MORAL CONCEPT THEORY: CULTURAL EVOLUTION, SOCIAL COGNITION, AND THE BRIDGE FROM COOPERATION TO MORALITY

Joshua Conrad Jackson

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirement for the degree of Doctor of Philosophy in the Department of Psychology and Neuroscience (Social Psychology).

Chapel Hill
2021

Approved by:

Kurt Gray

Masanori Takezawa

Katherine Gates

Keith Payne

Kristen Lindquist

ABSTRACT

Joshua Conrad Jackson: Moral Concept Theory: Cultural Evolution, Social Cognition, and the Bridge from Cooperation to Morality
(Under the direction of Kurt Gray)

Major puzzles remain in the science of moral psychology. Morality is theorized to be diverse and multidimensional, yet many people make moral judgments using a global dimension of "goodness" or "badness." Morality is theorized to be genetically evolved and universal, but the structure of morality varies widely across time, culture, and relationship. I address these puzzles with a new theory of morality. Moral concept theory claims that moral cognition co-evolved with human language as a mechanism for partner selection in cooperative dilemmas. In large social networks where people cannot observe and remember situation-specific moral concepts, generalized moral concepts become more functional, and spread throughout populations via prestige-biased transmission. Many people may now use a "global moral concept," a single concept of morality that people apply to all situations. Nine studies support these claims. Studies 1-2 shows evidence for a global moral concept that predicts cooperation across situations. Studies 3-4 find that social network size is correlated with belief in a global moral concept. Studies 5-6 show that a global moral concept outperforms granular moral concepts in large social networks (Study 5) and when interacting with unfamiliar partners (Study 6). Studies 7-8 shows evidence that moral concepts can be socially learned in a contemporary hunter-gatherer society (Study 7) and using historical patterns of word borrowing (Study 8). Study 9 is an agent-based model that formalizes predictions from moral concept theory. Moral concept theory reconciles biological and cultural evolutionary theories of moral cognition.

ACKNOWLEDGEMENTS

Thank you to my advisor, Kurt Gray, for supporting me these last five years. Kurt has made me a better thinker, writer, and researcher, all while giving me the freedom to make my own mistakes and learn from them. I also owe a debt of gratitude to all the old and new members of Kurt's Deepest Beliefs Lab, who made UNC a great place to exchange ideas and advice.

Thank you to my other academic mentors. I am particularly grateful for the mentorship of Michele Gelfand, Jamin Halberstadt, and Kristen Lindquist, who have all sacrificed time and energy to help me grow as a scholar.

Finally, a big thank you to friends and family who have made my time as a PhD student wonderful. My partner Nava Caluori, who has proofread many convoluted emails. My parents Francine Lorimer and Michael Jackson, who has suffered through many early drafts of talks. My close friends Joseph Leshin, Jason Hannay, and Tatum Jolink, who have always been there for a game night or a movie at the end of a long workday. And everyone else in the department who has given me a community and a family.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Many humans hold a deep conviction that some people are virtuous and good (moral*)* whereas others are nefarious and bad (immoral). People rate morality as more important than competence and warmth when they select surgeons, judges, close friends, coworkers, social acquaintances, and even storefront cashiers (Goodwin et al., 2014; Landy et al., 2018). They will endure physical pain to avoid even minor stains on their moral character (Vonasch et al., 2018), and both adults and children will endure personal costs in order to punish someone who has committed minor moral violations (Gummerum & Chu, 2014; McAuliffe et al., 2015; Vaish et al., 2011). More than 100 documented languages have a word for morality (Mikhail, 2007), and many modern religions view morality as the currency that determines how people experience the afterlife. Worldwide surveys show that people view moral character as more important than wealth, success, and even freedom (Vonasch et al., 2018).

Has morality always been so important? The answer depends on how you define morality. Defined as a cognitive system, morality is theorized to be as old as homo sapiens, and crucial to the early development of our species. Theories of evolution and human behavior argue that moral reputation to increase cooperation in groups, and it has served a similar function throughout human history (Curry et al., 2019; R. Dunbar & Dunbar, 1998; Haidt, 2001; Wright, 2010). Some scholars even claim that moral cognition has phylogenetic continuity in non-human animals such as chimpanzees, who use moral reputation to ostracize selfish conspecifics (DeWaal et al., 2006). Defined as a linguistic concept, however, *morality* is only several centuries old, adapted by Cicero to Latin as “moralis” from the Greek word “etikos.” In English

literature, “morality” was rarely used until the 19th century, and same is true in French (“moralité”), Russian (“мораль”), and Hebrew (“מוּסָרִיוּת”) (see Figure 1). These linguistic trends suggest that we have only been using words to communicate the concept of *morality* for a short period of time, and that the way we understand this concept may be very different from how early humans communicated about morality.

This paper introduces a new theory that combines insights from social psychology, cognitive anthropology, evolutionary biology, and comparative linguistics to show how an evolved process of moral cognition resulted in the historically recent concept of morality. Moral concept theory claims that the ability to recognize and track moral reputation co-evolved with human language, and that humans use different moral concepts (e.g., *caregiving ability, loyalty*) to communicate and store information about people’s cooperation in different situations (e.g., situations involving childcare, group-based conflict) which optimizes cooperative partner choice. As humans acquire language through social learning, they also acquire belief in moral concepts, and because social learning is sensitive to cultural context, so too are the moral concepts that people acquire. From this basis, I claim that the historical rise in social complexity and population size resulted in increasingly more general moral concepts which applied to a broader set of cooperative situations. Eventually this process resulted in a global moral concept which is communicated through the word “morality.” The global moral concept is unique because it applies to any cooperation-relevant situation, even though it is less accurate than situation-specific moral concepts. In this way, the global moral concept functions as a social heuristic that minimizes cognitive load at the occasional expense of accuracy (Kahneman et al., 1982).

Moral concept theory addresses a highly debated question about the dimensions of the moral domain. According to different theories, there are three (Shweder et al., 1997), five

(Graham et al., 2009), six (Iyer et al., 2012), or seven (Curry et al., 2019), dimensions of variance in moral judgment. Scholars have intensely debated the true number of moral dimensions, but here I suggest that there may not be any single solution to this dimension-reduction debate. Instead, morality's structure may vary based on the characteristics of social networks. In small groups of kin, moral cognition may be highly multi-dimensional, whereas in bustling cities of strangers or large online social media sites, judgments of morality may only vary along a single dimension of "goodness" or "badness." This variability may help humans optimize partner selection as their environments change.

In this paper, I draw from findings across the social and biological sciences to support moral concept theory. First, though, I briefly review some definitions of morality, which lay the groundwork for the theory.

Two Ways of Defining Morality

Morality: A process of moral cognition. Since the 1960s, a young field of moral psychology has studied "moral cognition" as the process in which humans make judgments of right and wrong. Early process theories of moral cognition focused on the roles of social learning and reasoning, with an emphasis on how humans learn to moralize some rules throughout the lifespan (Kohlberg & Hersh, 1977; Turiel, 2007). But these theories were displaced at the turn of the 21st century by bolder ideas about how emotion and intuition could influence the moral process. Drawing from new advances in evolutionary psychology and new methods of neuroimaging, these theories claimed that humans may have universal intuitions, which evolved genetically and were activated by specific neurobiological structures to make moral judgment (J. Greene & Haidt, 2002). One of the most successful of these theories, the social intuitionist model, claimed that "morality, like language, is a major evolutionary adaptation for an intensely

social species, built into multiple regions of the brain and body" (Haidt, 2001, p. 17). In social neuroscience, fMRI studies suggested that features of moral cognition such as mind-reading, empathy, and harm detection could be traced to specific brain areas (Farrow et al., 2001; J. D. Greene et al., 2001; Moll et al., 2002), and that there may be a fixed number of biologically-based moral intuitions shared by all humans.

For the last 20 years, scholars of moral cognition have searched for the universal moral intuitions that humans may have evolved as we became biologically modern. An early effort led to the CAD triad hypothesis, which argued that the three emotions of contempt, anger, and disgust had evolved to regulate moral judgments about communality, autonomy, and divinity (Rozin et al., 1999). Several years later, Moral Foundations Theory suggested that humans had evolved five major moral intuitions, made up of harm, loyalty, authority, purity, and fairness (Graham et al., 2009, 2013; Haidt & Joseph, 2004), adding liberty as a potential sixth intuition to accommodate American libertarians (Iyer et al., 2012). A more recent Morality-As-Cooperation hypothesis has proposed innate moral intuitions stemming from universal cooperation demands, such as helping kin, reciprocating, respecting property, and fighting in intergroup conflicts (Curry et al., 2019).

Theories of moral modularity have grown very popular in a very short time. Yet there have also been several challenges to the core assumptions of these models. The first challenge is neurobiological. An assumption of intuition-based models is that specific neurobiological modules regulate the emotions that give rise to moral judgment. However, large neuroimaging analyses have failed to find consistent brain structures or networks that regulate these kinds of emotional responses (Kober et al., 2008; Lindquist et al., 2012). Evidence has contradicted early claims that the orbitofrontal cortex regulates anger, the insula regulates disgust, and a meta-

analysis found no clear parallel between brain areas involved in moral emotions and those involved in moral judgments (Cameron et al., 2015). The second challenge is cultural. An assumption of intuition-based models is that certain moral intuitions are universal. Yet cross-cultural validation of moral judgment measures such as the moral foundations questionnaire have failed to find a single structure of morality that generalizes around the world (Iurino & Saucier, 2020). Other studies have found cultural variation in supposedly universal features of moral judgment, such as a bias towards condemning intentional (vs. unintentional) norm violations or an aversion to harmful behavior regardless of the status of the harmer (Barrett et al., 2016; Buchtel et al., 2015; Fessler et al., 2015). A third challenge is relational. Theories of modular morality suggest that people have a fixed set of moral values and a single moral domain, yet person-centered approaches to morality have consistently shown that people use a different standard of judgment to evaluate different relationship partners (Earp et al., 2020; Hester & Gray, 2020; Pizarro & Tannenbaum, 2012). Evidence from these person-centered approaches hint at a much more flexible moral domain which people continually adapt based on their interaction partner and their social context.

Evidence continues to accumulate that people do not have a universal set of discrete moral intuitions that can be traced to specific structures in the brain and generalize across relationships. Yet there is no widely accepted alternative theory to explain the evolution of moral cognition and current-day cultural diversity in moral cognition. Here I suggest that a good place to start formulating this alternative might be to briefly put aside the process of moral cognition and consider the concept of *morality*, or in other words, the meaning that people attach to words like "morality," and how these lay concepts of *morality* vary across individuals and groups.

Viewing morality as a concept helps shed light on how ordinary people may view the moral domain, and why words for *morality* have only recently emerged in history.

Morality: A global moral concept. In Raphael's famous fresco "School of Athens" (Figure 2), Plato and Aristotle are locked in debate about the nature of virtue. Plato, holding a copy of his *Timeo,* points upwards towards the heavens. His gesture represents his belief in the theory of non-physical forms, and the notion that there is a single "form" of the good that can be used as a yardstick to judge the inherent virtue of any behavior. Aristotle, holding his copy of *Etica,* points down to the earth, signaling his belief that virtue must be judged based on the affordances of specific people and situations. For example, when defining virtuous anger, Aristotle wrote that it must be "in accord with which we are angry, with whom we ought to be, at the things we ought, in the way we ought," alluding to the fact that anger alone was neither good nor evil but must be evaluated in its social context.

The fact that our current-day word "morality" was originally translated from Aristotle's "ēthikós" would suggest that Aristotle won his famous debate. But in an ironic twist, the modern meaning of the word bears stronger resemblance to Plato's form of the good. The Oxford English Dictionary defines "morality" as "pertaining to character or disposition, considered as good or bad, virtuous or vicious, of or pertaining to the distinction between right and wrong." Whereas Aristotle couched virtuous behavior in the local norms and situational affordances of a society, this modern definition paints *morality* as transcending place and person—a "global" concept that should predict any kind of good or evil behavior. The Oxford English Dictionary is not the only source of data on the conceptual breadth of *morality*. In a nationally representative survey of Americans, 73% agreed with the claim that "a person's morality is captured by that person's inherent goodness or badness, which manifests in different ways across different situations."

What can the emergence of a global moral concept tell us about human moral cognition? Some would argue that the emergence of concepts and words can tell us very little about psychology. There is a legacy in psychology of treating natural language as a smokescreen rather than a window into the mind, dating back to Freud's habit of dismissing intentional language and searching for slips of the tongue (Freud, 1989). Many social psychologists still assume that people use language to "tell more than they can know," and that real psychological processes must be understood using implicit or behavioral measures (Nisbett & Wilson, 1977). But this view is falling away, and many studies have shown that trends in natural language can yield insights into cultural and temporal variation in mental illnesses, social norms, emotion, and group identification (see Jackson et al., 2020). Many psychologists now agree that natural language is a valuable tool for studying psychological processes (Kennedy et al., 2021).

If changes to moral language do reflect changes in moral cognition, then linguistic trends suggest that people are making moral judgments in an increasingly global fashion. Other evidence supports this possibility. Studies of moral character perceptions show that ratings of moral traits such as "trustworthy," "honest," "fair," and "loyal" tend to form a one-factor solution with high intercorrelation, rather than the three-factor, five-factor, or seven-factor solution that would resemble theories of modular morality (Goodwin et al., 2014; Landy et al., 2016). Phrases such as "good person" and "bad person" are becoming increasingly more popular (see Figure 3), reflecting a growing number of words that allude to a global moral concept.

In the following sections, I explain how and why a global concept of morality may have emerged so recently in human history. First, I suggest that moral cognition may have emerged during the evolution of human language as a tool for choosing cooperative partners. I then show how historical trends in cultural evolution may have changed moral cognition and encouraged

the emergence of a global moral concept. Finally, I suggest some mechanisms for how moral concepts are learned and transmitted socially, focusing especially on prestige bias. At each stage, I provide evidence for these claims, drawing from experimental methods, cross-cultural surveys, ethnographic field studies, and archival analyses of language.

Moral Concept Theory

Moral cognition as a language-based tool for choosing cooperative partners. Theories of moral cognition may differ in many ways, but most of them agree that moral judgment has a close evolutionary relationship with sociality (Haidt, 2001), and with group-based cooperation in particular (Curry et al., 2019; Rai & Fiske, 2011). A wide range of unique animal behaviors facilitate cooperation. Dolphins will circle and protect a mother during childbirth, even at the cost of their physical safety (McBride & Kritzler, 1951). Vampire bats regurgitate blood to feed hungry conspecifics (Wilkinson, 1990). Belding ground squirrels make loud calls when they see a predator, giving their group time to find shelter but putting the caller in danger of attack (Dunford, 1977). These adaptations help group-based species of animals survive without widespread defection, and they have been the centerpiece of intense biological study for the last several decades (Jablonka & Lamb, 2007, 2007; Nowak et al., 2010).

Humans may not regurgitate our food for strangers or circle one another during childbirth, but humans can still maintain high levels of cooperation by remembering and transmitting information about each other's cooperative tendencies through moral reputation. Moral reputation is particularly useful for choosing cooperative partners, and this partner choice can make cooperation into a stable strategy that benefits individuals as well as groups (Apicella & Silk, 2019) (see Figure 4 for an illustration of these dynamics). Many animals show evidence of partner choice. Rhesus Macaques, velvet monkeys, elephant seals, lions, chimpanzees, and

even paper wasps will all stigmatize individuals who try to exploit the group (Cheney & Seyfarth, 1988; Hauser & Marler, 1993; Mills, 1991; Reeve & Nonacs, 1992; Reiter et al., 1978). But humans use partner choice more frequently than any other animal. Field studies of Dominican horticulturalists, Quichuan agro-pastoralists, and Australian Martu foragers have shown that partner choice can create patterns of competitive altruism where humans cooperate to gain social capital. Similar patterns of behavior emerge when students in the United States and United Kingdom play repeated economic games where individuals are able to select playing partners (Barclay & Willer, 2007; Sylwester & Roberts, 2010). Other strategies such as reciprocity and kin-based altruism can also create stable group-based cooperation (Gintis et al., 2008; Trivers, 1971), but there is evidence that partner choice in humans can lead to cooperation above and beyond these other mechanisms (Barclay, 2016).

Why are humans so adept at choosing cooperative partners? Many theories suggest that it is because of the human ability to use language. Humans use language via gossip to share information about cooperation and defection, meaning that people can track the moral reputation of individuals with whom they have little direct exposure (R. Dunbar & Dunbar, 1998). Several ethnographic field studies have supported the role of gossip for reducing free-riding and increasing cooperation in small-scale societies (Acheson, 1988; Boehm, 1997; Haviland, 1977; McPherson, 1991; Wilson et al., 2000), and experimental studies have found direct evidence that gossip can reduce selfishness in groups over time (Beersma & Van Kleef, 2012), especially in situations where a person might be vulnerable to exploitation because they have no previous exposure to an interaction partner (Feinberg et al., 2012, 2014; Fonseca & Peters, 2018). Case studies have observed potential forms of gossip in non-human animals (Leavens et al., 2014), but

the role of language in tracking and communicating moral reputation means that moral cognition likely had an outsized role in optimizing human cooperation compared to other animals.

Some theories claim that the benefits of moral gossip are so great that they may explain the evolution of language (R. Dunbar & Dunbar, 1998). However, there are many reasons to doubt this claim given the many other benefits of language for memory, categorization, and social cognition (Dor & Jablonka, 2001; Jablonka et al., 2012). A more likely claim is the reverse hypothesis: that language allowed for the evolution of moral reputation. Using language, humans could communicate moral reputation more easily and more frequently. They could also communicate moral reputation using either specific or general concepts. Situation-specific concepts like *caregiving ability* and *bravery* indicate whether people will cooperate in very select contexts, such as situations involving children or the elderly (*caregiving ability*) or in situations involving conflict (*bravery*). More general concepts like *loyalty* or *compassion* apply to a broader range of situations. For example, whereas *bravery* applies nearly exclusively to conflict, *loyalty* communicates cooperation in a wider range of situations where someone must defend the interests of their group. At the very extreme, a concept like *morality* seemingly applies to any situation involving cooperation (see **Studies 1-2** for evidence of the situational breadth of morality). In Figure 5, I use the Sierpinski triangle fractal to illustrate these situation-specific and more general ways of communicating morality.

The flexibility to use different moral concepts is important because there is a tradeoff to using situation-specific and situation-general information to predict cooperation. Using situation-specific moral concepts is useful because people's likelihood of cooperating varies widely across contexts. One of the defining studies of social psychology showed that many forms of social behavior only correlate at .20-.30 across situations (Mischel, 2013), a finding that launched a

crisis of conscience in personality psychology. In Pilot Studies a-c, I show that cooperative behavior also shows widespread situational variability. This situational variability means that situation-specific concepts of morality will nearly always yield more accurate predictions than situation-general concepts of morality since situation-general concepts average across meaningful context-dependent variance.

However, situation-specific concepts of morality are also cognitively costly. Theories of the social brain suggest that humans have strong constraints on their ability to remember and predict the social behavior of individuals in large groups (R. I. Dunbar, 2003, 2009). When a community is only made up of a few dozen kin-based ties, one will intimately know how these people cooperate. However, when groups consist of hundreds (or thousands) of people in a larger community, it is impossible to observe and encode rich information about one's potential interaction partners (R. I. Dunbar, 2014, 2016). These cognitive constraints should only escalate with the number of moral concepts that someone employs. It should be harder to remember how people in a large social group vary along five dimensions of morality than how they vary along a single dimension.

Morality is not the only cognitive system to face such a tradeoff. A large literature on judgment and decision-making shows that humans negotiate between costly algorithmic modes of making decisions and more efficient heuristic-based modes of making decisions about risk and reward (Johnson-Laird, 2006; Kahneman et al., 1982). Research now illustrates differences between these modes of thinking in decisions about resource gain (Allison & Messick, 1990) mating preference (Penke et al., 2008), and persuasion (Petty & Cacioppo, 1986). In each case, humans switch between "fast" and "slow" modes of reasoning depending on their social and affective context (Kahneman, 2011). For example, negative affect prompts more analytic

thinking (Baumann & Kuhl, 2002), and relationship closeness and perspective taking each decrease reliance on stereotypes in social judgments (Sherman et al., 2005; Wang et al., 2018). In the context of morality, people may adopt more situation-specific moral concepts when they interact with close partners or frequent interaction partners and may adopt more situation-general moral concepts when they interact with strangers or unfamiliar partners.

This evolutionary model of morality suggests that moral cognition co-evolved with language in humans, and that language provided humans with a way of categorizing and predicting cooperation across different situations using moral concepts. While some moral concepts are more situation-specific and cognitively costly, others are more situation-general and cognitively efficient. To balance the strengths and benefits of these concepts, humans can adapt the granularity of their moral domain based on interaction partners. And rather than a fixed number of moral intuitions, humans may have evolved a dynamic morality conceptual space that adapts based on social context.

This account of moral cognition is consistent with studies from functional neuroscience which have revealed little evidence for specific moral modules (Cameron et al., 2015), but considerable evidence for neural plasticity and connectivity between cognitive processes involving language and social cognition (Pehrs et al., 2018; Spreng & Andrews-Hanna, 2015; Spreng & Mar, 2012). It is also consistent with evidence from relationship science that people adapt moral judgments based on their level of familiarity with a relationship partner (Earp et al., 2020). Finally, this model of moral cognition sets the stage for understanding how morality may have changed across human history as groups grew larger and more complex.

