View publication stats