Moral Action as Cheater Suppression in Human Superorganisms

Robert Aunger*

London School of Hygiene and Tropical Medicine, London, UK

For it is peculiar to man as compared to the other animals that he alone has a perception of good and bad and just and unjust and other things [of this sort]; and partnership in these things is what makes a household and a city. (Aristotle, The Politics 37)

Aristotle. *The Politics.* Trans. Carnes Lord. Chicago: University of Chicago Press, 1984.

Developments in human technology and social organization have enabled the kinds of social roles that individuals can undertake to proliferate—creating a degree of interdependence not seen in other species. Human societies cannot rely on shared genetic interests or dyadic reciprocity to ensure social cohesion because genetic similarity is low while indirect reciprocity is rife; nevertheless, such societies cohere, due to the evolution of novel regulatory mechanisms that inhibit defaulting on social obligations: moral sentiments and actions. While the degree of social cooperation created by these mechanisms remains less than that of the eusocial insects, it is sufficient to suggest that contemporary human societies constitute crude "superorganisms" to which their members have wide-ranging responsibilities. The present paper argues that the domains and extent of moral regulation can be most usefully identified by defining the set of functions required to sustain a human superorganism. These functions are shown to be boundary, production, distribution, storage, control, structure, enforcement, signaling, memory, excretion, perception, and reproduction. Moral obligations to act then arise when individuals default on contributing to these functions. Major evolutionary transition theory is used to justify claims as to the crucial aspects of superorganism functioning, which enabled human superorganismal groups to form, and thus what aspects of morality had to be developed since the time of our common ancestor with other primates. Finally, comparison is made to Moral Foundations Theory, and the Model of Moral Motives and Dyadic Morality approaches, which suggests that Human Superorganism Theory is simultaneously more parsimonious while being more broadly explanatory. We believe this new approach to defining the moral domain has implications for fields ranging from psychology to legal theory.

Keywords: major transition theory, morality, moral psychology, ultrasociality, superorganisms

DEFINING THE MORAL DOMAIN

Until recently, morality was largely the domain of philosophers. A typical debate might involve defining why killing another human being is wrong (taken as one of the foundations for a good society). Answers might include because it will destabilize society and lead to general unhappiness

(the consequentialist position of Berkeley), or because it violates the Golden Rule of doing to others what we would want them to do to us (subscribed to by deontologists like Kant), or because people of good upbringing and hence virtuous character would not be inclined to do so (the position of virtue theorists like Aristotle). However, relatively little headway was made on such ethical issues because the standard philosophical toolkit of linguistic analysis, logical argument, and introspection provided insufficient means to settle such arguments.

Currently, a variety of disciplines have begun to engage in empirical investigations into the nature of morality. Indeed, morality is currently a hot topic in biology (Alexander, 1987), psychology (Hauser, 2006), philosophy (Joyce, 2007), law (Sunstein, 2005), primatology (de Waal, 2009), and the social sciences (Nichols, 2004). Though controversies continue to rage, there is now general agreement about some aspects of morality. It is thought to be a quintessentially human trait, although the behavioral roots of morality (such as loyalty to kin, intolerance of theft, and punishment of cheats) can be seen in related primate species (Boyd et al., 2003; Boyd and Richerson, 2006; Hauser et al., 2009).

Moral systems have psychological, social, and behavioral components: the psychological mechanisms enable individuals to recognize actions that are moralized and produce intentions to conform to normative expectations as well as to punish moral offenders; social mechanisms like ostracism or shunning help to enforce moral norms; and behavioral violations invoke the activation of the psychological and social mechanisms in the first place (i.e., immoral actions like incest or murder). All of these components need to be in place for moral systems to function properly.

Recent work in psychology has established a number of principles relevant to morality. It is broadly accepted that moral judgments have specific emotional underpinnings (e.g., shame, guilt and retribution) (Haidt, 2003; Moll et al., 2005; Nelissen and Zeelenberg, 2009) and are made rapidly without rational calculation as to harms and benefits (Haidt, 2001)—although when presented with moral dilemmas, people may derive judgments from the interaction of intuitive and rational faculties (Greene et al., 2001). Further, Mikhail's notion of a "moral grammar" may help explain how moral intuitions are generated *via* a set of both rational and emotional components in the brain (Hauser, 2006; Mikhail, 2011).

Debate continues, however, as to exactly what functions morality serves—except to say that moral judgments and action facilitate social cohesion in some way (Hauser, 2006; Joyce, 2007; Bekoff and Pierce, 2009; de Waal, 2009). Indeed, there is a growing consensus that morality evolved to enable human cooperation (Rawls, 1971; Hume, 1998; Haidt and Kesebir, 2010; Ellemers and Van der Toorn, 2015; Curry, 2016). For example, Joyce (2007) argues that making actions moral (i.e., normative) increases their motivational impetus, thus increasing the likelihood of cooperative activity. But why do we think of assault, theft, eating meat, abduction, public nudity, treason, rape, counterfeiting money, and denying the Holocaust as all being immoral? Punishment would not seem to explicitly facilitate cooperation in all these cases. Here, we will argue these concerns all constitute violations of implicit obligations to obey social rules. These obligations derive from a requirement that emerged from developments in social organization during the recent evolutionary history of our species.