The cultural evolution of moral cognition. Humans have probably lived in a diverse array of social groups since the late Pleistocene. In early anthropological theories of human

history, a “nomadic-egalitarian” model claimed that Pleistocene-era groups lived in small bands of 30-100 individuals (Fry et al., 2020; Hill et al., 2011; Kelly, 2013). Yet new evidence suggests that Pleistocene hunter-gatherers lived in surprisingly diverse groups, and that even the same hunter-gatherer band could move between relatively large sedentary settlements and smaller nomadic communities depending on the season (Singh & Glowacki, n.d.). Recent studies have shown that foragers in Papua New Guinea and the Pacific Northwest exploit coastal resources to maintain relatively large groups exceeding 1,000 individuals (Ames, 1994; Roscoe, 2006). Studies of nomadic foragers in Alaska and northern Australia observed that hunter-gatherers in both regions seasonally shifted between large sedentary settlements and dispersed mobile groups depending on rainfall, megafauna migration patterns, and temperature (Wengrow & Graeber, 2015; White & Peterson, 1969). Archaeological evidence from the Neolithic and Upper Paleolithic eras have found evidence of sedentary hunter-gatherer communities in Africa, East Asia, and Melanesia (Heinsohn, 2010; Marean, 2016; Snir et al., 2015).

The diversity of human groups in the late Pleistocene suggests that early humans probably needed to use different strategies for finding cooperative partners depending on seasonality, region, and resource availability. In difficult climates or inland jungles where hunter gatherer bands were small, situation-specific morality may have helped Pleistocene era humans optimize partner choice across different situations. Yet in coastal communities or in resource-abundant seasons with larger populations, more general moral concepts may have helped people predict cooperation of a larger number of potential social interaction partners. In other words, partner choice strategies likely had no universal optimum in early human history, making it unlikely that humans evolved a universal neurobiologically based structure to moral cognition. A more successful strategy could use language to communicate moral concepts that were

appropriate to a specific social network, and to adapt morality's conceptual space to the size and scale of a community.

This adaptation may have accelerated during the Holocene era, as technological and agricultural advances fundamentally changed the structure of social groups. As climates stabilized following the Ice Age, more human groups were able to live in sustainable sedentary communities that grew their own food (Gupta, 2004). Archeological sites have revealed groups such as the Hamburgian, Natufian, and Tell es-Sultan across Europe and Asia (Belfer-Cohen, 1991; Bruins & Van Der Plight, 1995; Grimm & Weber, 2008). Gobekli Tepe in modern-day Turkey showed evidence of large-scale religion as early as the 9th millennium BCE (Schmidt, 2000). Technology rapidly advanced during the Holocene, and by 3,000 BCE, cities with populations as large as 50,000 emerged in Mesopotamia, Ancient Egypt, and along the Yellow River in China (Turchin et al., 2018). In the following millennia, ancient civilizations rose and fell in Rome, Mongolia, and Zimbabwe. Christianity and Islam spread around the world, and the world witnessed the growth of the city-state, globalization, and the internet.

This latest chapter of human cultural evolution has had an outsized impact on social cognitive processes (Muthukrishna et al., 2020). Cultural evolutionary models suggest that perceptions of personality may have grown more multidimensional throughout the rise of social complexity as people filled more occupational and social niches (Smaldino et al., 2019). Studies of religious belief suggest that views of supernatural agents grew more punitive and intervening in socially complex societies (Norenzayan & Shariff, 2008), partly because these societies had stricter norms and harsher rules for deviants (Caluori et al., 2020; J. C. Jackson et al., 2020). A large-scale analysis of psychological variation found that the Catholic Church may have transformed many aspects of social cognition in Europe, increasing out-group trust, creativity,

individualism, and analytic thinking (Schulz et al., 2019). Linguistic analysis suggests that people's conceptualizations of emotion categories such as "fear" and "anger" vary across phylogenetic heritage, suggesting that emotion categories may have changed in their meaning over time as human groups diverged through migration and interacted through trade and warfare (J. C. Jackson et al., 2019). The psychologist and philosopher Cecilia Heyes has used these changes to suggest that many forms of social cognition are socially learned "gadgets" which can be adapted and taught like literacy (C. Heyes, 2018; C. M. Heyes & Frith, 2014).

Moral cognition may have also changed during this process of cultural evolution. But unlike personality, which may have grown more multidimensional, morality may have become increasingly one-dimensional as societies have grown larger and more complex. This hypothesis is grounded in the dynamics of partner choice in large-scale societies, and the prediction tradeoff between cognitive cost and accuracy during partner choice. In small-scale networks of kin, it is functional to use a variety of different cooperation-relevant concepts to choose cooperation partners. However, using specific traits comes with cognitive load, and in large-scale networks of kin it will be challenging to remember so many different pieces of information for all partners in one's social network. Under these conditions, a *social memory hypothesis* suggests that it may be more functional to use a smaller number of more general moral concepts. More general moral concepts will not be as accurate as granular concepts, but they will apply to a wider range of situations and will allow people to remember information for many interaction partners.

More general concepts of morality can also be adaptive by encouraging inferences of cooperation from past behavior in different contexts. With a general concept of loyalty, one can assume that a person who sacrificed personal gains to help their family will also be likely to sacrifice for their town. At the extreme end of generality, a global concept of morality will allow

inferences from any situation to generalize to any other situation (e.g., if someone is a bad caregiver, one may assume this person is also selfish, disloyal to their community, and deceitful). These inferences will sometimes be incorrect. In fact, social psychology has referred to this inference process as the "fundamental attribution error," because of its high error rate (Heider, 1983; Ross, 1977). But the strategy may still be functional on average within large and anonymous networks. Since people's cooperative behavior correlates moderately across situations, an *observation-inference hypothesis* suggests that someone using a global concept of morality will outperform someone with uninformed priors in these interactions. In support of this point, social psychological studies have found that people are more likely to make the fundamental attribution error when judging strangers compared to family members (Taylor, 1981), suggesting that they use situation-behavior inferences strategically.

Four studies in this paper are designed to test these hypotheses about social network size and moral structure. **Studies 3-4** are correlational surveys of people across the United States and around the world which test whether people living in larger social networks are more likely to endorse a global concept of morality. **Studies 5-6** are experimental designs that show how manipulating social network size and partner anonymity can increase the functionality of a global moral concept for predicting cooperation in social dilemmas. These studies suggest that cultural evolution has not only shaped how people use moral cognition to predict how different relationship partners will cooperate; this process has also shaped people's lay definitions of morality and encouraged belief in a global moral concept which dominates person-perception.

Evidence from the ethnographic record also supports this trend by showing that people in small-scale groups may view morality in terms of more differentiated and situation-specific concepts than people in complex Western societies. In his ethnography of the Tikopia people of

the current-day Solomon Islands, Raymond Firth wrote that the Tikopia morality was "more in the direction of protecting the interests of their own people than promoting moral principles" (Firth, 1936, p. 370). Raymond De Coccola made similar observations in his ethnography of the Copper Innuit people of the Arctic circle, writing that "Killing is a personal affair only if it does not weaken the group as a whole" (Coccola, King, & Houston, 1986, p. 167). Similar observations are visible in ethnographies of the Mbuti hunter-gatherer people of the current-day Congo (Beierle, 1995), and the Yanomamo people of South America (Chagnon, 2012), where murder was customary so long as it was between tribes and as part of a young man's initiation. Apart from being small-scale non-globalized societies, these groups had little else in common, procuring foods in different ways, speaking different languages, and following different religious traditions. But their intuitions about morality were each more context- and person-dependent than the Western emphasis on broad moral concepts.

These predictions show how more general concepts of morality may have grown more functional and prevalence in large social networks, but they do not illustrate the mechanism whereby general concepts of morality emerged and spread in these networks. To this end, new evidence suggests that humans may socially learn modes of moral cognition using many of the same strategies that they use to learn technological innovations or symbolic traditions.

Social Learning and Prestige Bias in the Transmission of Moral Cognition. Animals have a range of strategies that they use to transmit and receive culture (Kendal et al., 2018). Of these strategies, the "prestige bias" towards copying prestigious vs. low-status individuals may be particularly important for the cultural evolution of morality (Henrich & Broesch, 2011; Henrich & Gil-White, 2001). Prestige bias is also sometimes called "payoff bias" because it suggests that many animals decide to copy conspecifics when they see that some strategy has

yielded a successful payoff (Wood et al., 2013). Among non-human animals, stickleback fish will copy successful foraging behavior (Pike & Laland, 2010), and great tit birds will copy the strategy that conspecifics use to open a puzzle box and retrieve food (Aplin et al., 2015). Among humans, Fijian villagers will copy prestigious individuals' styles of fishing, growing yams, and using medicinal plants (Henrich & Broesch, 2011). Prestige bias can lead to "runaway selection" when prestigious individuals engage in maladaptive behavior (e.g., vaccine rejection) that is widely copied (Boyd & Richerson, 1988). But for the most part, prestige bias promotes functional cultural innovation since prestigious people have usually earned resources and status through successful behavior.

Prestige biased social learning is typically applied to visible behaviors such as foraging, agricultural practices, and symbolic markers like tattoos, but there is good reason to believe that moral cognition can also be learned and transmitted in a similar way. The inner workings of moral cognition may be invisible, but the language that humans use to communicate moral concepts can easily be learned and communicated. Words such as "loyalty," "generosity," and "honesty" all signal general moral concepts that apply across a range of situations. Words like "moral" and "bad person" go even further, signaling a global moral concept that applies across situations. In this sense, language provides a mechanism for modes of thinking to be socially transmitted just like visible behavior (C. Heyes, 2018). Individuals who use general moral concepts may become increasingly prestigious in large and complex societies, and their moral language will allow other people to copy their style of moral cognition.

The final two empirical studies in this paper are designed to test whether moral judgment strategies can be socially transmitted. **Study 7** investigates the structure of morality among a hunter-gatherer group, testing whether exposure to Western society predicts the structure of

moral judgments. **Study 8** tests whether moral information has been systematically exchanged between large and small societies over human history using patterns of word borrowing.

We acknowledge that other mechanisms such as conformist transmission (transmission based on descriptive norms) are also plausible candidates for transmitting moral concepts (Henrich & Boyd, 1998). However, prestige bias is often uniquely suited for transmitting functional cultural adaptations (Henrich & Gil-White, 2001), and so it is more plausible as a mechanism for socially learned morality, at least when moral concepts are in their early stages of transition.

The Present Studies

Moral concept theory claims that a significant evolutionary function of moral cognition was to optimize partner selection in cooperative dilemmas, and that the evolution of language allowed humans to communicate and categorize moral concepts that regulated reputation in social groups. These concepts were likely dynamic, meaning that humans can adapt the dimensionality of their moral space ranging from many fine-grained granular moral concepts (e.g., "loyalty to leaders," "willingness to share food") to more general concepts ("loyalty," "generosity"). This flexibility may have helped early humans during the Pleistocene as they moved between small nomadic bands and larger sedentary communities. It also laid the base for changes to moral cognition during the Holocene era as humans moved into increasingly larger and more complex societies. Throughout the Holocene, moral concepts may have become increasingly more general, and a global moral concept may have emerged, facilitated by the prestige-biased transmission of new words like "morality." People continue to vary today on their conceptual moral space, this variation may be particularly acute in small-scale societies that have recently experienced contact with Western culture.

I test these predictions across several empirical studies and an agent-based model. Studies 1-2 test whether the concept of "morality" is more global than other cooperation-relevant traits that have received attention in the literature. Studies 3-4 use large correlational surveys to test whether belief in the global concept of morality is more prevalent among people living in large social networks. Studies 5-6 experimentally test whether global morality is more functional than granular moral concepts because it helps remember people information about a greater number of interaction partners (Study 5) and because it helps people infer cooperation when interacting with unfamiliar partners (Study 6). Studies 7-8 test whether moral cognition can be socially learned by investigating moral judgments among the Hadza hunter-gatherers of southeastern Africa (Study 7) and by analyzing historical patterns of word borrowing across 44 languages (Study 8). Study 9 is an agent-based model that synthesizes these findings and predictions into a mathematical framework. Before presenting these studies, however, I outline three pilot studies that test the assumption in moral concept theory that people's likelihood of cooperation varies widely across different situations.

PILOT STUDY A: SITUATIONAL VARIABILITY ACROSS VIGNETTES

Will someone who is willing to share food also be willing to reciprocate labor? Will someone who offers to fight in battle also offer to take care of their neighbor's children? Although past studies of situation and behavior suggest that behaviors only correlate moderately (.20-.30) across situations (Mischel, 2013), most models of the evolution of cooperation treat cooperation as a global trait. An important assumption of moral concept theory is that the likelihood of cooperation varies across domains. My pilot studies test this assumption three times. First with stylized vignettes, second with economic game decisions, and third with real-world cooperative behavior. In each study, I use the intraclass correlation coefficient (ICC) to

calculate the proportion of variance that can be explained at the person level, and the average correlation between a person's behavior across situations. I predict that there will be substantial variability in cooperative behavior across situations in each context, with ICC values of below .50 representing less than 50% of variability in cooperation explained at the person level.

Method

Sample. I advertised for 300 participants on Amazon Mechanical Turk. In total, 289 participants (184 men, 101 women, 4 non-binary; M_{age} = 37.83, SD_{age} = 10.79) completed the study.

Measures. Participants viewed a series of 7 dilemmas that captured seven important domains of cooperation: generosity, caregiving ability, bravery, reciprocity, public goods, respect, and equity. These domains of cooperation are loosely adapted from Curry and colleagues (2019), who identified "helping kin," "helping group," "reciprocating," "being brave," "respecting superiors," "dividing resources," and "respecting property" as seven universally positive cooperative behaviors.

Before viewing and responding to the vignettes, participants read the prompt "Below are various dilemmas people will face in their daily lives. If you were in these dilemmas, how likely would you be to act as the dilemma describes? '1' means very unlikely, while '7' means very likely. When answering these questions **think about what you would actually do.** This is an anonymous survey, and we are hoping to capture how people would make these decisions in real life" Participants will read each vignette and respond to a Likert-type scale anchored at 1 (very unlikely) and 7 (very likely). Below is an example of a vignette for "generosity":

An acquaintance of yours has been living in the same house for their whole life, and it is falling apart. You have recently bought a large amount of building materials to build an extension to your house, and even after you are done building, you have enough to repair their house. On the other hand, if you keep the materials for yourself, you would have enough building materials for a renovation project in the future. How likely would you be to give them the materials?

Analytic plan. I restructured the data so that cooperative decisions were nested in participants. I then fit a random-effects analysis of variance (ANOVA), which decomposed variance of cooperative decisions into within- and between-subject parts. After fitting a random-effects ANOVA, I calculated the model's ICC using the formula $\frac{\tau_{00}}{\tau_{00}+\sigma^2}$, which allowed me to assess the variance explained by nesting within person. An ICC of 1 would suggest that all variance in cooperation decisions is explained at the person level, whereas an ICC of 0 would suggest that the nesting of the data is irrelevant because person effects do not explain any variance.

Results

Participants were most likely to cooperate in the reciprocity scenario (M = 6.10, SD = 1.23), followed by the generosity scenario (M = 5.36, SD = 1.23), the equity scenario (M = 5.38, SD = 1.66), the public goods scenario (M = 5.13, SD = 1.75), the respect scenario (M = 4.83, SD = 1.67), the caregiving scenario (M = 4.28, SD = 2.05), and the bravery scenario (M = 4.26, SD = 2.02).

A random effects ANOVA revealed an ICC of .14 across the scenarios, which translates to 14% of variation of cooperation explained by person-level effects. Variation in the means for

each scenario could plausibly result in an artificially low ICC. To examine this possibility, I estimated the ICC separately for the three scenarios with the highest mean level of cooperation (reciprocity, generosity, and equity) and again for the three scenarios with the lowest mean level of cooperation (respect, caregiving, bravery). The first model containing responses on the three highest scenarios yielded an ICC of .17 whereas the second model containing responses on the three lowest scenarios yielded an ICC of .08. Together, these models suggest that the low ICC was not driven by variation in the mean rate of cooperation across scenarios.

PILOT STUDY B: SITUATIONAL VARIABILITY ACROSS ECONOMIC GAMES

Method

Sample. I advertised for 300 participants on Amazon Mechanical Turk. In total, 306 participants (183 men, 117 women; M_{age} = 38.43, SD_{age} = 11.03) completed the study.

Measures and procedure. After consenting to participate, participants read that they were in a study focused on how people make decisions about social dilemmas. Instructions stated "these dilemmas are not arbitrary but will affect how much you are bonused for this study. In each dilemma, you can win anywhere from 0-10 additional 'bonus cents.' There are six dilemmas, and so your study 'bonus' will range from 0-60 cents (10 cents for each dilemma). For some of these decisions, you will be randomly paired with other people who are taking the study. However, these dilemmas are completely anonymous, and we will not share your decision with your partner and you will not be communicating with your partner. Please read the instructions for each dilemma carefully."

After reading these instructions, participants completed adapted versions of six common economic games: the prisoner's dilemma, the stag hunt, the public goods game, the dictator game, the traveler's dilemma, and the trust game. Each of these games is designed to measure

cooperative intent in different contexts. For example, the prisoner's dilemma captures contexts where cooperation comes at a personal expense, whereas the stag hunt represents a dilemma where mutual cooperation yields the highest payout to each partner. The dilemmas and their accompanied instructions to participants in this study are listed in Table 1.

Participants were not actually paired with partners, and this was disclosed during debriefing. However, the study's cover story was essential for ensuring that participants made realistic decisions rather than behaving in socially desirable ways (cooperating in all situations).

Analytic plan. Pilot Study B had the same analysis plan as Pilot Study A. Participants responded to economic games on different scales. For example, the response to a prisoner's dilemma is binary (cooperate or defect) whereas the response to a dictator game is continuously scaled from 1-10. I standardized responses to cooperative games during data cleaning so that the donation scales in the games did not artificially deflate the ICC value.

Results

Participants' decision tendencies for each game are displayed in Table 2. The random effects ANOVA revealed an ICC of .25 across the games, which translates to 25% of variation of cooperation explained by person-level effects. Within-subject variation in economic games was lower than for vignette responses, but still showed substantial variability.

PILOT STUDY C: SITUATIONAL VARIABILITY IN REAL-WORLD OUTCOMES

Method

Sample. Pilot Study C re-analyzed responses in the General Social Survey (GSS). The GSS is a large representative survey of Americans that assesses a variety of social, political, and spiritual attitudes, and has a representative sample of participants, which mitigates the non-

representative samples from Pilot Studies A-B. I will use the combined longitudinal dataset of GSS surveys from 1972-2010, which has approximately 53,000 data-points.

Measures. The primary measure of cooperation across situations was a series of items assessing cooperation in the last 12 months. Participants first read the prompt "During the past 12 months, how often have you done each of the following things?" and then indicated responses on a 6-point scale ranging from "Not at all in the past year" to "More than once a week" for the following behaviors: (a) donated blood, (b) given food or money to a homeless person, (c) returned money to a cashier after getting too much change, (d) allowed a stranger to go ahead of you in line, (e) done volunteer work for a charity, (f) given money to charity, (g) offered your seat on a bus or in a public place to a stranger who was standing, (h) looked after a person's plants, mail, or pets while they were away, (i) carried a stranger's belongings, like groceries, a suitcase, or shopping bag, (j) given directions to a stranger, and (k) let someone you didn't know well borrow an item of some value like dishes or tools.

Some of these items may not be applicable for participants (e.g., participants may never have gotten too much change). However, the GSS allows participants to indicate "don't know," "no answer," and "not applicable," meaning that they will not be forced to respond if they do not have an appropriate answer. These responses were converted to "NA" values prior to analysis.

Analytic plan. I used the same analysis plan as Studies 1-2. I replicated analyses for the full range of behaviors, and for only the behaviors that do not constitute everyday interactions (a,b,e,f from the list in the "measures" section) so that variation in the frequency of cooperative behaviors was not confounded with variation in the feasible frequency of behaviors. For example, it is only possible to give directions when you live in an area where people are frequently lost and only feasible to look after a person's plants if you know someone who is

about to travel but giving blood and giving to charity requires active initiation, so it is less susceptible to bias because of opportunity.

Results

Table 3 shows the mean frequency for each form of cooperative behavior in this study. The most frequent form of cooperation was allowing a stranger to go ahead of you in line and the least frequent was donating blood.

The random effects ANOVA revealed an ICC of .10 across the behaviors, which translates to 10% of variation of cooperation explained by person-level effects. Narrowing the behaviors to everyday interactions resulted in an ICC of .12, showing that the low ICC could not purely be attributed to variation in people's opportunity to cooperate.

Pilot Studies Discussion

The purpose of these pilot studies was to confirm that cooperation varies significantly across situations, just like many other forms of social behavior (Mischel, 2013). I found that people's cooperation in hypothetical vignettes, economic games, and real-world outcomes all varied substantially across situations, with only 10-25% of variance reduced to person effects. This finding is important because an assumption of moral concept theory is that the optimal way of predicting cooperation in groups is to develop situation-specific moral concepts, whereas a global moral concept will be less accurate because of situational variability in people's likelihood to cooperate. My pilot studies supported this assumption. I next turned to testing the core predictions of moral concept theory, such as the prediction that the word "morality" is a global moral concept that people see as similarly predictive of any cooperative situation.

STUDY 1: EVIDENCE FOR A GLOBAL MORAL CONCEPT

The word "morality" may reflect a global moral concept that contrasts with more situation-specific concepts such as "bravery," "childcaring ability," and "loyalty" that are highly useful for predicting behavior, but only in certain select situations. If "morality" reflects a global moral concept, people should perceive morality as predicting behavior, on average, at a greater level than any other trait across situations. Morality should also vary less in its predictive validity across situations since it is a situation-general concept. Finally, morality should never be the most predictive moral concept for any given situation. In other words, morality will always be perceived as moderately important for predicting cooperation, whereas other traits will either be very important or not at all important. Studies 1-2 tested these hypotheses with a vignette-based design where participants read about the cooperative dilemmas from Pilot Study A and rated the predictiveness of a series of different traits. Study 1 used a box-checking measure where people rated whether a trait was relevant or irrelevant for predicting cooperation in a situation, whereas Study 2 used a Likert-style measure.

Sample. I advertised for 200 participants on Amazon Mechanical Turk and included an attention check to exclude low-quality participants. Near the end of the study, participants were asked to choose their favorite hobby from a list and asked to write "gardening" if they were paying attention. In total, 193 participants (90 men, 101 women, 2 non-binary; M_{age} = 38.92, SD_{age} = 11.59) completed the study and passed the attention check.

Measures and Procedure. Participants read the same 7 vignettes depicting cooperation across situations as in Study 1. However, rather than reporting how they would act in these vignettes, they will be asked to reflect on how someone else would act, and the type of information that would help them predict someone else's cooperation. For example, the vignette depicted in the Pilot Study A measures section was adjusted so that participants read:

You have been staying in your current house your whole life, and it is falling apart. An acquaintance of yours has recently bought a large amount of building materials to build an extension to their house, and even after they are done building, they have enough to repair your house. You are thinking of asking them for some of their materials, but you are afraid of being rejected.

The full set of scenarios is listed in Appendix A. After each scenario, participants were asked "which of the following traits would be relevant for deciding whether this person can be trusted in this situation? You may check multiple traits, and we would like you to select all that apply." Participants saw the traits "morality," "caregiving ability," "group loyalty," "reciprocity," "interpersonal respect," "allocation fairness," and "resource generosity." These traits have all received attention in research on the moral domain (Curry et al., 2019; Haidt & Joseph, 2007).

Analytic plan. I tested my predictions with three analyses. First, I used a series of multi-level regression model to test for the most chosen trait in each scenario. Since responses in this study were binary, I used logistical regression with intercepts randomly varying across participants. I predicted that morality would never be the most clicked trait in a specific situation. Second, I used a second multi-level regression to test which trait was clicked most *on average* across the studies. Averaging the values across situations resulted in a normal distribution of proportions, so I fit a multi-level model with a gaussian distribution and intercepts varying randomly across participants. I predicted that morality would have the highest mean rate of clicking across situations, even though it was not the most clicked trait in any specific situation.

Finally, I used the same modeling approach to test for the standard deviation of traits across situations. I predicted that morality would have the lowest standard deviation of any trait, since other traits would be viewed as better matches to the situation.

Results

My prediction that morality would not be the most selected trait for a given situation was supported in all but one scenario. For the discretion scenario, morality was selected more frequently than respect, which was the next highest selected trait ($OR = 1.59$, $p < .001$). For all other scenarios, other traits were selected more than morality. For the property scenario, generosity was selected more frequently than morality ($OR = 6.51$, $p < .001$). For the kin scenario, caregiving ability was selected more frequently than morality ($OR = 2.34$, $p < .001$). For the group conflict scenario, loyalty was selected more frequently than morality ($OR = 5.73$, $p < .001$). For the building materials scenario, reciprocity was selected more frequently than morality ($OR = 7.74$, $p < .001$). For the public goods scenario, loyalty was selected more frequently than morality ($OR = 5.96$, $p < .001$). And for the land plot scenario, fairness was selected more frequently than morality ($OR = 4.26$, $p < .001$).

I next tested for the average mean and standard deviation for each trait across scenarios. As predicted, morality had the highest mean across scenarios, but the lowest standard deviation across scenarios (see Table 4). This suggests that, even though morality was only deemed the most relevant trait in one scenario, it was significantly selected more than any other trait when averaging across all scenarios, and it was selected at a similar level across scenarios. These results support the role of morality as a global concept that is viewed as predictive across diverse situations. Figure 6 displays the mean and standard deviation across scenarios for each trait in Study 1.

STUDY 2: REPLICATING EVIDENCE FOR A GLOBAL MORAL CONCEPT

Method

Sample. I advertised for 200 participants on Amazon Mechanical Turk and included the same attention check as Study 1. In total, 198 participants (102 men, 96 women; M_{age} = 39.34, SD_{age} = 11.08) completed the study and passed the attention check.

Measures, and analytic plan. Study 2 had the same measures, procedure, analytic plan, and predictions as Study 1 with one key change: participants used a Likert-type response scale rather than a binary measure. That is, instead of clicking traits that are relevant to a situation, participants rated each trait on its importance for predicting cooperation in the situation on a scale from 1 ("Not at all important") to 7 ("Very important").

Results

How predictive did people rate different traits across situations? Whereas Study 1 showed partial support for the prediction that morality was never be the most endorsed trait for a given situation, Study 2 fully supported this hypothesis. For the property scenario, generosity was selected more frequently than morality (b = 1.68, p < .001). For the kin scenario, caregiving ability was selected more frequently than morality (b = .55, p < .001). For the group conflict scenario, loyalty was selected more frequently than morality (b = 1.14, p < .001). For the building materials scenario, reciprocity was selected more frequently than morality (b = 1.33, p < .001). For the public goods scenario, loyalty was selected more frequently than morality (b = 1.20, p < .001). For the discretion scenario, respect was selected more (albeit not significantly more) than morality (b = .20, p = .16). And for the land plot scenario, fairness was selected more frequently than morality (OR = .59, p < .001).

I next tested for the average mean and standard deviation for each trait across scenarios. As predicted, morality had the highest mean across scenarios, but the lowest standard deviation across scenarios (see Table 5). This suggests that, even though morality was only deemed the most relevant trait in one scenario, it was significantly selected more than any other trait when averaging across all scenarios, and it was selected at a similar level across scenarios. However, in this study, the mean of morality was not significantly greater than the mean of respect, and the standard deviation of morality was not significantly lower than the standard deviation of respect. Together with Study 1, however, these results offer support for the idea that morality is a concept that people apply broadly across situations, and not exclusively to specific situations.

Discussion

Studies 1-2 offered further support that the word "morality" is a global concept that people assume predicts cooperation across situations. In both studies, morality was rated as more predictive of cooperation than other traits such as loyalty, reciprocity, and respect, and ratings of morality's predictiveness varied less across scenarios—with a lower standard deviation across scenarios—than these other traits. In Study 2, the difference between morality and respect did not reach significance. This suggests that respect may also be a highly general moral concept that is judged as predictive of cooperation across situations.

Even though people viewed morality as highly connected to cooperation, they almost never judged morality as the *most* predictive trait for cooperation in any single situation. For example, when people predicted someone's likelihood of sharing resources, they valued the trait of generosity over morality. In Study 1, there was one exception to this rule: participants viewed morality as the best trait for predicting cooperation in a scenario where someone needed to be discrete with a shameful secret. However, this effect disappeared in Study 2. This pattern of

results suggests that, while many people do use a global concept of morality, they are still sensitive to variation in cooperation across situations and the ability of more granular traits to predict cooperation-specific behavior.

Studies 1-2 demonstrated the existence of a global moral concept. In Studies 3-4, we examined variation in people's belief in this concept, testing whether people living in larger social networks would be more likely to endorse and use a global moral concept than people living in smaller social networks.

STUDY 3: BELIEFS IN A GLOBAL MORAL CONCEPT IN THE UNITED STATES

Studies 3-4 developed a scale to test whether people living in larger social networks would be more likely to endorse and use a global moral concept in everyday cooperative dilemmas. There are many ways of assessing social network size, but these studies focused on two metrics of social network size that would capture the breadth of people's everyday interactions in person, via people's estimates of their daily interaction partner, and online, via the number of friends people had on Facebook. We predicted that individuals with more Facebook friends and more daily interaction partners would be more likely to endorse and use a global moral concept. We also predicted that this relationship would replicate controlling for age, gender, education status, religious belief, and socioeconomic status. Study 3 tested these predictions across a representative sample of Americans, and Study 4 tested the predictions across four nations.

Method

Sample. I advertised for 2000 participants using the Qualtrics panels service. I determined sample size by recruiting as many participants as possible given the cost constraints of the panel service. Participants were pseudo-representative in the sense that they were recruited

to be nationally representative on the key dimensions of age, political party affiliation, race, and region of the country (South, Northeast, Midwest, West). In total, 2011 participants (504 men, 1501 women, 6 non-binary; M_{age} = 33.49, SD_{age} = 16.40) signed up for the study and completed all measures.

Measures.

Belief and importance of morality. I designed a simple scale to assess participants' belief in a global moral concept, and their self-reported reliance on a global moral concept in everyday life. I designed the scale as a two-factor measure with one 3-item factor measuring belief in a global moral concept ("at their core, people are either morally good or morally evil," "every person has a basic good or evil moral character," "All human behavior stems from people's underlying morality") and another 3-item factor measuring reliance on a global moral concept ("I often think about people's morality when I interact with them," "when deciding whether to trust someone, I try to gauge their underlying moral character," "I rarely, if ever, need to gauge someone's underlying moral character"; final item is reverse-coded). Participants responded to each item on a 1-100 scale anchored at 1 ("Strongly Disagree") and 100 ("Strongly Agree").

Factor analysis revealed support for a 2-factor solution with one factor (Eigenvalue = 2.57; Global morality belief) containing loadings over .40 from items 1-3 and a second factor (Eigenvalue = 1.07; Global morality reliance) containing loadings over .40 from items 4-6. I averaged these two sets of items into scores representing each of the factors.

Size of social network. I operationalized social network size using two key metrics: participants' self-reports of how many people they interact with every day (as a proxy for the size of their in-person social network), and participants' self-reported number of friends on Facebook (as a proxy for the size of their virtual social network). Participants were excluded

from these analyses if they listed more than 5,000 friends on Facebook since Facebook does not allow more than 5,000 friends. Participants were also excluded if they reported interacting with more than 1,000 unique people per day, which would require a unique interaction every 43.2 seconds of a 12-hour waking day. This procedure excluded 11 participants, since that 2000 were included in analysis.

Demographics. I measured age, self-identified gender, SES, religiosity, and education. SES was operationalized based on people's responses to the McArthur Ladder item, which asks people to rate themselves higher or lower on an 11-rung ladder, where higher values represent people with the most money, highest education, and best jobs. Religiosity was operationalized using responses to the Supernatural Beliefs Scale (Jong et al., 2013). Education was operationalized using a dummy-coded measure of whether people have completed a 4-year college degree.

Analytic plan. Both number of Facebook friends and number of daily interaction partners were positively skewed, so I log-transformed them prior to analysis. I began testing hypotheses with zero-order correlations between each factor of my global morality scale and each metric of social network size. I then replicated these analyses with more rigorous multiple regressions which controlled for demographic characteristics.

Results

Global morality belief was correlated robustly with global morality reliance, $r(1998) = .42$, $p < .001$, suggesting that the same people who belief in a global concept of morality also use this concept to predict people's behavior. Zero-order correlations showed that global morality belief was significantly associated with number of Facebook friends, $r(1998) = .05$, $p = .02$, but not the number of everyday interactions partners, $r(1998) = -.003$, $p = .88$. Global morality

reliance was correlated with both number of number of Facebook friends, $r(1998) = .11$, $p < .001$, and number of everyday interactions partners, $r(1998) = .10$, $p < .001$.

I next estimated how these associations changed when controlling for other demographic characteristics. Table 6 displays results from four multiple regressions where global morality belief and global morality importance were separately regressed on each of the social network size predictions controlling for demographic characteristics. These regressions showed broadly similar results to our zero-order correlations. Both social network size proxies predicted increased global morality reliance, and number of Facebook friends (but not number of everyday interaction partners) predicted global morality belief. This pattern of results may arise because people's reliance on global morality is more sensitive to social conditions than their belief in global morality, which may be more stable across the lifespan, but it could also reflect sampling error or random noise. Study 4 allowed me to test whether these findings replicated in an international sample.

STUDY 4: BELIEFS IN A GLOBAL MORAL CONCEPT IN FOUR COUNTRIES

Method

Sample. I advertised for 1000 participants from four nations (The United States, Brazil, Singapore, and Germany) using Qualtrics panels. I chose these nations because they cover different world regions and they vary in their cultural tightness: the strictness of cultural norms (Gelfand et al., 2011). Since cultural tightness is related to moralization (Jackson et al., 2020), I sought to test whether the relationship between moral beliefs and social network size was robust in a sample including people from both tight and loose cultures. I recruited 1000 participants because this was the largest sample I could afford given Qualtrics panels pricing. The United States and Singapore are native English-speaking countries, but Brazil and Germany are not.

Therefore, native speakers translated the survey from English into Portuguese and German for these speakers using standard translation and back-translation procedures.

In total, 1044 participants (484 men, 560 women; M_{age} = 45.44, SD_{age} = 16.04) completed the survey. Of these participants, 267 were from Singapore, 261 were from the United States, 260 were from Germany, and 256 were from Brazil.

Measures, procedure, and analytic plan. Study 4 had the same measures, procedures, and analytic plan as Study 3. For regression analyses, I added dummy-coded variables representing each country to remove country-specific variance. The measure of global morality belief and global morality reliance showed the same two-factor structure as in Study 3.

Results

Global morality belief correlated with global morality reliance, $r(970) = .32$, $p < .001$, replicating Study 3's finding that the same people who belief in a global concept of morality also use this concept to predict people's behavior. As in Study 3, zero-order correlations showed that global morality belief was significantly associated with number of Facebook friends, $r(970) = .14$, $p < .001$, but not the number of everyday interactions partners, $r(970) = .06$, $p = .07$. Global morality reliance was correlated with both number of number of Facebook friends, $r(970) = .14$, $p < .001$, and number of everyday interactions partners, $r(970) = .12$, $p < .001$.

Regression estimates also mirrored these correlations. Table 7 displays these results. Each measure of global morality correlated with each measure of social network size. Of note, the relationship between global morality belief and number of everyday interaction partners reached significance when controlling for demographic characteristics, which I did not find in Study 3. These results offered consistent support for the association between using a global moral concept and social network size.

To summarize the results of Studies 3-4, Figure 7 displays the mean of the global morality scale (collapsed across the two subscales) regressed against each of the measures of network size. This figure displays a clear positive relationship between global morality endorsement and social network size across each sample.

Discussion

Studies 3-4 found evidence for a correlation between social network size and endorsement of a global concept of morality. People who had larger networks of Facebook friends believed in a global moral concept and reported relying on a global moral concept at greater rates than people with smaller numbers of Facebook friends. People's self-reported number of daily interaction partners was robustly associated with reliance on a global concept of morality and was linked to belief in a global concept of morality in Study 4 but not Study 3. Associations were consistent across a representative sample of Americans and a sample from four different nations, and were robust to controlling for SES, age, gender, religiosity, and education.

These findings show evidence that people are more likely to believe in a global concept of morality in larger compared to smaller social networks. Since this evidence was correlational, it cannot establish a mechanism behind this relationship, but Studies 5-6 test two hypotheses for why social network size could causally increase the prevalence of a global moral concept.

STUDY 5: EXPERIMENTALLY TESTING THE SOCIAL MEMORY HYPOTHESIS

Studies 5-6 experimentally tested whether a global concept of morality can optimize decisions in cooperative dilemmas. Moral concept theory offers two hypotheses for how more general concepts of morality can maximize cooperation in large social networks. The first of these hypotheses is the social memory hypothesis: that using a smaller number of more general

morality concepts can help people remember information about a larger number of potential social partners than using a larger number of more granular morality concepts. Study 5 examined support for the social memory hypothesis by manipulating social network size and testing whether a more global concept of morality (which I both manipulated and measured) predicted cooperation-optimizing trust decisions in the large social network condition vs. the small social network condition.

Method

Sample. I advertised for 1000 participants from Amazon Mechanical Turk to take part in this study. In total, 2,055 participants signed up to take part, and 1,135 participants (493 men, 638 women, 4 non-binary; M_{age} = 38.10, SD_{age} = 13.11) completed the study and passed our attention check, which was the same attention check we used in Studies 1-2.

Measures and procedure. All participants in the study followed a similar procedure and completed a similar set of measures. Participants were asked to imagine that they lived in an imaginary community where they frequently needed to anticipate whether other people would cooperate with them. Participants were told that, throughout the study, they would be able to see the characteristics of each person in their community and take notes on these characteristics. They would then be able to memorize their notes and use them to make decisions in hypothetical cooperative dilemmas.

After reading the initial instructions, each participant viewed profiles, one at a time, for each member of their hypothetical community. These profiles contained the person's name ("Alice"), and 1-100 scores on seven traits such as "generosity" and "reciprocity" that matched the traits in Studies 1-2. Scores on each profile were simulated to match the ICC from Pilot Study A so that profiles had realistic variation across different kinds of cooperation. When

viewing the traits, participants could take notes on any information that would help them predict the person's cooperation later in the study.

After viewing each trait, participants were shown their notes from the profiles on a single screen and were given the opportunity to review and memorize their notes before engaging in hypothetical cooperative dilemmas. Then, in the critical phase of the study, participants were presented with the target profiles but without any information on cooperative traits, and they were then asked to predict their partner's likelihood of cooperating using the same scenarios from Studies 1-2. I used the data from Studies 1-2 to determine the optimal cooperation predictions for each scenario. For example, participants from Studies 1-2 had rated "caregiving ability" as the most predictive of cooperation in a scenario that involved trusting someone to take care of one's children for a day, and so the optimal strategy for that scenario would be to predict the partner's level of cooperation based on their caregiving ability. Figure 8 illustrates each phase of Study 5's procedure.

Our central measure in this task was participants' "prediction error," the absolute value of the difference between participants' 1-100 estimates of partner cooperativeness and the partner's real cooperativeness. For example, if participants gave Alice a trustworthiness score of 78 in the caregiving scenario—which was previously deemed to be most relevant to childcare ability—participants prediction error would be the absolute value of the difference between 78 and Alice's level of caregiving ability (which participants viewed in the first part of the study).

Social network size manipulation. In the "small network" condition, participants only needed to memorize the characteristics of three profiles, whereas they needed to memorize information about twelve targets in the "large network" condition. Participants completed the same number of cooperation dilemmas in both conditions. However, in the large network

condition, participants viewed twelve different profiles, whereas in the small network condition, participants viewed the same three profiles four times each to complete the twelve dilemmas.

Moral strategy manipulation and measurement. In addition to network size, I also manipulated how participants made moral decisions in Study 5. In the "global morality" condition, participants were told "people in this community have a fundamental underlying moral character, which determines how they behave in a range of situations in everyday life. People's individual traits give you hints about what their moral character might be." In the "differentiated morality" condition, participants were told "people in this community have no single underlying moral character, which means that they may be more cooperative in some situations than others." This manipulation was intended to manipulate whether participants used a global concept of morality (a single moral trait that informed their decisions in all situations) or tried to store information about many different situation-specific moral concepts.

I also measured how people made moral decisions with an item that followed the cooperative dilemmas. After participants completed these dilemmas, I asked them about the strategy that they used to remember and take notes on cooperative information, with options including: (a) "I tried to get a general impression of the person's moral character by averaging across their traits," (b) "I tried to remember specific traits so I could get a sense of people's behavior in different contexts," and (c) "I tried a different strategy (please specify)." I used this to measure participants' strategy for estimating morality since averaging across cooperative information (strategy a) indicated using a global moral concept whereas using specific cooperative traits in different contexts (strategy b) indicated using more specific moral concepts. This analysis excluded 33 participants who indicated option c, meaning that they used a different strategy.

Scoring of Prediction Error.

Analytic plan. The primary hypothesis in Study 5 was that the morality strategy manipulation and the social network size manipulation would interact such that the global morality strategy would be relatively better at predicting cooperation than the differentiated morality strategy in the large network condition. I therefore predicted a negative effect of global morality strategy on prediction error in the large network condition, since higher prediction error values signified worse predictions. I predicted a null or positive effect of global morality strategy on prediction error in the smell network condition. I also tested two hypotheses related to the measured morality variable, predicting that (a) people would adopt the global morality strategy more frequently in the large network condition than the small network condition, and (b) that using the global morality strategy would be relatively better at predicting cooperation than the differentiated morality strategy in the large network condition. These were logistical regressions since moral strategy was a binary variable. Finally, I conducted exploratory analyses of the text that participants entered about each of their hypothetical community members during the first phase of the study, since this text indicated the amount of information that participants were encoding about each of their potential interaction partners.

Results

Moral strategy manipulation results. In the primary model with network condition with moral strategy condition, I did not find results for my hypothesized interaction, $b = 1.01$, $SE = .80$, $t(1130) = 1.26$, $p = .21$. I also did not find significant main effects of either manipulation.

Moral strategy measurement results. In my secondary model with network condition and measured moral strategy, I found support for my hypothesized interaction, $b = -2.71$, $SE = .87$, $t(1097.93) = -3.13$, $p = .002$. Participants in the small network condition performed

significantly worse when they used a global morality strategy compared to when they used situation-specific cooperative traits, $b = 2.05$, $SE = .63$, $t(1097.99) = -3.26$, $p = .001$. However, they performed non-significantly better in the large network condition, $b = -.67$, $SE = .60$, $t(1097.85) = -1.12$, $p = .26$. This interaction is illustrated in Figure 9.

I also found that participants were more likely to adopt the global morality strategy in the large network condition, $OR = 1.28$, $SE = .03$, $t(1101) = 6.05$, $p < .001$, but that they were not significantly more likely to use the global morality strategy when they were in the global morality condition, $OR = 1.01$, $SE = .03$, $t(1101) = .25$, $p = .80$. In other words, participants may have intuited that a global morality strategy would be more functional in large social networks compared to smaller social networks, and these intuitions may have bypassed the effect of my manipulation on strategy. This pattern of results would also explain why I didn't find a significant interaction between morality condition and network condition in my primary analysis.

Text entry. My final analysis examined how much text participants wrote down about each of their hypothetical interaction partners. A long text entry would indicate that someone wrote down extensive notes about their cooperation partners, but it should also be more difficult to remember long text entries. The effect of text entry on performance in cooperation dilemmas should therefore depend on network size, showing diminishing marginal returns in larger vs. smaller social networks.

I found an interaction between network size and text entry length that was consistent with this prediction, $b = 3.90$, $SE = 1.52$, $t(1131) = 2.57$, $p = .01$. Text entry length was associated with better performance in the small network condition, $b = -4.10$, $SE = 1.12$, $t(1131) = -3.65$, $p < .001$, but not in the large network condition, $b = -.19$, $SE = 1.02$, $t(1131) = -.19$, $p = .85$. This

effect, displayed in Figure 9, offers evidence that encoding rich information about cooperation partners shows diminishing marginal returns as social network size grows.

Discussion

Study 5 found some evidence that people using a global concept of morality predict cooperation better in large social networks than in small social networks, compared to people using more differentiated and situation-specific concepts of morality. My manipulation of moral strategy showed no effect on performance. But measuring self-reported moral strategies showed that participants using a global moral concept performed worse in the small network condition but similarly well in the large network condition compared to people who used more differentiated moral concept. I also found that keeping detailed notes about cooperation partners was associated with better performance in the small network condition but not the large network condition, which shows that detailed information shows diminishing marginal returns as network size grows.

Despite these findings, Study 5 did not show that a global concept of morality could outperform using situation-specific concepts in large social networks—only that it performed similarly in a large social network. Study 6 tested the observation-inference hypothesis from moral concept theory to determine whether global concepts of morality could outperform more differentiated concepts of morality when people were interacting with relatively anonymous interaction partners.

STUDY 6: EXPERIMENTALLY TESTING THE OBSERVATION-INFERENCE HYPOTHESIS

Study 6, like Study 5, used an experimental design to test whether a global concept of morality can optimize decisions in cooperative dilemmas. However, Study 6 tested a different mechanism whereby global morality can optimize cooperation in large social networks: the

observation-inference hypothesis. The observation-inference hypothesis is that using a global concept of morality can make up for the lack of interaction that people have about their social partners in large social networks. Larger social networks are characterized by relative anonymity, meaning that people will have sparse information about their interaction partners. Using a global concept of morality—which applies across many situations—means that people can infer a partner's potential cooperation across a wide range of situations using sparse data about how they have cooperated in the past. These inferences will not be as accurate as more complete situation-specific information about cooperation, but it should be more accurate than guessing cooperation with uninformed priors.

To test the observation-inference hypothesis, I used a similar paradigm to Study 5. Participants engaged in cooperation dilemmas within a hypothetical community. Rather than manipulating social network size, I manipulated whether participants had access to complete or incomplete information about their partners' cooperation-relevant traits. I hypothesized that using a global concept of morality would perform better than using more differentiated concepts of morality when participants had sparse information about morality but would perform similarly or worse when participants had complete information about a partner's level of cooperation across situations.

Method

Sample. I advertised the study for 1,000 participants on Amazon Mechanical Turk, but only 795 participants signed up for the study, and only 692 participants (326 men, 362 women, 4 non-binary; $M_{age} = 38.44$, $SD_{age} = 12.24$) completed all the measures and passed the manipulation check, which was the same check as in Studies 1,2, and 5.

Measures and procedure. Study 6 followed a similar procedure to Study 5. Participants were given the same instructions about participating in a hypothetical community where they would need to engage in cooperative dilemmas, and the dilemmas were identical to those in Study 5. However, Study 6 did not have a memorization period, and I did not vary social network size. Instead, participants immediately engaged in cooperative dilemmas after reading the instructions.

Information completeness manipulation. To manipulate the information that participants had about their interaction partners, I randomly assigned half of the participants in Study 6 to see profiles that lacked key cooperative traits. In these incomplete profiles, the most relevant trait to make cooperation decisions was always missing (see Figure 10). For instance, in the scenario where participants needed to trust a partner to look after their children, information about the partner's caregiving ability would be missing. This manipulation was designed to mimic the lack of information that people have about their interaction partners in large and anonymous societies.

Moral strategy manipulation and measurement. I manipulated and measured global vs. differentiated morality using the same manipulation as in Study 5. The manipulation did not affect self-reported strategy in Study 5, but this may have been because people naturally adapted their strategy when they were in large vs. small social networks. This was less of a concern in Study 6 since network size was constant.

Analytic plan. I used the same analytic plan as in Study 5. Since participants did not enter notes at any point in the study, I did not analyze text responses in Study 6.

Results

Moral strategy manipulation results. In my primary model with network condition and manipulated moral strategy, I found support for my hypothesized interaction, b = -2.62, SE = .89,

$t(688.14) = -2.95$, $p = .003$. Participants in the complete information condition performed non-significantly worse when they used a global morality strategy compared to when they used situation-specific cooperative traits, $b = .81$, $SE = .63$, $t(688.14) = 1.28$, $p = .20$. However, they performed significantly better in the incomplete information condition, $b = -.67$, $SE = .60$, $t(1097.85) = -1.12$, $p = .26$. This interaction is illustrated in Figure 10.

Moral strategy measurement results. In my secondary model with network condition and measured moral strategy, I found further support for my hypothesized interaction, $b = -3.32$, $SE = .92$, $t(658.01) = -3.60$, $p < .001$. Participants in the complete information condition performed significantly worse when they used a global morality strategy compared to when they used situation-specific cooperative traits, $b = 2.04$, $SE = .66$, $t(658.01) = 3.08$, $p = .002$. However, they performed significantly better in the incomplete information condition, $b = -1.28$, $SE = .64$, $t(658.01) = -1.99$, $p = .047$.

Discussion

Study 6 showed that using a global moral concept can help people predict cooperation when they interact with relatively anonymous partners. When participants had complete information about how their partners cooperated in different situations, using a global moral concept resulted in lower performance than using differentiated moral concepts. But when participants had incomplete information about their cooperation partners, using a global moral concept resulted in better cooperation predictions. These results support the observation-inference hypothesis that more general concepts of morality allow people to infer cooperation when they have sparse information about their cooperation partners. In sum, Studies 5-6 provide causal evidence that a global concept of morality can be functional for cooperative partner selection in large social networks.

STUDY 7: SOCIAL LEARNING AND MORALITY IN THE HADZA

How can more general concepts of morality spread throughout human groups? Studies 5-6 identified the conditions where a global concept of morality could be functional, but it did not provide a mechanism for how a global concept of morality emerges and spreads over time. Studies 7-8 addressed this mechanistic gap by testing whether modes of moral cognition could be socially learned at the micro (Study 7) and the macro (Study 8) level.

In Study 7, we focused on whether cultural exposure could predict variation in how people made moral judgments among the Hadza, a nomadic hunter-gatherer group living along the Central Rift Valley in northern Tanzania. The Hadza typically inhabit bands of about 30 children and adults. There are little more than 1,000 Hadza people today, but much fewer Hadza still maintain a hunter-gatherer lifestyle (Marlowe, 2010). Even among full-time foragers, there is increased access to outside society via the number of aid workers, missionaries, and ethnotourists visiting Hadza which continues to rise every year (Apicella, 2018; Apicella et al., 2014; Pollom et al., 2021). In a group of Hadza hunter-gatherers sampled in 2019, 40% reported living outside of Hadzaland at some point, 25% reported having held a job that pays money, and nearly 60% claimed to have heard of the former United States President, Barack Obama (Smith & Apicella, 2020). The growing exposure to Tanzanian and Western culture has led to changes to Hadza diet (Crittenden et al., 2017) and foraging strategies (Pollom et al., 2021), but it could also have plausibly led to differences in moral cognition via exposure to Western concepts of morality.

Moral concept theory suggests that more general concepts of morality emerge and spread as societies grow larger and more complex, and that people learn to use more general concepts of morality by copying prestigious individuals. The Hadza represent a good population to test this

prediction because Hadza social networks are beginning to change as more and more Hadza live outside of Hadzaland, work in jobs, and learn about life outside of Tanzania. Since Hadza with greater cultural exposure have larger social networks, I predicted that they would adopt more general concepts of morality. Moreover, since Western culture is associated with prestige in southeastern Africa, I also predicted that exposure to Western culture might predict using more generalized moral cognition above and beyond social network changes. In sum, accessing data on moral judgments among the Hadza provided an ideal venue for testing changes in moral cognition in a rapidly changing society.

Method

Sample. I re-analyzed data from Smith & Apicella (2020) for this study. Smith and Apicella (2020) collected data from 85 Hadza individuals in 2019 through one-on-one interviews. This sample included 41 women and 45 men (M_{age} = 36.33, SD_{age} = 13.87) from 12 different camps.

Moral judgment measure. Participants ranked eight individuals in their camp on morally relevant attributes, such as generosity ("Who is the most generous?"), honesty ("Who is the most honest?"), effort towards foraging ("Who works the hardest to get food"), partner choice preference ("who would you most like to live with if you were to move camp tomorrow?"), and having a good heart ("Who has a good heart?"). These rankings were done using cards in one-on-one interviews with research assistants, and the full details of the procedure are available in Smith and Apicella (2020).

I used these rankings to create two measures. The first measure quantified contextual variability in participants' moral rankings. A more situation-specific conception of morality would mean that these ratings would be independent, but a more situation-general conception of

morality would mean that these ratings would be related, such that someone who is highly generous would also be perceived as highly honest and as having a good heart. I took the standard deviation across rankings within individuals with the assumption that a higher standard deviation would indicate greater contextual sensitivity, whereas a lower standard deviation would indicate more situation-general moral judgment.

The second measure quantified whether participants used a global concept of morality in their moral judgments. The term "having a good heart" was selected by Smith and Apicella as a rough Hadza equivalence of the English word "morality," meaning that I could operationalize global conceptions of morality in terms of how strongly individuals' predicted their mean ranking of all other attributes. If participants rated having a good heart as strongly correlated with other moral attributes, this would indicate greater belief in a global concept of morality.

Cultural exposure. I measured cultural exposure using the same measure as in Smith and Apicella (2020). This measure contained 10 different indicators: Years of school (log transformed), whether participants could count to 10 in Swahili, whether participants had held a job outside of Hadzaland, whether participants could identify the president of Tanzania, whether participants could identify Barrack Obama, whether participants could identify Nelson Mandela, whether participants could identify Mahatma Gandhi, whether participants had lived outside of Hadzaland, and whether participants had lived in the neighboring city to Hadzaland.

An exploratory factor analysis identified two factors within this cultural exposure measure. One factor (Eigenvalue = 1.41) contained the three items about knowledge of foreign figures (Gandhi, Obama, and Mandela). The other factor (Eigenvalue = 3.79) contained the remaining items. The first factor appeared to be tapping foreign knowledge, whereas the second factor appeared to be tapping local exposure via how much time participants had spent outside of

Hadzaland working and in school. These factors were highly correlated ($r = 40$, $p < .001$), and Smith and Apicella (2020) collapsed all items onto a single factor. I used the single factor measure as well as replicating analyses with the separate factors.

Analytic plan. I tested my predictions using multi-level regressions nested within judge (the person doing the ratings) and subject (the person being rated). To measure contextual specificity in moral judgments, I regressed the standard deviation of the judge's rankings on judge's cultural exposure, controlling for the judge's age and gender, and controlling for the subject's age, gender, and whether the subject was a spouse of the judge. To measure global morality, I regressed the mean of judge's rankings on the interaction of cultural exposure and the good heart ranking controlling for the same covariates. I then replicated these analyses while breaking the cultural exposure score into the two subfactors that I had identified in an exploratory factor analysis and entering these factors into a multiple regression.

Results

Cultural Exposure and Contextual-Specificity in Moral Judgment. A multiple regression found a negative robust association between cultural exposure and the standard deviation of all moral rankings, $b = -.51$, $SE = .15$, $t(81.39) = -3.42$, $p < .001$, suggesting that Hadza with more cultural exposure ranked their campmates as similarly cooperative or uncooperative across all attributes, rather than using contextual specificity. Figure 13 shows the zero-order association between ranking standard deviation and cultural exposure. When I re-ran this model with separate cultural exposure factors, the foreign exposure factor was significant, $b = -.27$, $SE = .13$, $t(78.69) = -2.08$, $p = .04$, but the local exposure factor was not, $b = -.02$, $SE = .02$, $t(80.14) = -.95$, $p = .34$, suggesting that foreign exposure could predict using more generalized moral judgment above and beyond exposure to local education and workplace

systems. Table 8 displays each of these regressions, along with the effects of the control variables.

Cultural Exposure and Use of a Global Moral Concept. I also found the predicted interaction between cultural exposure and good heart rankings, b = -.02, SE = .02, t(80.14) = -.95, p = .34. Good heart was more strongly predictive of other moral attributes for participants with higher, b = -.02, SE = .02, t(80.14) = -.95, p = .34, compared to lower, b = -.02, SE = .02, t(80.14) = -.95, p = .34, cultural exposure (see Table 9 and Figure 11). The separate cultural exposure scores model was singular, so I did not interpret the results.

Discussion

Hadza hunter-gathers with higher levels of cultural exposure showed less situation-specific moral judgment and more evidence of a global moral concept when they ranked the morality of their campmates. These results are significant because they show preliminary evidence that styles of moral cognition can be transmitted and learned across cultures. Of course, these data were correlational so they cannot prove directionality. It may have been that Hadza with more general moral cognition sought out exposure outside their community. However, it is promising that these results were robust to controlling for the age and gender of judges and the subjects that they rated. It is also interesting that foreign exposure predicted generalized moral cognition above and beyond exposure to the surrounding local culture.

STUDY 8: WORD BORROWING SHOWS THE LEXICAL TRANSMISSION OF MORAL CONCEPTS

A significant proportion of words from every language are borrowed. The English word "boulevard" and "adroit" are both borrowed from French, whereas the French words "cool" and "airbag" are borrowed from English. Unlike cognates (words that are derived from a common ancestral form like the English "one" and the French "une"), borrowed words are not handed

down by a parent language, and words are frequently borrowed because they are useful for speakers of a language (Grzega, 2003). For example, French borrowed the concept of *airbag* because it helped communicate an important technological innovation.

In Study 8, I used historical word borrowing patterns to see if concepts resembling the concept of *morality* have been historically borrowed at higher rates in large societies compared to small societies. This finding would be significant in two ways. First, it would provide macro-level evidence that information about morality can be socially learned, which would complement Study 7's micro-level evidence. Second, it would provide evidence that larger and more complex societies have had an especially high demand for concepts resembling *morality* throughout human history. This would complement Study 7's finding which suggests that large, industrialized societies are transmitting global concepts of morality through current-day cultural exposure.

Method

Sample. I obtained data on word borrowing from the World Loanword Database (WOLD) database (https://wold.clld.org/meaning/16-11#2/24.3/-4.8), which contains borrowing data from approximately 44 large-scale and small-scale languages. WOLD groups concepts into categories (e.g., "emotions and values," "the body") and then rates each concept within these categories from 1 ("Clearly Borrowed") to 5 ("No evidence for borrowing") for each language. For some words that are either probably borrowed or clearly borrowed, WOLD also provides the word's source language and original word form. For example, the Kanuri word "ro" (meaning the spirit or soul) is labeled as "clearly borrowed" from the Arabic word "ruuh." All borrowing rates are compiled by linguistics experts.

Coding procedure. Study 8 focused on the "emotions and values" category of concepts, which contains 48 independent concepts. Some of these concepts are clearly related to *morality* (e.g., *morally wrong*) and others are not related to *morality* (e.g., *good luck*). Two research assistants coded the word-forms for each of these concepts and its borrowing score in WOLD.

One advantage of sampling from a variety of words with various relationships to morality is that it naturally controlled for borrowing rates of non-moral words. For example, one possibility is that larger languages are simply more likely to borrow any kind of word related to emotions and values, but sampling from the full emotions and values list ensures that our effect focuses on likelihood of borrowing words about morality *relative* to other kinds of words about emotions and values.

Population size. I retrieved data on the size of each language from Ethnologue, a large linguistic database that indexes data about world languages. Language size was highly skewed, and I log-transformed the variable prior to analysis. Ethnologue also provides categories to indicate the status of the language, including "endangered," "extinct," and "national." I used these categories to create a dummy-coded variable separating current-day languages (e.g., "national," "provincial") from languages that are disappearing or have disappeared (e.g. "endangered," "extinct").

Norming concepts. I normed each concept in this study on its relevance to morality using a sample of 219 English-speaking participants. This strategy has been employed in past research (Jackson et al., 2019), and it is appropriate because the "emotion and values" concepts have English-language meanings. All participants rated each of the 48 concepts on how relevant it was to the concept of morality. Before rating the concepts, participants read the instructions: "Some people define 'morality' as the code that separates fundamentally 'moral or good' acts or people

from ‘fundamentally immoral or bad’ acts or people. We would like you to rate a series of concepts on how related there are to the concept of morality. To give you an example, the concept of sandwich would be highly unrelated to the concept of ‘morality’ whereas the concept of ‘evil’ would be highly related to the concept of ‘morality.’” Participants then rated concepts using a scale from 1 (“highly unrelated to morality”) to 7 (“highly related to morality”).

Analytic plan. I used a regression to test whether larger languages are more likely to borrow concepts related to morality than smaller languages. One concern with this analysis would be confounding a general likelihood to borrow concepts with the specific tendency to borrow concepts about morality. I addressed this confound by estimating the relationship between borrowing status and moral relevance for each language and then regressing these estimates on language size. In a subsequent model, I added control variables which captured whether languages were national, and whether they were contemporary (vs. extinct) to ensure that language size was not confounded with whether a language was contemporary or historical. In this second model, I also estimated a multi-level model with intercepts randomly varying across language families to ensure that a significant relationship was not confounded with interdependence of datapoints. Figure 12 displays the geographic coordinates of each language in Study 8’s analysis, with node size and color illustrating the language size and the link between borrowing status and morality relevance.

Results

As predicted, larger languages had a greater likelihood of borrowing words for concepts related to morality, $b = .04$, $SE = .02$, $t(36) = 2.33$, $p = .03$. This relationship replicated controlling for nestedness across language families, and contemporary language status, $b = .05$,

SE = .02, $t(33.97) = 2.07$, $p = .046$. Figure 13 displays the relationship between language size and the share of borrowed concepts related to morality.

STUDY 9: MODELING MORAL CONCEPT THEORY

Experiments, surveys, historical language analysis, and ethnographic field data support different predictions in moral concept theory. My final study synthesized this evidence and formalized moral concept theory's predictions and assumptions with an agent-based model using cultural evolutionary dynamics. To summarize this model, I present a plain text description of the model dynamics, followed by a more symbolic summary of the model parameters.

Plain Text Description of Model

Imagine that you are living in a small community, which grows every year. You frequently must decide whether you trust people in your community to cooperate with you in specific situations, and people's likelihood of cooperation correlates loosely from situation to situation. As you interact with people more, you learn how they behave in different situations, but you can only have so many interactions in a year, and you can only remember so much information about your cooperative partners. Given these constraints, it is more difficult to predict how people will behave as more people enter your social network. What is the best way of predicting cooperation in this growing community?

Symbolic Description of Model

In more symbolic terms, some number of agents *n* are situated in a small-world network where they are more likely to interact with neighbors than with cross-cutting ties. Agents have a starting level of resources ϑ that changes over time. In each round *i* of the model, connected agents *a* in the model have an opportunity to interact with each other by playing a trust game, where an actor agent a_1 must choose to send some amount of their resources to a partner agent

a_2. This amount is bounded at 0 and 1 and is denoted as prediction *P.* The resources triple after they transfer, and a_2 can then send some proportion of the tripled amount back to a_1. A trust game mimics partner choice dynamics because $a_1$can choose to send nothing to a_2 if they anticipate defection. In this model, however, agents do not have global cooperative values. They have different levels of cooperation across *s* situations, with cooperation coefficients simulated so that the situation-total correlation coefficient is .25 (the coefficient that we estimated in our survey of economic games). Each agent uses a $n \times s$ prediction matrix to estimate their partners likely level of cooperation in each round of the trust game, and they update this prediction matrix after each interaction by observing the behavior of their partner. With enough interactions, each agent can therefore develop a complete set of predictions that will optimize their partner choice decisions.

The model's critical manipulation is that agents use different partner choice strategies. If a_1 is a local optimizer, they will try to optimize *P* using knowledge about a_2 cooperation in situation s_k. If they have no prediction data because they have not interacted with a_2 in situation s_k, they will generate a random prediction. In contrast, if the actor agent is a global optimizer, they will try to optimize their cooperation prediction by equally weighting the previous knowledge that they have about their partner in situations *s* by averaging all coefficients into $\bar{s}$. Global optimizers will nearly always make informed predictions, but their predictions will sometimes be inaccurate because they are biased by cooperation values from less relevant situations. The difference between local and global optimizers resembles the difference between people who use a global moral concept that applies across situations vs. more granular situation-specific moral concepts. A global moral concept will encourage people to make inferences about someone's likelihood of cooperation in nearly every situation, but these inferences will often be

less accurate than if they had an accurate situation-specific prediction. Equations 1-2 describe the prediction dynamics in the trust game for agents with local prediction (Equation 1) and agents with global prediction (Equation 2).

$$\sum_{a=1}^{n} P_{a,s} = a_{2,s_k} \quad (1)$$

$$\sum_{a=1}^{n} P_{a,s} = a_{2,\bar{s}} \quad (2)$$

In addition to the cooperation phase, each round also has a learning phase, an expansion phase, and a forgetting phase. In the learning phase, connected agents will have the opportunity to change their prediction strategy (local vs. global prediction) to match their neighbor's strategy. Agents learn via cultural evolutionary dynamics with payoff-biased transmission, meaning that they are more likely to learn from agents who have more resources than they do. During the expansion phase, φ agents are added to the network and assigned connections with other agents. A dispersion parameter δ regulates the network characteristics, such that lower values will mean that new agents will be more densely connected to their neighbors and higher values will translate to more diverse connections across the network. New agents are randomly assigned prediction strategies, and every agent is assigned the same initial resource value ϑ, which means that new agents will be more likely to socially learn than agents who have accrued more resources over time. During the forgetting phase, agents with more than γ values in their prediction matrix will forget some proportion ψ of their prediction values, which mimics the limit to social information that humans can store in memory.

Results

Running this model for 200 iterations reveals clear dynamics. For example, as agents are continually added to the model, each agent's prediction matrix grows sparser. Each agent's mean number of interactions increases as a function of new ties that are formed with incoming agents, but agents still do not observe their partners as frequently because their rising number of interactions is outpaced by their exponentially increasing number of potential social ties. Agents' social memory limit also inhibits the number of situation-specific traits they can store for their partners. These observation and memory barriers prevent agents from retaining dense prediction matrices for each of their interaction partners over time. To illustrate these dynamics, Figure 14 displays the rising sparsity in agents' prediction matrices over time on the left, and the mean number of interactions along with the number of interactions per network edge on the left.

As agents' prediction matrices grow sparser, global prediction becomes a more valuable strategy. In small networks of highly familiar agents, global prediction introduces error because cooperation in less relevant situations biases agents' predictions. But when agents are not able to predict situation-specific cooperation because of sparsity in their prediction matrices, it is functional to infer cooperation using information from other situations because cooperation correlates moderately across situations. Figure 15's top-left panel illustrates that payoff-biased learning results in a rising number of conversions to a global prediction strategy and a declining number of conversions to a local prediction strategy over time, and the top right panel illustrates the rising proportion of agents who pursue global prediction over time. The bottom panel of Figure 15 displays the network of agents with agents shaded based on their prediction strategy. Over time, the network becomes saturated with agents who pursue a global prediction strategy. Table 10 lists key results at the conclusion of 10 different runs of the model: (a) mean sparsity of

prediction matrices, (b) ratio of conversions to global vs. local prediction strategies, and (c) proportion of agents using the global prediction strategy.

These results are robust to variability in the model's parameters, including variation in the dispersion of the network (δ), how quickly the network grows (φ), and agents' memory capacity (γ). In our primary models, we set γ at 150 given the significance of the number in evolutionary research on social memory (R. I. Dunbar, 2003, 2014), but results were highly similar when doubling or even tripling agents' memory capacity. Table 11 presents prediction matrix sparsity and the proportion of global prediction agents at different values of $\delta, \varphi,$ and γ.

Discussion

This model represents a stylized version of moral concept theory and does not include the intricacies of how language can communicate cognitive style or how social norms and institutions such as religion can shape moral convictions. However, with very basic assumptions about cooperative dilemmas and social network structure, this model shows how local and global partner choice strategies can change through social learning, and how generalized concepts of morality can rapidly spread in large and complex social networks. Perhaps most importantly, this model provides a mathematical and visual framework for understanding some of the conceptual claims within moral concept theory.

GENERAL DISCUSSION

Many people may claim that "morality" is one of the most important words in their vocabulary. Humans use morality to make life-changing judgments about partners, co-workers, and strangers, and go to great lengths to preserve their own morality (Vonasch et al., 2018). However, I have suggested here that the idea of morality is neither universal nor is it historically enduring. Morality may have emerged as a socially learned concept for predicting people's

cooperation in large complex societies where people have many ties and largely anonymous interaction partners. In these large societies, it is impossible to remember and observe information about people's cooperation across many different contexts, and so more general concepts of cooperation can help people remember information about a greater number of interaction partners and make inferences about how these partners may behave in novel contexts. A global concept of morality represents the extreme end of this process of generalization—it can predict how someone will behave in any cooperative situation and it has very little memory cost—although these benefits come at the expense of accuracy.

Three pilot studies and 9 empirical studies support this theory of morality. Pilot Studies A-C provided crucial evidence that cooperation varies meaningfully across situations, which is an important assumption of moral concept theory. Studies 1-2 then showed that people view the word "morality" to relevant to a wider range of situations than six other cooperation-relevant traits. Studies 3-4 show that people in large social networks endorse and use the global concept of morality more than people in smaller social networks. Studies 5-6 demonstrate the functionality of a global concept of morality in large social networks, showing support for the social memory hypothesis and observation-inference hypothesis. Studies 7-8 found evidence of social learning in moral judgment, both at the micro-level among individuals in a remote hunter-gatherer society (Study 7) and among languages throughout human history (Study 8). Finally, an agent-based model synthesized this evidence into a dynamic simulation where agents learned a global concept of morality through payoff-biased transmission as they interacted in an ever-growing dynamic social network.

Situating Moral Concept Theory with Other Theories of Morality

Moral concept theory builds on several other theories of the evolution and structure of morality. Perhaps most notably, it provides an alternative to social intuitionist models that suggest humans evolved a fixed number of moral intuitions such as loyalty and reciprocity (Curry et al., 2019; Haidt & Joseph, 2004). Moral concept theory accommodates this plurality of moral judgments, but it suggests that there is no single number of moral intuitions that have been hardwired into human moral psychology. Rather than neurobiologically based modules, traits such as "loyalty" and "fairness" are likely semantic concepts that organize predictions about people's cooperation in different situations. Similarly, humans may view morality along twenty, seven, five, three, or one dimension based on their social context and their relationship partner. We may view the morality of a familiar partner like a family member along a vast number of different dimensions while viewing a faceless political opponent along a single dimension of immorality. Social intuitionist models have been called a "first draft" of the moral mind (Haidt & Joseph, 2007), but moral concept theory suggests that the first draft of the moral mind involves the natural human ability to use language and memory to categorize social behavior.

Moral concept theory also joins a recent push towards "person-centered morality" in moral psychology. Person-centered morality argues that the major function of moral cognition is to understand and predict people's behavior, rather than formulate attitudes about abstract moral principles (Pizarro & Tannenbaum, 2012; Uhlmann et al., 2015). Another more recent addition to person-centered morality has focused on how moral judgment can be different for people in close relationships compared to acquaintances (Earp et al., 2020), which aligns with moral concept theory's claim that the structure of morality should be different for close and distant social ties. Other person-centered perspectives have emphasized the potential for moral judgments to interact with social cognitive categories like race, gender, and political ideology

(Hester & Gray, 2020; Schein, 2020). To this end, a fruitful area of future research for moral concept theory could identify variation in how humans use moral concepts when judging racial, cultural, and political in-groups and out-groups.

Another theory of dyadic morality has argued that humans make moral judgments through a universal dyadic template where an agent harms a patient, and that this template typically involves negative affect, perceived harm, and the violation of social norms (Gray et al., 2012; Schein & Gray, 2018). In this sense, dyadic morality offers a process-based account of moral cognition that complements moral concept theory's structure-based account of moral cognition. Whereas moral concept theory explains how the conceptual space of morality can change and evolve over time, dyadic morality explains the phenomenology of moral judgment. By emphasizing the common role of a dyadic template and perceived harm in moral judgment, dyadic morality also suggests that many different perceived moral violations ranging from dishonesty to disloyalty to discrimination may invoke the same affective and cognitive processes, even though they are conceptualized differently in language.

Finally, moral concept theory provides an alternative to existing evolutionary accounts of morality. Many of these accounts rely invoke social value orientation—the orientation someone has towards allocating resources to the self and to the other (McClintock & Allison, 1989)—and the closely related idea of the "welfare tradeoff ratio," the cost that one agent will impose on another for some benefit *t* (Delton & Robertson, 2016). Claims in evolutionary psychology suggest that humans may have evolved modular devices for calculating welfare tradeoff ratios using characteristics of the situation (public vs. private context, amount at stake) and cooperation partner (kinship, formidability) (Sell et al., 2009), and that these devices may even be linked with genetic adaptations such as the monoamine oxidase A (MAOA) gene (McCullough et al., 2011).

Together, these systems may have given humans an efficient method of calculating the benefits of cooperative behavior, while gauging the morality of others.

Individual differences in people's willingness to help and harm others are undeniable, and social value orientation is a useful measurement tool (Murphy et al., 2011). However, theories of asocial evolutionary adaptations for detecting morality are both too simple and too complex. A hardwired mechanism for detecting and predicting cooperation is too simple to account for the nuanced nature of cooperation based on context, social partner, and cultural norms, and it is simultaneously too complex to envision an independent genetically evolved module for predicting and categorizing cooperation when humans can flexibly categorize and communicate cooperation using domain-general processes like language, memory, and social learning.

Another evolutionary model named Dynamic Coordination Theory claims that morality may have evolved to provide universal laws around which people can coordinate in social dilemmas (DeScioli & Kurzban, 2013). This perspective is original and creative, and it may be true that morality helps people pick sides in debates. But in other ways, this model is implausible. For example, dynamic coordination theory assumes that certain moral preferences are biologically evolved to coordinate behavior and resolve conflict, yet many conflicts focus on moral behavior, suggesting that morality can often be a source of division rather than resolution. Moreover, morality can only function as a tool to resolve conflicts if people agree on moral principles, but evidence shows widespread disagreement in judgments of morality (Skitka & Morgan, 2014).

Situating moral concept theory among other theories of moral cognition highlights two ways that moral concept theory may advance previous research on human morality. First, moral

concept theory shows how humans can develop moral pluralism without needing distinct biologically based moral intuitions. Second, moral concept theory shows how human morality may have evolved without a specific universal structure that coordinates moral judgment. Beyond these insights for evolutionary and moral psychology, moral concept theory also has implications for historical changes in how humans have viewed gods, and why moralization and political polarization is rising.

Is Conceptual Broadening Specific to Morality?

I propose that the moral cognition becomes conceptually broader as societies grow larger and more complex. However, an alternative possibility is that all social cognitive information becomes broader as societies grow, such that people also develop broader concepts of personality, emotion, etc. Conceptually broader social cognitive categories could, in principle, help people remember information about a larger group of potential interaction partners.

However, there are theoretical and empirical reasons to believe that morality is more sensitive to population size than these other social cognitive categories. Theoretically, a broader view of morality is unique for predicting cooperation since judgments of moral character are more directly linked to trustworthiness than judgments of personality or emotion (Curry et al., 2019). In this way, a broader concept of morality will not only help people remember information about interaction partners in larger groups; it will also help them use this knowledge to exploit higher payoffs than broader concepts of emotion or personality. Some empirical evidence also suggests that morality is uniquely sensitive to population size. For example, studies of culture and personality have found that conceptions of personality are *higher* dimensional in larger complex societies than small-scale societies (Smaldino et al., 2019). And in Study 8 of this paper, I found that morality words were borrowed by larger societies at a greater

rate than words about other concepts related to emotions and values. Taken together, this evidence suggests that morality may become broader at a more rapid rate than other social cognitive information as groups grow larger, although this is an interesting question for research to explore further.

Broader Implications of Moral Concept Theory

Implications for the Rise of Moralization. While moral concept theory primarily focuses on historical changes to moral cognition, it also relates to more recent trends in morality. Recent research has identified "concept creep": the growing moralizations of previously innocuous behavior (Haslam, 2016). Over the last several decades, a growing number of behaviors have been grandfathered into the moral domain and are now viewed as diagnostic of moral good or evil. Concept creep has coincided with exponential increases to the size of social networks via population size and urbanization (Li, 2020). The rise of the internet means that social networks can now also include ties outside of a person's local environment.

Concept creep may have little to do with social network expansion, but there are good reasons to believe that these trends are linked. Here I have shown that social network size correlates with a global moral concept, and this may occur because global morality allows inferences from cooperative behavior in one situation to make predictions in practically any other cooperative dilemma. People may expand the global moral concept for very similar reasons. Expanding the global moral concept to include events that are ambiguously related to cooperation means that one can make cooperative inferences about a novel social partner without having clear information about this person's previous cooperative behavior. This mechanism is closely related to the notion of "prevalence-induced concept change," wherein the diminishing prevalence of overtly harmful behavior (e.g., murder) encourages people to expand the scope of

immorality to include more borderline behavior (Levari et al., 2018). Prevalence-induced concept change makes it plausible that morality will only grow in its scope and importance, since population growth and urbanization show no signs of slowing.

Implications for the Evolution of Religion. One of the more enduring puzzles in the study of religion is the rise of moralizing high gods such as the Christian and Jewish God. Historical analyses suggest that high gods with universal moral codes likely emerged in the last 12,000 years, coinciding with the Neolithic Revolution when many human groups transitioned to more sedentary agriculture-based communities rather than nomadic hunter-gatherer groups (Watts et al., 2015; Whitehouse et al., 2019). One reason for this co-incidence may be because moralizing high gods helped regulate cooperation in these communities, encouraging people to cooperate by fostering the expectation that defection would be met with divine punishment (Biopolitics, 2016; Norenzayan & Shariff, 2008). This mechanism has been explored in a large literature on the rise of moralizing religion.

Another plausible mechanism, however, is that people developed stronger beliefs in morality as societies grew larger, and they projected these beliefs onto the content of gods minds. Some evidence supports this projection account. For example, people frequently project their own demographic and personality traits onto their perceptions of gods (Epley et al., 2009; Jackson et al., 2018; Purzycki, 2013), and project punitive characteristics onto perceptions of gods when they seek to punish norm violators (J. C. Jackson et al., 2020). This evidence makes it plausible that, once communities adopted a belief in fundamentally moral acts and people, their gods became arbiters of these moral beliefs, and the instruments of punishment for those who were deemed immoral. This bottom-up explanation could complement existing top-down

explanations that emphasize religious moralization as a tool for reinforcing large-scale cooperation.

Implications for Political Polarization. Political polarization is rising in many countries, with a particularly acute rise over the last several decades of American history (Finkel et al., 2020). In political psychology, affective polarization has emerged as a particularly acute problem because many Americans express more hostility and mistrust of opposing political parties in addition to disagreeing about policy stances (Iyengar et al., 2019; Rogowski & Sutherland, 2016; Webster & Abramowitz, 2017). To explain the recent rise in affective polarization, different models have pointed to increasing political segregation (Motyl et al., 2014), the rise of politically polarized media (Martin & Yurukoglu, 2017), and the rising rate of political interactions on the internet (Brady et al., 2020). Predictions from moral concept theory suggest that the rise of the internet and social media may be a particularly acute factor in the rise of political polarization.

The internet is unique both because it rapidly expands social network size, it involves a high density of interactions with strangers, and it rewards negatively valenced information (Brady et al., 2017). The first two of these conditions makes it likely that people will use a global concept of morality to interact with people on social media, whereas the third makes it likely that people will view social media interaction partners negatively. These three conditions create a perfect recipe for affective polarization, because adopting a global concept of morality will encourage people to make wide-ranging negative inferences about political opponents online.

These dynamics of political polarization show how the functionality of a global moral concept can backfire when global morality intersects with in-group out-group dynamics. It also shows how interventions that aim to reduce affective polarization may benefit from encouraging more granular moral concepts that allow people to appreciate the different ways that their

political can be cooperative. A different kind of intervention could change the way that people find partners on social media. Algorithms that encourage repeated interactions with the same people from opposing political orientations could build the familiarity that leads people to use more granular moral concepts rather than simply labeling their political opponents as globally immoral.

Conclusion

Moral cognition may be one of the most frequently essentialized constructs in the study of social behavior. Over the last several millennia, scholars have developed theories of morality which dictate fundamental ways that humans can be good or evil. In the early days of moral philosophy, normative theories of morality claimed to show how to be moral by invoking the will of the gods. More recently, intuitionist theories have claimed to show how people judge morality by invoking the brain and the body. All these theories are alike in their habit of claiming a universal and invariant structure to morality that has been invariant throughout human history.

I have suggested that moral cognition has no essential structure. People instead make moral judgments using socially learned concepts that communicate differences in cooperation across situations. Unlike past theories, I suggest that moral cognition may have changed significantly across history. Through payoff-biased transmission, humans may have culturally evolved increasingly general concepts of morality, and many people today may use a global concept of morality. In this view, moral cognition is not a universal system that we can locate in the brain or body. It is an emergent product of language, memory, and social learning, and the interaction of these processes endows morality with a constantly changing structure.

Many centuries ago, Plato and Aristotle debated whether virtue is a global and abstract form or whether it is a complex interaction of person and context. Here I suggest that both

answers are correct depending on whom you ask, where you ask, and when you ask. Understanding this variation in the structure of morality may help us chart morality's evolution throughout human history and could be the key for predicting the future of moral cognition in a world of rising globalization and digitization.

Table 1.

Cooperative Dilemmas in Study 2

Cooperative Dilemma	Instructions to Participants
Prisoner's Dilemma	*We have just paired you with another random participant in this study, and you must each decide whether to "cooperate" or "defect." If both you and your partner choose to "cooperate," you will both receive an additional bonus of 6 cents. If you and your partner both choose to "defect," you will both receive an additional bonus of 2 cents. If you choose to "defect" and your partner chooses to "cooperate," you will receive a bonus of 10 cents and they will receive a bonus of 0 cents. If you choose to "cooperate" and your partner chooses to "defect," you will receive a bonus of 0 cents and your partner will receive a bonus of 10 cents.*
Stag Hunt	*We have just paired you with another random participant in this study, and you must each decide whether to pursue a "large prize" or a "small prize." The large prize is 10 cents, but you will only get this money if both you and your partner choose to pursue the large prize. The small prize is 5 cents, but you are guaranteed to get this money, regardless of whether your partner chooses to pursue the small prize or large prize. For example, if you choose to pursue the large prize and your partner chooses to pursue the small prize, they will get 5 cents bonus while you will receive 0 cents bonus.*

Public Goods Game	*We have just paired you with other three random participants in this study, and will give each one of you 4 cents. If you choose, you can invest your money into a common pool. The money will be doubled and divided equally among all four participants, irrespective of your investments. You also keep the money you do not contribute. This setup means that you will personally get the most money if you keep your 4 cents and everyone else invests their 4 cents. However, the group will get the most money if everyone invests their 4 cents.*
Dictator Game	*We have just paired you with another random participant in this study. Now, you need to split 10 cents between the two of you. You can make the decision freely with no restriction at all, and the recipient has no influence over the outcome.*
Traveler's Dilemma	*We have just paired you with another random participant in this study. Now, each of you need to write down some amount of money from 2 to 10 cents. If you both write down the same number, we will give you each that amount of money. However, if you write down different amounts, the person who writes down the lower number will get that value + 1 additional cent, whereas the person who writes down the higher number will get nothing.*
Trust Game	*We have just paired you with another random participant in this study. In this dilemma, you have 10 cents, and you can choose to "send" any portion of your money to the other player. The money that you send will be tripled, and then your partner can choose to*

send any proportion of this money back to you. For example, if you sent the entire 10 cents, this would become 30 cents, and a totally egalitarian partner would send back 15 cents to you.

Table 2.

Decision-Making Tendencies in Pilot Study B Economic Games

Game	Metric	Value
Prisoner's Dilemma	Proportion Cooperated	73%
Stag Hunt	Proportion Cooperated	74%
Public Goods Game	Mean Contribution	3.20c/4.00c
Dictator Game	Mean Contribution	4.06c/10.00c
Traveler's Dilemma	Mean Contribution	6.01c/10.00c
Trust Game	Mean Contribution	6.91c/10.00c

Table 3.

Mean of Cooperative Behaviors in Pilot Study C

Behavior	**Mean Frequency**
Donated blood	1.25
Gave food or money to a homeless person	2.41
Returned money to a cashier after getting too much change	1.80
Allowed a stranger to go ahead of you in line	3.19
Done volunteer work for a charity	2.11
Given money to a charity	2.90
Offered your seat on a bus or in a public place to a stranger who was standing	1.93
Looked after a person's plants, mail, or pets while they were away	2.17
Carried a stranger's belongings	1.93
Given directions to a stranger	3.10
Let someone you didn't know borrow an item of some value	1.78

Table 4.

Mean and Standard Deviation of Each Trait in Study 1 Contrasted Against Morality

Predictor	*b*(SE)	*t*	*df*	*p*
Mean Model				
Caregiving Ability	-.17 (.02)	-7.87	1152	< .001
Loyalty	-.12 (.02)	-5.46	1152	< .001
Reciprocity	-.24 (.02)	-10.91	1152	< .001
Respect	-.20 (.02)	-9.46	1152	< .001
Fairness	-.24 (.02)	-11.09	1152	< .001
Generosity	-.20 (.02)	-9.19	1152	< .001
Standard Deviation Model				
Caregiving Ability	.09 (.01)	6.29	1152	< .001
Loyalty	.10 (.01)	6.69	1152	< .001
Reciprocity	.07 (.01)	5.24	1152	< .001
Respect	.03 (.01)	2.01	1152	.04
Fairness	.07 (.01)	4.93	1152	< .001
Generosity	.07 (.01)	5.11	1152	< .001

Note. All predictors are contrasted against morality.

Table 5.

Mean and Standard Deviation of Each Trait in Study 2 Contrasted Against Morality

Predictor	*b*(SE)	*t*	*df*	*p*
Mean Model				
Caregiving Ability	-.62 (.08)	-8.25	1182	< .001
Loyalty	-.26 (.08)	-3.48	1182	< .001
Reciprocity	-.51 (.08)	-6.88	1182	< .001
Respect	-.12 (.08)	-1.50	1182	.10
Fairness	-.64 (.08)	-8.62	1182	< .001
Generosity	-.51 (.08)	-6.76	1182	< .001
Standard Deviation Model				
Caregiving Ability	.44 (.05)	9.40	1182	< .001
Loyalty	.31 (.05)	6.54	1182	< .001
Reciprocity	.38 (.05)	8.04	1182	< .001
Respect	.03 (.05)	.66	1182	.51
Fairness	.51 (.05)	10.93	1182	< .001
Generosity	.48 (.05)	10.36	1182	< .001

Table 6.

Global Morality Belief and Reliance on Social Network Size Metrics from Study 3

Predictor	*b*(SE)	*t*	*df*	*p*
Global Morality Belief (1)				
Facebook Friends	1.51 (.55)	2.76	1993	.006
SES	-.01 (.25)	-.54	1993	.96
Age	.04 (.03)	1.12	1993	.26
Gender	-3.31 (1.15)	-2.89	1993	.004
Religiosity	2.11 (.34)	6.26	1993	< .001
College Degree	-4.39	-3.88	1993	< .001
Global Morality Belief (2)				
Everyday Interactions	.47 (1.02)	.45	1993	.65
SES	.002 (.25)	.006	1993	.99
Age	.01 (.03)	.38	1993	.71
Gender	-2.94 (1.14)	-2.57	1993	.01
Religiosity	2.19 (.34)	6.51	1993	< .001
College Degree	-4.33 (1.14)	-3.80	1993	< .001
Global Morality Reliance (1)				
Facebook Friends	1.63 (.61)	2.68	1993	.007
SES	.26 (.28)	.93	1993	.36
Age	-.23 (.04)	-6.48	1993	< .001
Gender	-3.23 (1.27)	-2.54	1993	.01
Religiosity	2.35 (.38)	6.30	1993	< .001

College Degree	-.46 (1.25)	-.37	1993	.71
Global Morality Reliance (2)				
Everyday Interactions	2.89 (1.14)	2.54	1993	.01
SES	.20 (.28)	.74	1993	.46
Age	-.24 (.04)	-6.85	1993	< .001
Gender	-2.74 (1.26)	-2.17	1993	.03
Religiosity	2.40 (.37)	6.47	1993	< .001
College Degree	-.74 (1.26)	-.59	1993	.56

Table 7.

Global Morality Belief and Reliance on Social Network Size Metrics from Study 4

Predictor	*b*(SE)	*t*	*df*	*p*
Global Morality Belief (1)				
Facebook Friends	1.47 (.71)	2.06	962	.04
SES	1.15 (.30)	3.87	962	< .001
Age	-.01 (.05)	-.30	962	.77
Gender	-1.87 (1.26)	-1.49	962	.14
Religiosity	1.84 (.41)	4.53	962	< .001
College Degree	-3.63 (1.49)	-2.44	962	.02
Global Morality Belief (2)				
Everyday Interactions	3.27 (1.29)	2.53	962	.01
SES	1.16 (1.16)	3.93	962	< .001
Age	-.03 (.05)	-.65	962	.51
Gender	-1.48 (1.26)	-1.17	962	.24
Religiosity	1.67 (.40)	4.64	962	< .001
College Degree	-3.87 (1.49)	-2.60	962	.009
Global Morality Reliance (1)				
Facebook Friends	1.31 (.62)	2.11	962	.03
SES	.56 (.26)	2.19	962	.03
Age	-.01 (.04)	-.25	962	.81
Gender	-.74 (1.09)	-.68	962	.50
Religiosity	1.65 (.35)	4.68	962	< .001

College Degree	.55 (1.29)	.43	962	.67
Global Morality Reliance (2)				
Everyday Interactions	2.23 (1.12)	1.99	962	.047
SES	.58 (.26)	2.30	962	.02
Age	-.03 (.04)	-.73	962	.46
Gender	-.47 (1.09)	-.42	962	.67
Religiosity	1.69 (.35)	4.82	962	< .001
College Degree	.39 (1.29)	.30	962	.76

Note. Country-specific fixes effects were included in each model, but they are not displayed for parsimony.

Table 8.

Cultural Exposure on Moral Ranking Standard Deviations

Predictor	***b*(SE)**	***t***	***df***	***p***
Model 1				
Cultural Exposure	-.51 (.15)	-3.42	81.39	< .001
Judge Gender	.05 (.09)	.55	81.29	.46
Judge Age	-.002 (.003)	-.75	80.59	.46
Subject Gender	-.05 (.05)	-1.00	80.98	.32
Subject Age	-.0002 (.002)	-.13	79.16	.90
Subject Spouse	.08 (.14)	.58	616.52	.56
Model 2				
Foreign Cultural Exposure	-.27 (.13)	-2.08	78.69	.04
Local Cultural Exposure	-.02 (.02)	-.95	80.14	.35
Judge Gender	.01 (.09)	.11	80.37	.91
Judge Age	-.02 (.003)	-.61	79.50	.55
Subject Gender	-.05 (.05)	-.98	79.94	.33
Subject Age	-.003 (.002)	-.14	78.19	.89
Subject Spouse	-.09 (.14)	.61	614.30	.54

Table 9.

Cultural Exposure on Relationship Between Good Heart and Other Moral Rankings

Predictor	***b*(SE)**	***t***	***df***	***p***
Cultural Exposure (centered)	.003 (.20)	.02	663.39	.99
Good Heart (centered)	.37 (.02)	15.17	663.69	< .001
Judge Gender	.02 (.12)	.18	648.33	.86
Judge Age	-.001 (.004)	-.30	655.69	.76
Subject Gender	.16 (.15)	1.07	81.00	.29
Subject Age	.01 (.005)	2.18	75.74	.03
Subject Spouse	.77 (.31)	2.46	633.16	.01
Cultural Exposure * Good Heart	.23 (.08)	2.97	642.81	.003

Table 10.

Key Parameters Across 10 Runs of the Model

	Prediction Sparsity	**Global Conversion Rate**	**Global Prediction %**
Run 1	.72	3.02	73%
Run 2	.73	1.13	53%
Run 3	.74	2.61	70%
Run 4	.71	1.57	63%
Run 5	.71	2.91	76%
Run 6	.74	1.11	52%
Run 7	.71	.82	44%
Run 8	.72	1.89	61%
Run 9	.71	2.29	69%
Run 10	.72	1.72	60%

Note. "Global conversion rate" is the ratio of agents converting to global vs. local prediction strategies at the end of the model. A rate of above 1 suggests that more agents are converting to global vs. local prediction.

Table 11.

Model Robustness Checks

Parameter	Value	Prediction Sparsity	Global Prediction %
δ	n	.72	73%
	n*.5	.61	75%
	n*2	.75	60%
φ	1	.72	73%
	2	.81	55%
	3	.86	61%
γ	100	.80	73%
	150	.72	73%
	200	.71	66%

Note, δ values are scaled to the population because the same constant will lead to less dispersion in larger groups compared to smaller groups.

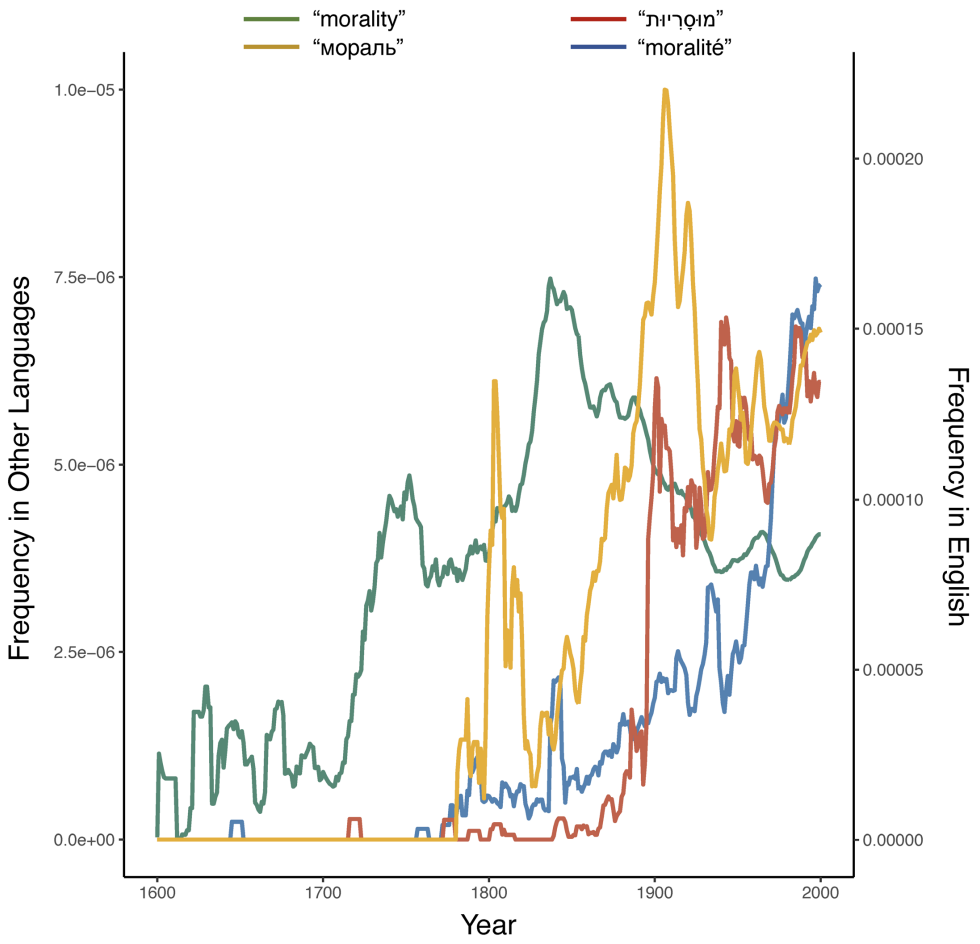

Figure 1. The rise of words communicating the concept of morality in language.

Frequency represents the rate of each word as a proportion of all words in books. The right axis represents the frequency of "morality" in English (and corresponds to the green line), and the right axis represents the frequency of words that represent the concept of morality in other languages (and corresponds to the other lines).

Figure 2. Raphael's School of Athens fresco.

Plato (center left) points to the heavens, while Aristotle (center right) points to the earth.

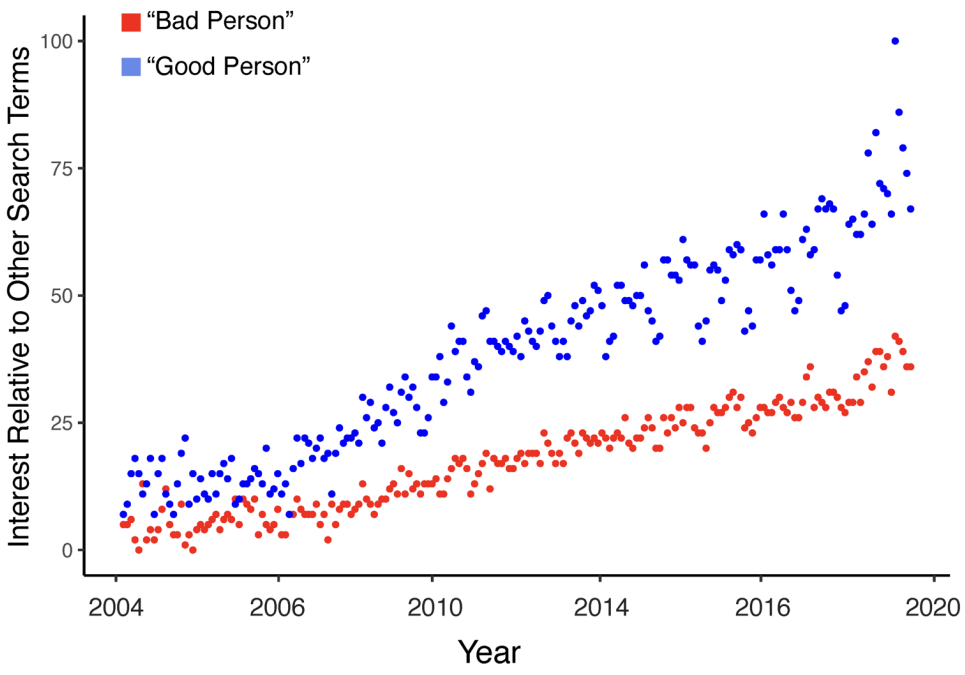

Figure 3. Interest in the phrases "good person" and "evil person" over time

Data come from Google trends, and represent search frequency for each term within the United States, scaled so that a score of 100 represents the highest-frequency data point.

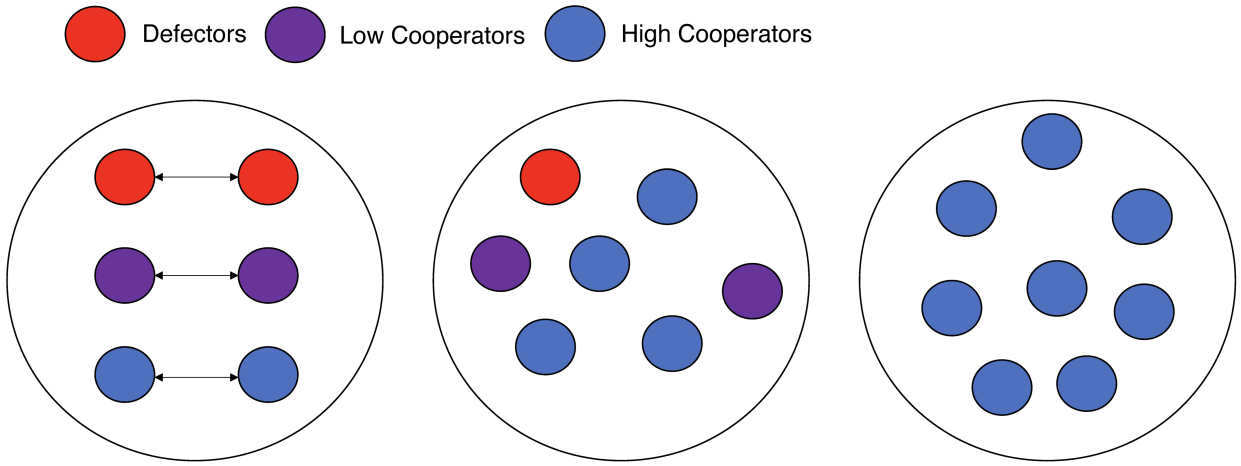

Figure 4. An illustration of partner choice in the evolution of cooperation

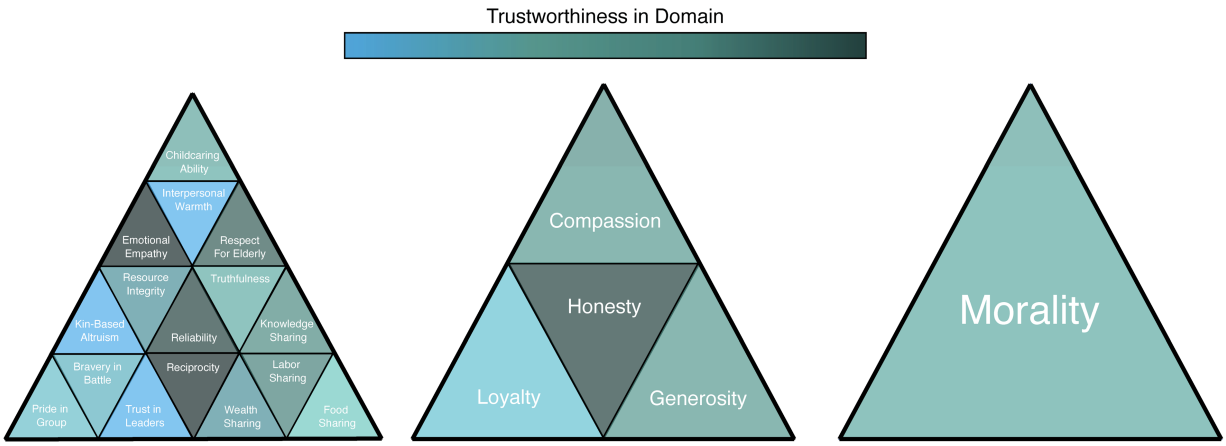

Figure 5. A dynamic conceptual space to morality, illustrated using the Sierpinski triangle

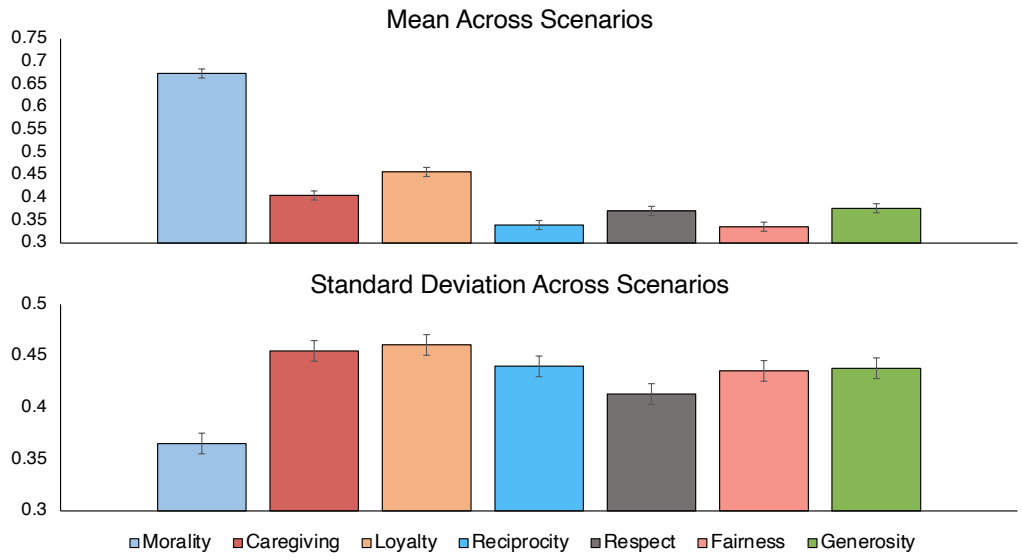

Figure 6. Mean and standard deviation of cooperative traits from Study 1

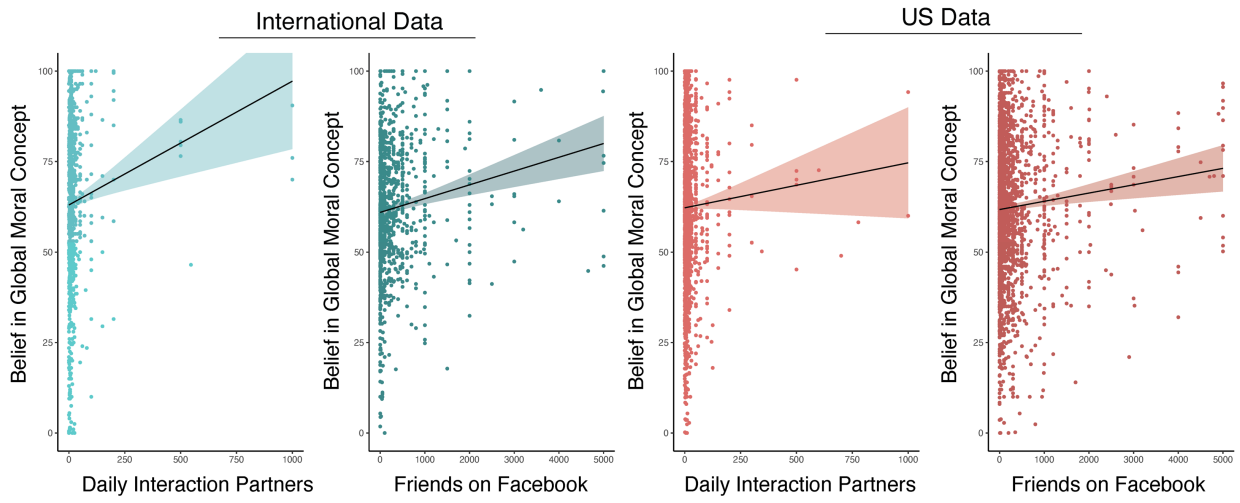

Figure 7. Social network size and the global moral concept

Belief in the global moral concept represents the average of global morality belief and global morality reliance. The left two plots are from Study 4, and the right two plots are from Study 3.

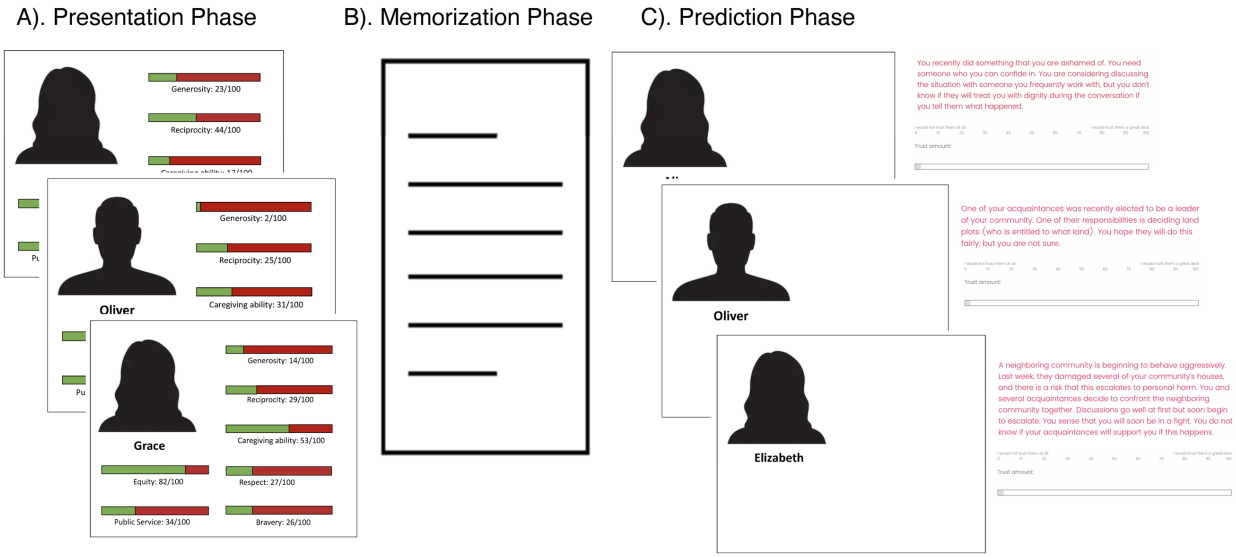

Figure 8. An illustration of Study 5's procedure

Participants first viewed a series of profiles of possible interaction partners in a hypothetical community and took notes on the characteristics of each potential partner. Participants then had the opportunity to memorize these notes on a separate screen. Finally, they viewed the profiles again without trait information and decided how much to trust each partner in a series of cooperative dilemmas.

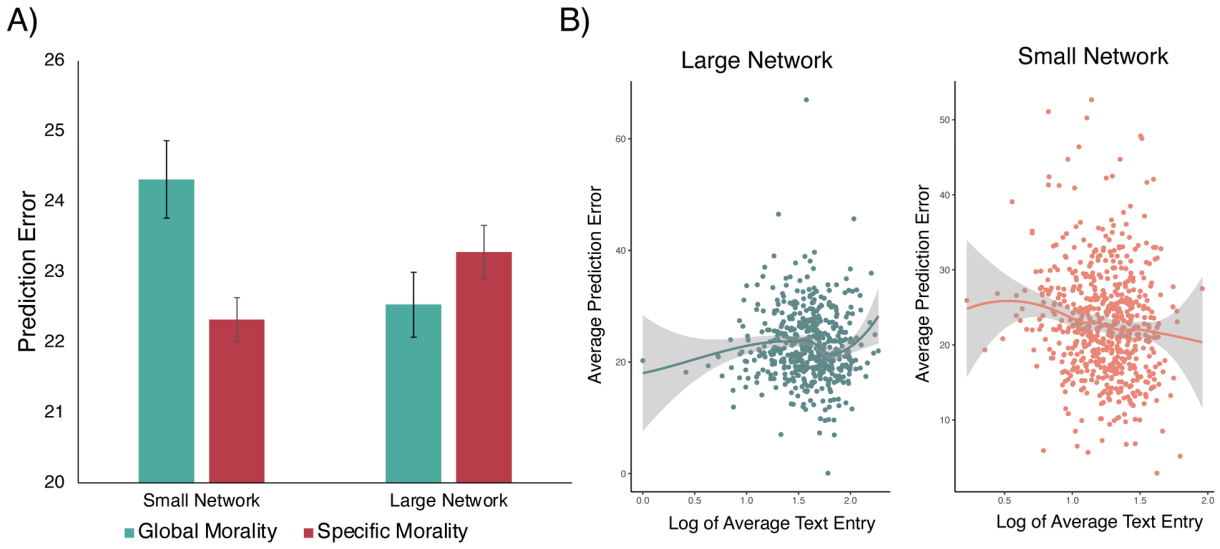

Figure 9. Results from Study 5

Left) The interaction of network size condition and morality strategy on prediction error. Participants using global morality performed significantly worse in the small network condition, but non-significantly better in the large network condition. Right) The correlation of text entry length with prediction error in each condition. In the small (but not large) network condition, text length was negatively associated with prediction error.

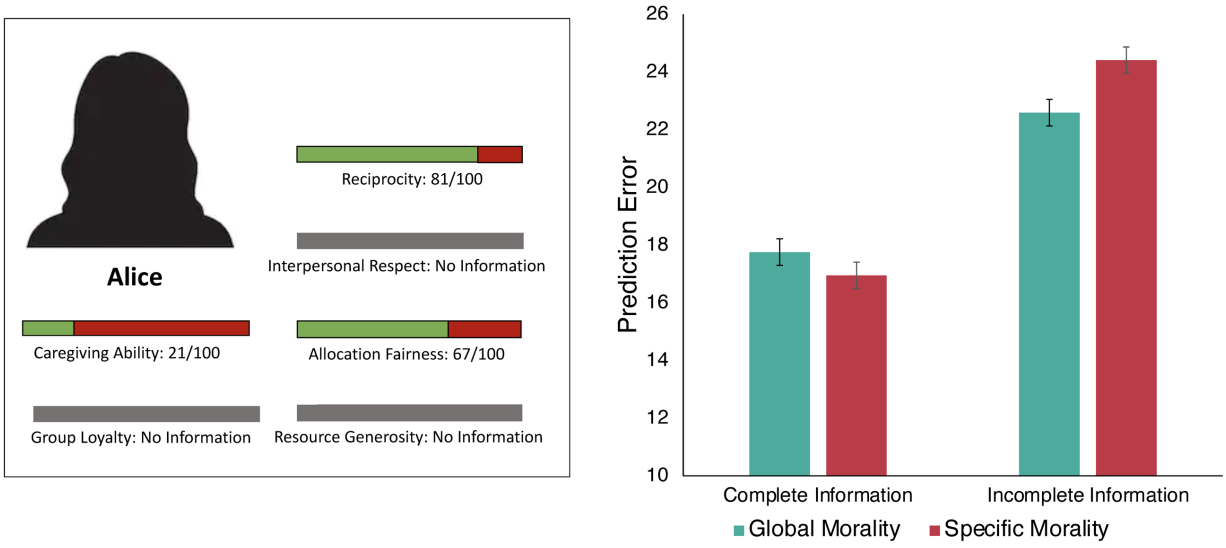

Figure 10. Stimuli and findings in Study 6

Left) An example of an incomplete profile from Study 6. Right) The interaction of information condition and morality condition on prediction error. Participants with complete information about cooperative traits performed non-significantly better when they were in the specific morality condition, whereas participants with incomplete information performed significantly better when they were in the global morality condition.

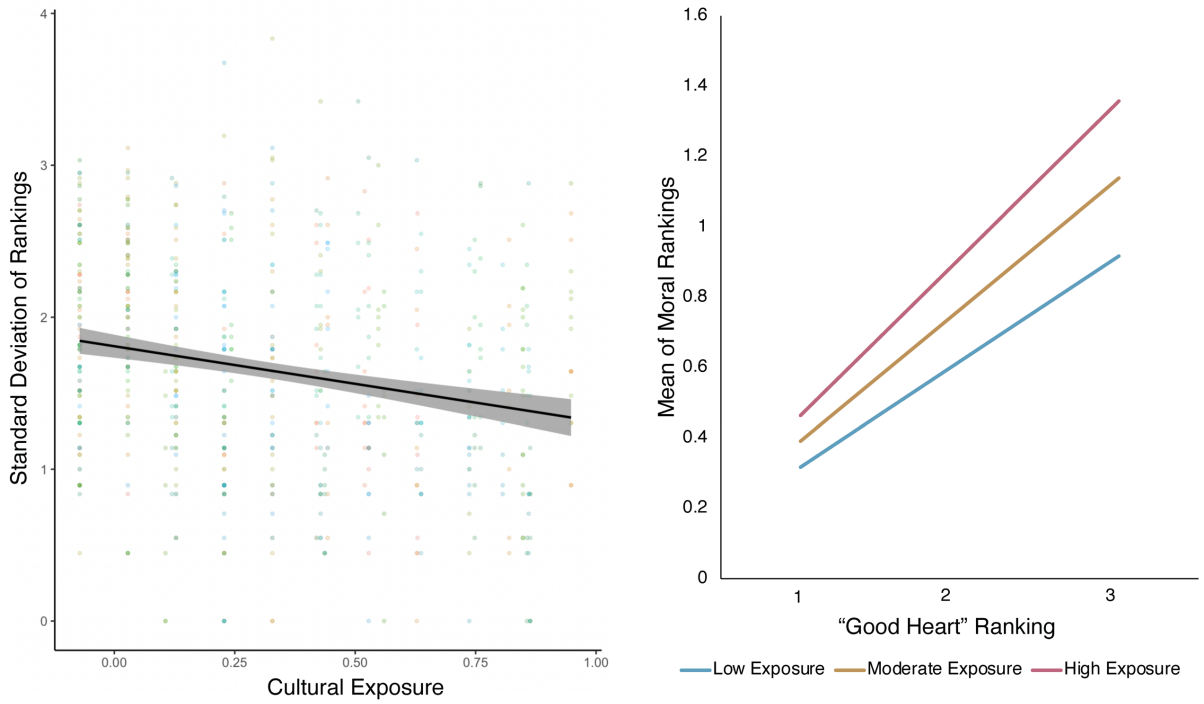

Figure 11. Patterns of moral judgment in the Hadza

Left) The correlation between cultural exposure and the standard deviation of moral attribute rankings in the Hadza. Right) The correlation between rankings of "good heart" and other moral attributes at low, medium, and high levels of cultural exposure.

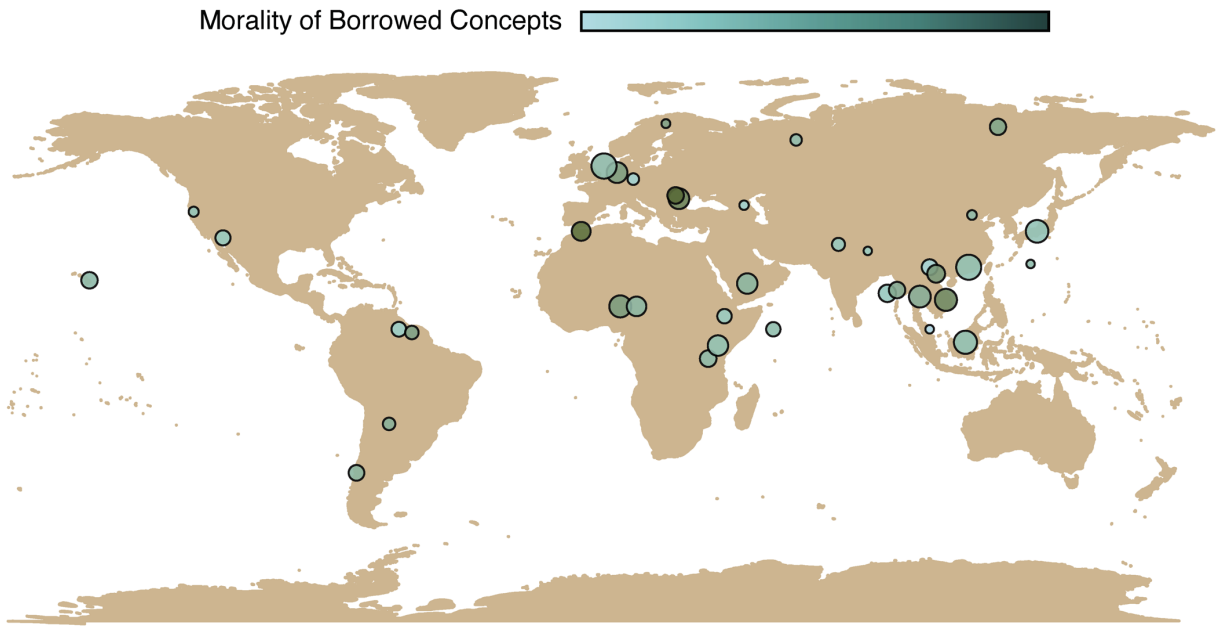

Figure 12. A map of languages in Study 8

A map of languages in Study 8. Node size represents the log of language size (number of speakers). Darker nodes borrow a greater proportion of words about morality.

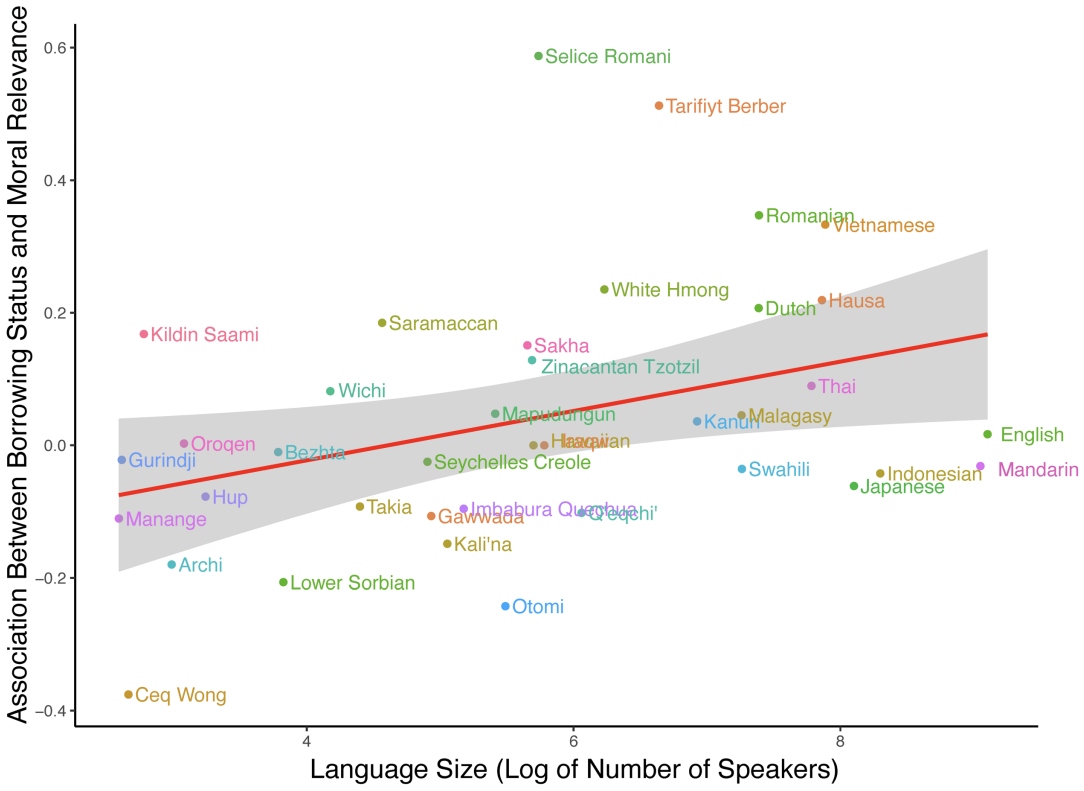

Figure 13. Language size and borrowing patterns in Study 8

The association between language size (log of number of speakers) and the likelihood of borrowing moral vs. non-moral concepts (the association of borrowing status and moral relevance in a language). Node colors represent language family.

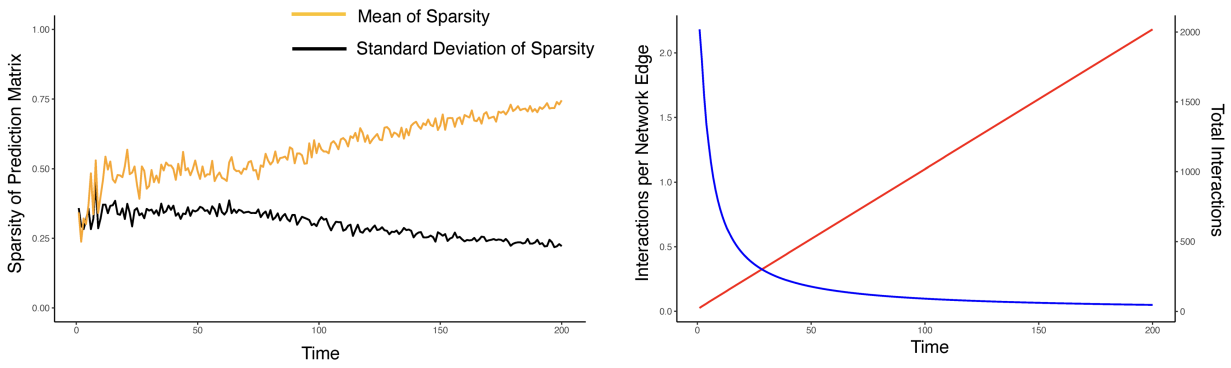

Figure 14. Network sparsity over time in Study 9

Left) Characteristics of network sparsity over time. The mean level of network sparsity is rising, and the standard deviation is falling. Right) The total number of interactions over time plotted against the interactions per network edge. Agents are having more total interactions over time (red), but they are interacting with an increasingly smaller proportion of their social network over time (blue).

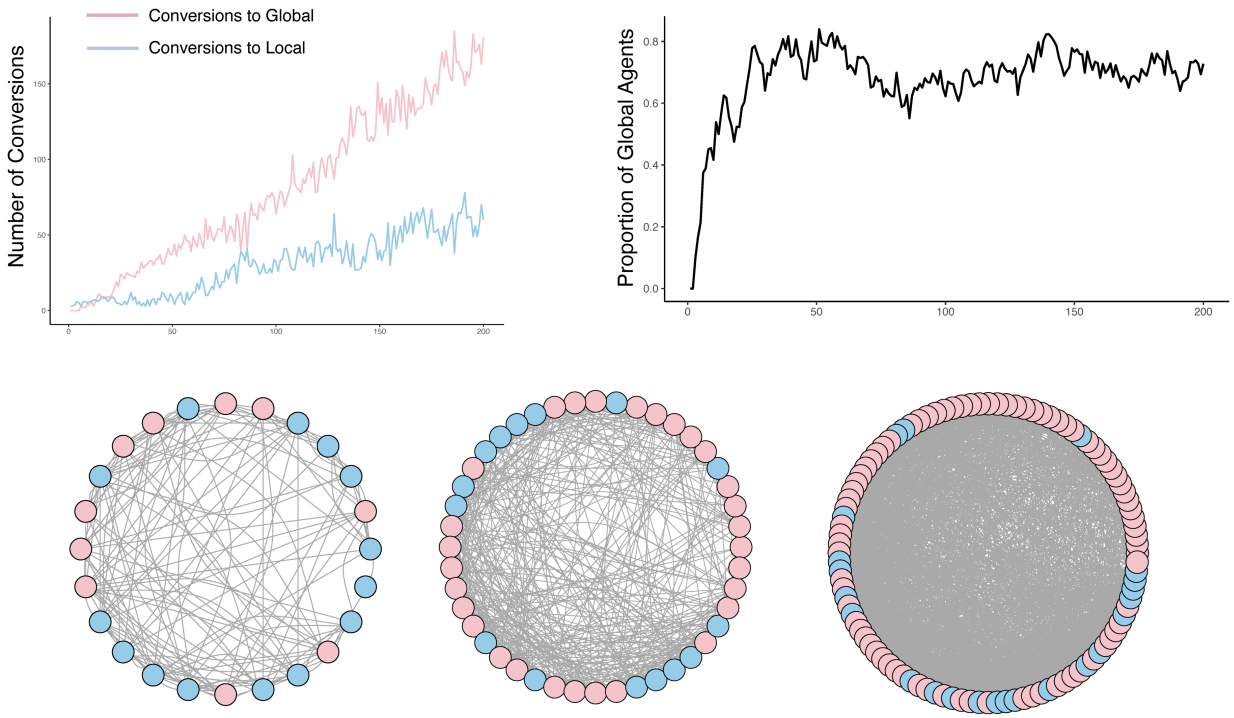

Figure 15. Cooperation prediction strategy and network size in Study 9

Top left) agents are increasingly converting to the global prediction strategy. Top right) Agents with a global prediction strategy represent a growing proportion of agents in the model. Bottom) The network after 10 rounds, 20 rounds, and 50 rounds, where red nodes represent agents with a global prediction strategy and blue nodes represent agents with a local prediction strategy.

REFERENCES

Acheson, J. M. (1988). *The lobster gangs of Maine*. Upne.

Allison, S. T., & Messick, D. M. (1990). Social decision heuristics in the use of shared resources. *Journal of Behavioral Decision Making*, *3*(3), 195–204.

Ames, K. M. (1994). The Northwest Coast: Complex hunter-gatherers, ecology, and social evolution. *Annual Review of Anthropology*, *23*(1), 209–229.

Apicella, C. L. (2018). High levels of rule-bending in a minimally religious and largely egalitarian forager population. *Religion, Brain & Behavior*, *8*(2), 133–148.

Apicella, C. L., Azevedo, E. M., Christakis, N. A., & Fowler, J. H. (2014). Evolutionary origins of the endowment effect: Evidence from hunter-gatherers. *American Economic Review*, *104*(6), 1793–1805.

Apicella, C. L., & Silk, J. B. (2019). The evolution of human cooperation. *Current Biology*, *29*(11), R447–R450.

Aplin, L. M., Farine, D. R., Morand-Ferron, J., Cockburn, A., Thornton, A., & Sheldon, B. C. (2015). Experimentally induced innovations lead to persistent culture via conformity in wild birds. *Nature*, *518*(7540), 538–541.

Barclay, P. (2016). Biological markets and the effects of partner choice on cooperation and friendship. *Current Opinion in Psychology*, *7*, 33–38.

Barclay, P., & Willer, R. (2007). Partner choice creates competitive altruism in humans. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1610), 749–753.

Barrett, H. C., Bolyanatz, A., Crittenden, A. N., Fessler, D. M., Fitzpatrick, S., Gurven, M., Henrich, J., Kanovsky, M., Kushnick, G., & Pisor, A. (2016). Small-scale societies exhibit fundamental variation in the role of intentions in moral judgment. *Proceedings of the National Academy of Sciences*, *113*(17), 4688–4693.

Baumann, N., & Kuhl, J. (2002). Intuition, affect, and personality: Unconscious coherence judgments and self-regulation of negative affect. *Journal of Personality and Social Psychology*, *83*(5), 1213.

Beersma, B., & Van Kleef, G. A. (2012). Why people gossip: An empirical analysis of social motives, antecedents, and consequences. *Journal of Applied Social Psychology*, *42*(11), 2640–2670.

Belfer-Cohen, A. (1991). The natufian in the Levant. *Annual Review of Anthropology*, 167–186.

Johnson, D. (2016). *God is Watching You: How the Fear of God Makes Us Human*. Oxford University Press.

Boehm, C. (1997). Impact of the human egalitarian syndrome on Darwinian selection mechanics. *The American Naturalist*, *150*(S1), S100–S121.

Boyd, R., & Richerson, P. J. (1988). *Culture and the evolutionary process*. University of Chicago press.

Brady, W. J., Crockett, M. J., & Van Bavel, J. J. (2020). The MAD model of moral contagion: The role of motivation, attention, and design in the spread of moralized content online. *Perspectives on Psychological Science*, *15*(4), 978–1010.

Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences*, *114*(28), 7313–7318.

Bruins, H. J., & Van Der Plight, J. (1995). Tell Es-Sultan (Jericho): Radiocarbon results of short-lived cereal and multiyear charcoal samples from the end of the Middle Bronze Age. *Radiocarbon*, *37*(2), 213–220.

Buchtel, E. E., Guan, Y., Peng, Q., Su, Y., Sang, B., Chen, S. X., & Bond, M. H. (2015). Immorality east and west: Are immoral behaviors especially harmful, or especially uncivilized? *Personality and Social Psychology Bulletin*, *41*(10), 1382–1394.

Caluori, N., Jackson, J. C., Gray, K., & Gelfand, M. (2020). Conflict changes how people view God. *Psychological Science*, *31*(3), 280–292.

Cameron, C. D., Lindquist, K. A., & Gray, K. (2015). A constructionist review of morality and emotions: No evidence for specific links between moral content and discrete emotions. *Personality and Social Psychology Review*, *19*(4), 371–394.

Cheney, D. L., & Seyfarth, R. M. (1988). Assessment of meaning and the detection of unreliable signals by vervet monkeys. *Animal Behaviour*, *36*(2), 477–486. https://doi.org/10.1016/S0003-3472(88)80018-6

Crittenden, A. N., Sorrentino, J., Moonie, S. A., Peterson, M., Mabulla, A., & Ungar, P. S. (2017). Oral health in transition: The Hadza foragers of Tanzania. *PloS One*, *12*(3), e0172197.

Curry, O., Whitehouse, H., & Mullins, D. (2019). Is it good to cooperate? Testing the theory of morality-as-cooperation in 60 societies. *Current Anthropology*, *60*(1).

De Waal, F., Macedo, S. E., & Ober, J. E. (2006). *Primates and philosophers: How morality evolved.* Princeton University Press.

Delton, A. W., & Robertson, T. E. (2016). How the mind makes welfare tradeoffs: Evolution, computation, and emotion. *Current Opinion in Psychology*, *7*, 12–16.

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

Dor, D., & Jablonka, E. (2001). From cultural selection to genetic selection: A framework for the evolution of language. *Selection*, *1*(1–3), 33–56.

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

Dunbar, R. I. (2003). The social brain: Mind, language, and society in evolutionary perspective. *Annual Review of Anthropology*, *32*(1), 163–181.

Dunbar, R. I. (2009). The social brain hypothesis and its implications for social evolution. *Annals of Human Biology*, *36*(5), 562–572.

Dunbar, R. I. (2014). The social brain: Psychological underpinnings and implications for the structure of organizations. *Current Directions in Psychological Science*, *23*(2), 109–114.

Dunbar, R. I. (2016). Do online social media cut through the constraints that limit the size of offline social networks? *Royal Society Open Science*, *3*(1), 150292.

Dunford, C. (1977). Kin selection for ground squirrel alarm calls. *The American Naturalist*, *111*(980), 782–785.

Earp, B. D., McLoughlin, K., Monrad, J., Clark, M. S., & Crockett, M. (2020). *How social relationships shape moral judgment*.

Epley, N., Converse, B. A., Delbosc, A., Monteleone, G. A., & Cacioppo, J. T. (2009). Believers' estimates of God's beliefs are more egocentric than estimates of other people's beliefs. *Proceedings of the National Academy of Sciences*, *106*(51), 21533–21538.

Farrow, T. F., Zheng, Y., Wilkinson, I. D., Spence, S. A., Deakin, J. W., Tarrier, N., Griffiths, P. D., & Woodruff, P. W. (2001). Investigating the functional anatomy of empathy and forgiveness. *Neuroreport*, *12*(11), 2433–2438.

Feinberg, M., Willer, R., & Schultz, M. (2014). Gossip and ostracism promote cooperation in groups. *Psychological Science*, *25*(3), 656–664.

Feinberg, M., Willer, R., Stellar, J., & Keltner, D. (2012). The virtues of gossip: Reputational information sharing as prosocial behavior. *Journal of Personality and Social Psychology*, *102*(5), 1015.

Fessler, D. M., Barrett, H. C., Kanovsky, M., Stich, S., Holbrook, C., Henrich, J., Bolyanatz, A. H., Gervais, M. M., Gurven, M., & Kushnick, G. (2015). Moral parochialism and contextual contingency across seven societies. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1813), 20150907.

Finkel, E. J., Bail, C. A., Cikara, M., Ditto, P. H., Iyengar, S., Klar, S., Mason, L., McGrath, M. C., Nyhan, B., & Rand, D. G. (2020). Political sectarianism in America. *Science*, *370*(6516), 533–536.

Fonseca, M. A., & Peters, K. (2018). Will any gossip do? Gossip does not need to be perfectly accurate to promote trust. *Games and Economic Behavior*, *107*, 253–281.

Freud, S. (1989). *The psychopathology of everyday life*. WW Norton & Company.

Fry, D. P., Keith, C. A., & Söderberg, P. (2020). *Social complexity, inequality and war before farming: Congruence of comparative forager and archaeological data*. McDonald Institute for Archaeological Research.

Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., Duan, L., Almaliach, A., Ang, S., & Arnadottir, J. (2011). Differences between tight and loose cultures: A 33-nation study. *Science*, *332*(6033), 1100–1104.

Gintis, H., Henrich, J., Bowles, S., Boyd, R., & Fehr, E. (2008). Strong reciprocity and the roots of human morality. *Social Justice Research*, *21*(2), 241–253.

Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, *106*(1), 148.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. In *Advances in experimental social psychology* (Vol. 47, pp. 55–130). Elsevier.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*(5), 1029.

Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, *23*(2), 101–124.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*(5537), 2105–2108.

Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences*, *6*(12), 517–523.

Grimm, S. B., & Weber, M.-J. (2008). The chronological framework of the Hamburgian in the light of old and new 14C dates. *Quartär*, *55*, 17–40.

Grzega, J. (2003). Borrowing as a word-finding process in cognitive historical onomasiology. *Onomasiology Online*, *4*, 22–42.

Gummerum, M., & Chu, M. T. (2014). Outcomes and intentions in children's, adolescents', and adults' second-and third-party punishment behavior. *Cognition*, *133*(1), 97–103.

Gupta, A. K. (2004). Origin of agriculture and domestication of plants and animals linked to early Holocene climate amelioration. *Current Science*, 54–59.

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.

Haidt, J., & Joseph, C. (2007). The moral mind: How five sets of innate intuitions guide the development of many culture-specific virtues, and perhaps even modules. *The Innate Mind*, *3*, 367–391.

Haslam, N. (2016). Concept creep: Psychology's expanding concepts of harm and pathology. *Psychological Inquiry*, *27*(1), 1–17.

Hauser, M. D., & Marler, P. (1993). Food-associated calls in rhesus macaques (Macaca mulatta): I. Socioecological factors. *Behavioral Ecology*, *4*(3), 194–205.

Haviland, J. B. (1977). *Gossip, reputation, and knowledge in Zinacantan*. University of Chicago Press Chicago.

Heider, F. (1983). *The psychology of interpersonal relations*. Psychology Press.

Heinsohn, T. E. (2010). Marsupials as introduced species: Long-term anthropogenic expansion of the marsupial frontier and its implications for zoogeographic interpretation. *Altered Ecologies: Fire, Climate and Human Influence on Terrestrial Landscapes*, 133–176.

Henrich, J., & Boyd, R. (1998). The evolution of conformist transmission and the emergence of between-group differences. *Evolution and Human Behavior*, *19*(4), 215–241.

Henrich, J., & Broesch, J. (2011). On the nature of cultural transmission networks: Evidence from Fijian villages for adaptive learning biases. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 1139–1148.

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.

Hester, N., & Gray, K. (2020). The moral psychology of raceless, genderless strangers. *Perspectives on Psychological Science*, *15*(2), 216–230.

Heyes, C. (2018). *Cognitive gadgets: The cultural evolution of thinking*. Harvard University Press.

Heyes, C. M., & Frith, C. D. (2014). The cultural evolution of mind reading. *Science*, *344*(6190).

Hill, K. R., Walker, R. S., Božičević, M., Eder, J., Headland, T., Hewlett, B., Hurtado, A. M., Marlowe, F., Wiessner, P., & Wood, B. (2011). Co-residence patterns in hunter-gatherer societies show unique human social structure. *Science*, *331*(6022), 1286–1289.

Iurino, K., & Saucier, G. (2020). Testing measurement invariance of the Moral Foundations Questionnaire across 27 countries. *Assessment*, *27*(2), 365–372.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, *22*, 129–146.

Iyer, R., Koleva, S., Graham, J., Ditto, P., & Haidt, J. (2012). Understanding libertarian morality: The psychological dispositions of self-identified libertarians. *PloS One*, *7*(8), e42366.

Jablonka, E., Ginsburg, S., & Dor, D. (2012). The co-evolution of language and emotions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1599), 2152–2159.

Jablonka, E., & Lamb, M. J. (2007). Précis of evolution in four dimensions. *Behavioral and Brain Sciences*, *30*(4), 353.

Jackson, J. C., Caluori, N., Abrams, S., Beckman, E., Gelfand, M., & Gray, K. (2020). *Tight Cultures and Vengeful Gods: How Culture Shapes Religious Belief*.

Jackson, J. C., Hester, N., & Gray, K. (2018). The faces of God in America: Revealing religious diversity across people and politics. *PloS One*, *13*(6), e0198745.

Jackson, J. C., Watts, J., Henry, T. R., List, J.-M., Forkel, R., Mucha, P. J., Greenhill, S. J., Gray, R. D., & Lindquist, K. A. (2019). Emotion semantics show both cultural variation and universal structure. *Science*, *366*(6472), 1517–1522.

Jackson, J., Watts, J., List, J.-M., Puryear, C., Drabble, R., & Lindquist, K. (2020). *From Text to Thought: How Analyzing Language Can Advance Psychological Science*. https://doi.org/10.31234/osf.io/qat4r

Johnson-Laird, P. N. (2006). *How we reason*. Oxford University Press, USA.

Jong, J., Bluemke, M., & Halberstadt, J. (2013). Fear of death and supernatural beliefs: Developing a new supernatural belief scale to test the relationship. *European Journal of Personality*, *27*(5), 495–506.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kahneman, D., Slovic, S. P., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge university press.

Kelly, R. L. (2013). *The lifeways of hunter-gatherers: The foraging spectrum*. Cambridge University Press.

Kendal, R. L., Boogert, N. J., Rendell, L., Laland, K. N., Webster, M., & Jones, P. L. (2018). Social Learning Strategies: Bridge-Building between Fields. *Trends in Cognitive Sciences*, *22*(7), 651–665. https://doi.org/10.1016/j.tics.2018.04.003

Kennedy, B., Ashokkumar, A., Boyd, R. L., & Dehghani, M. (2021). *Text Analysis for Psychology: Methods, Principles, and Practices*.

Kober, H., Barrett, L. F., Joseph, J., Bliss-Moreau, E., Lindquist, K., & Wager, T. D. (2008). Functional grouping and cortical–subcortical interactions in emotion: A meta-analysis of neuroimaging studies. *Neuroimage*, *42*(2), 998–1031.

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

Landy, J. F., Piazza, J., & Goodwin, G. P. (2016). When it's bad to be friendly and smart: The desirability of sociability and competence depends on morality. *Personality and Social Psychology Bulletin*, *42*(9), 1272–1290.

Landy, J. F., Piazza, J., & Goodwin, G. P. (2018). Morality traits still dominate in forming impressions of others. *Proceedings of the National Academy of Sciences*, *115*(25), E5636–E5636.

Leavens, D. A., Taglialatela, J. P., & Hopkins, W. D. (2014). From grasping to grooming to gossip: Innovative use of chimpanzee signals in novel environments supports both vocal and gestural theories of language origins. In *The evolution of social communication in primates* (pp. 179–194). Springer.

Levari, D. E., Gilbert, D. T., Wilson, T. D., Sievers, B., Amodio, D. M., & Wheatley, T. (2018). Prevalence-induced concept change in human judgment. *Science*, *360*(6396), 1465–1467.

Li, Q. (2020). Urbanization Since 1949: History, Current State and Problems. In *China's Development Under a Differential Urbanization Model* (pp. 1–19). Springer.

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: A meta-analytic review. *The Behavioral and Brain Sciences*, *35*(3), 121.

Marean, C. W. (2016). The transition to foraging for dense and predictable resources and its impact on the evolution of modern humans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1698), 20150239.

Marlowe, F. (2010). *The Hadza: Hunter-gatherers of Tanzania* (Vol. 3). Univ of California Press.

Martin, G. J., & Yurukoglu, A. (2017). Bias in cable news: Persuasion and polarization. *American Economic Review*, *107*(9), 2565–2599.

McAuliffe, K., Jordan, J. J., & Warneken, F. (2015). Costly third-party punishment in young children. *Cognition*, *134*, 1–10.

McBride, A. F., & Kritzler, H. (1951). Observations on pregnancy, parturition, and postnatal behavior in the bottlenose dolphin. *Journal of Mammalogy*, *32*(3), 251–266.

McClintock, C. G., & Allison, S. T. (1989). Social value orientation and helping behavior 1. *Journal of Applied Social Psychology*, *19*(4), 353–362.

McCullough, M. E., Kurzban, R., & Tabak, B. A. (2011). *Evolved mechanisms for revenge and forgiveness.*

McPherson, N. M. (1991). A question of morality: Sorcery and concepts of deviance among the Kabana, West New Britain. *Anthropologica*, 127–143.

Mikhail, J. (2007). Universal moral grammar: Theory, evidence and the future. *Trends in Cognitive Sciences*, *11*(4), 143–152.

Mills, M. G. L. (1991). Conservation management of large carnivores in Africa. *Koedoe*, *34*(1), 81–90.

Mischel, W. (2013). *Personality and assessment*. Psychology Press.

Moll, J., de Oliveira-Souza, R., Eslinger, P. J., Bramati, I. E., Mourão-Miranda, J., Andreiuolo, P. A., & Pessoa, L. (2002). The neural correlates of moral sensitivity: A functional magnetic resonance imaging investigation of basic and moral emotions. *Journal of Neuroscience*, *22*(7), 2730–2736.

Motyl, M., Iyer, R., Oishi, S., Trawalter, S., & Nosek, B. A. (2014). How ideological migration geographically segregates groups. *Journal of Experimental Social Psychology*, *51*, 1–14.

Murphy, R. O., Ackermann, K. A., & Handgraaf, M. (2011). Measuring social value orientation. *Judgment and Decision Making*, *6*(8), 771–781.

Muthukrishna, M., Henrich, J., & Slingerland, E. (2020). Psychology as a historical science. *Annual Review of Psychology*, *72*.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231.

Norenzayan, A., & Shariff, A. F. (2008). The origin and evolution of religious prosociality. *Science*, *322*(5898), 58–62.

Nowak, M. A., Tarnita, C. E., & Wilson, E. O. (2010). The evolution of eusociality. *Nature*, *466*(7310), 1057–1062.

Pehrs, C., Zaki, J., Taruffi, L., Kuchinke, L., & Koelsch, S. (2018). Hippocampal-temporopolar connectivity contributes to episodic simulation during social cognition. *Scientific Reports*, *8*(1), 1–13.

Penke, L., Todd, P. M., Lenton, A. P., & Fasolo, B. (2008). *How self-assessments can guide human mating decisions.*

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In *Communication and persuasion* (pp. 1–24). Springer.

Pike, T. W., & Laland, K. N. (2010). Conformist learning in nine-spined sticklebacks' foraging decisions. *Biology Letters*, *6*(4), 466–468.

Pizarro, D. A., & Tannenbaum, D. (2012). *Bringing character back: How the motivation to evaluate character influences judgments of moral blame.*

Pollom, T. R., Cross, C. L., Herlosky, K. N., Ford, E., & Crittenden, A. N. (2021). Effects of a mixed-subsistence diet on the growth of Hadza children. *American Journal of Human Biology*, *33*(1), e23455.

Purzycki, B. G. (2013). The minds of gods: A comparative study of supernatural agency. *Cognition*, *129*(1), 163–179.

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.

Reeve, H. K., & Nonacs, P. (1992). Social contracts in wasp societies. *Nature*, *359*(6398), 823–825.

Reiter, J., Stinson, N. L., & Le Boeuf, B. J. (1978). Northern elephant seal development: The transition from weaning to nutritional independence. *Behavioral Ecology and Sociobiology*, *3*(4), 337–367.

Rogowski, J. C., & Sutherland, J. L. (2016). How ideology fuels affective polarization. *Political Behavior*, *38*(2), 485–508.

Roscoe, P. (2006). Fish, game, and the foundations of complexity in forager society: The evidence from New Guinea. *Cross-Cultural Research*, *40*(1), 29–46.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In *Advances in experimental social psychology* (Vol. 10, pp. 173–220). Elsevier.

Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology*, *76*(4), 574.

Schein, C. (2020). The importance of context in moral judgments. *Perspectives on Psychological Science*, *15*(2), 207–215.

Schein, C., & Gray, K. (2018). The theory of dyadic morality: Reinventing moral judgment by redefining harm. *Personality and Social Psychology Review*, *22*(1), 32–70.

Schmidt, K. (2000). Göbekli Tepe, southeastern Turkey: A preliminary report on the 1995-1999 excavations. *Paléorient*, 45–54.

Schulz, J. F., Bahrami-Rad, D., Beauchamp, J. P., & Henrich, J. (2019). The Church, intensive kinship, and global psychological variation. *Science*, *366*(6466).

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.

Sherman, J. W., Stroessner, S. J., Conrey, F. R., & Azam, O. A. (2005). Prejudice and stereotype maintenance processes: Attention, attribution, and individuation. *Journal of Personality and Social Psychology*, *89*(4), 607.

Shweder, R. A., Much, N. C., Mahapatra, M., & Park, L. (1997). *The" big three" of morality (autonomy, community, divinity) and the" big three" explanations of suffering.*

Singh, M., & Glowacki, L. (n.d.). *Human social organization during the Late Pleistocene: Beyond the nomadic-egalitarian model*.

Skitka, L. J., & Morgan, G. S. (2014). The social and political implications of moral conviction. *Political Psychology*, *35*, 95–110.

Smaldino, P. E., Lukaszewski, A., von Rueden, C., & Gurven, M. (2019). Niche diversity can explain cross-cultural differences in personality structure. *Nature Human Behaviour*, *3*(12), 1276–1283.

Smith, K. M., & Apicella, C. L. (2020). Partner choice in human evolution: The role of cooperation, foraging ability, and culture in Hadza campmate preferences. *Evolution and Human Behavior*, *41*(5), 354–366.

Snir, A., Nadel, D., Groman-Yaroslavski, I., Melamed, Y., Sternberg, M., Bar-Yosef, O., & Weiss, E. (2015). The origin of cultivation and proto-weeds, long before Neolithic farming. *PLoS One*, *10*(7), e0131422.

Spreng, R. N., & Andrews-Hanna, J. R. (2015). The default network and social cognition. *Brain Mapping: An Encyclopedic Reference*, *1316*, 165–169.

Spreng, R. N., & Mar, R. A. (2012). I remember you: A role for memory in social cognition and the functional neuroanatomy of their interaction. *Brain Research*, *1428*, 43–50.

Sylwester, K., & Roberts, G. (2010). Cooperators benefit through reputation-based partner choice in economic games. *Biology Letters*, *6*(5), 659–662.

Taylor, S. E. (1981). A categorization approach to stereotyping. *Cognitive Processes in Stereotyping and Intergroup Behavior*, *832114*.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, *46*(1), 35–57.

Turchin, P., Currie, T. E., Whitehouse, H., François, P., Feeney, K., Mullins, D., Hoyer, D., Collins, C., Grohmann, S., & Savage, P. (2018). Quantitative historical analysis uncovers a single dimension of complexity that structures global variation in human social organization. *Proceedings of the National Academy of Sciences*, *115*(2), E144–E151.

Turiel, E. (2007). The development of morality. *Handbook of Child Psychology*, *3*.

Uhlmann, E. L., Pizarro, D. A., & Diermeier, D. (2015). A person-centered approach to moral judgment. *Perspectives on Psychological Science*, *10*(1), 72–81.

Vaish, A., Missana, M., & Tomasello, M. (2011). Three-year-old children intervene in third-party moral transgressions. *British Journal of Developmental Psychology*, *29*(1), 124–130.

Vonasch, A. J., Reynolds, T., Winegard, B. M., & Baumeister, R. F. (2018). Death before dishonor: Incurring costs to protect moral reputation. *Social Psychological and Personality Science*, *9*(5), 604–613.

Wang, C. S., Lee, M., Ku, G., & Leung, A. K. -y. (2018). The cultural boundaries of perspective-taking: When and why perspective-taking reduces stereotyping. *Personality and Social Psychology Bulletin*, *44*(6), 928–943.

Watts, J., Greenhill, S. J., Atkinson, Q. D., Currie, T. E., Bulbulia, J., & Gray, R. D. (2015). Broad supernatural punishment but not moralizing high gods precede the evolution of political complexity in Austronesia. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1804), 20142556.

Webster, S. W., & Abramowitz, A. I. (2017). The ideological foundations of affective polarization in the US electorate. *American Politics Research*, *45*(4), 621–647.

Wengrow, D., & Graeber, D. (2015). Farewell to the 'childhood of man': Ritual, seasonality, and the origins of inequality. *Journal of the Royal Anthropological Institute*, *21*(3), 597–619.

White, C., & Peterson, N. (1969). Ethnographic interpretations of the prehistory of western Arnhem Land. *Southwestern Journal of Anthropology*, *25*(1), 45–67.

Whitehouse, H., Francois, P., Savage, P. E., Currie, T. E., Feeney, K. C., Cioni, E., Purcell, R., Ross, R. M., Larson, J., & Baines, J. (2019). Complex societies precede moralizing gods throughout world history. *Nature*, *568*(7751), 226–229.

Wilkinson, G. S. (1990). Food sharing in vampire bats. *Scientific American*, *262*(2), 76–83.

Wilson, D. S., Wilczynski, C., Wells, A., & Weiser, L. (2000). Gossip and other aspects of language as group-level adaptations. *The Evolution of Cognition*, 347–365.

Wood, L. A., Kendal, R. L., & Flynn, E. G. (2013). Copy me or copy you? The effect of prior experience on social learning. *Cognition*, *127*(2), 203–213.

Wright, R. (2010). *The moral animal: Why we are, the way we are: The new science of evolutionary psychology*. Vintage.