In particular, moral actions seek to control defections from public obligations by other members of large-scale human groups. This problem is acute because humans depend on each other—due to a significant division of labor—for everyday necessities, but cannot rely on shared genes or direct reciprocity to ensure ready supplies of what is needed to stay alive. Social organization in large groups of unrelated individuals instead depends on threats of retaliation for failures to abide by cooperative obligations, including fulfilling one's own roles in the social group. These threats are made real by an underlying moral psychology that motivates the punishment of social infractions, even in the face of no direct benefit to the punisher. In effect, anyone in such a social group can potentially punish (or reward) anyone else in the group for "bad" (or "good") behavior. Thus, we will defend a particular proposition that morality evolved to solve particular problems of cooperation—problems associated with ultrasociality. Essentially, morality can be hypothesized to have evolved as a system of behavioral control to aid cooperation in large groups, which cannot rely on shared genes to cause overlapping interests. Further, we will argue that to the degree that these large-scale groups cohere and function (at least partly thanks to moral threats and actions), they can be called human "superorganisms."

The remainder of this paper will outline this new approach to elucidation of the moral domain, Human Superorganism Theory (HSoT). We conclude that identifying the adaptive function of moral action should go a long way toward clarifying the ancillary issues that dominate the contemporary study of morality.

HUMAN SOCIETY AS A SUPERORGANISM

The idea that human societies form superorganisms can be derived from a recent development in evolutionary biology called "major transition theory" (Maynard Smith and Szathmary, 1995). The basic premise of this theory is that the means by which information can be stored, translated, and transmitted from one generation to the next has itself changed a number of times during the history of evolution on Earth. Major transition theory thus suggests that the way evolution works has itself evolved, with the consequence that new levels of functional complexity are achieved with each transition (e.g., the transition from RNA to DNA as the store of genetic information, or unicellular to multicellular life, or multicellular to social life).

The central difficulty in using major transition theory for explaining any evolutionary problem is showing how cheating can be suppressed at the lower level so a higher level of organization can evolve (Buss, 1987; Maynard Smith and Szathmary, 1995; Michod, 1999). This is because there are always incentives (in strategic, or even genetic terms) for the members of an organization to defect from the cooperative actions that sustain the new level of cohesion in favor of their individual interests. Mechanisms for regulating these "selfish" behaviors must therefore evolve to consolidate activities that ensure the new level of

organization coheres and persists. These mechanisms are likely to be specific to each case of transition.

Once the first step of suppression has occurred, there is typically a second stage during which differentiation of function occurs within the new structure (Szathmáry, 2015; West et al., 2015). For example, explaining the evolution of multicellular from unicellular life forms may have to do with a mutation that causes cells not to separate after division, thus suppressing their independence, and then acquiring a reproductive/soma differentiation of function (Gavrilets, 2010).

The most recent transition recognized by major transition theory is that to ultrasociety (Maynard Smith and Szathmáry, 1998). The human species is unique in living in ultrasocial groups—that is, large groups of cooperative members who are essentially unrelated to one another (e.g., in tribal groups, or cities) (Richerson and Boyd, 1998). Recall that for a major transition to occur, fitness differences within constituent groups must be suppressed so that selection at the level of the group becomes the primary evolutionary force at work. Major transition theory argues that a reproductive bottleneck helps reduce conflicts of interest at the group level (Bourke, 2011; Szathmáry, 2015). Some insect societies exhibit extreme restrictions on genetic variability (and hence limit social conflict) by having a single female, the queen, serve as the sole reproducer (Hölldobler and Wilson, 2008).

Human societies have not (yet) achieved such a group-level adaptation. Nevertheless, there is a strong division in the human superorganism between "somatic" functions and the reproductive function. In particular, a specialized institution, the human family, has evolved as the vehicle for reproduction (Emlen, 1995; Geary and Flinn, 2001), which serves to restrict both genetic variability in human populations (which is very low compared to other primate species) (Bowden et al., 2012), and overall group population size. While human females are born with around two million eggs, a variety of mechanisms have evolved to reduce actual completed fertility to a very small number. Sequestering of newly reproductive females in puberty huts and restriction of their social movement until pair-bonded ensure that exposure to males is minimized early in life. Concealed ovulation ensures that males do not fight over access to a woman during a brief, obvious period of fertility, reducing sexual tensions in a group, and making it possible for some reproductive cycles to pass without insemination, reducing fertility (Geary and Flinn, 2001; Marlowe and Berbesque, 2012). During the majority of a human female's reproductive career, she is (except in extreme ecological circumstances) limited to forming a single breeding pair (i.e., monogamy) (Zeitzen, 2008; Chapais, 2013). These single pair-bonds are reinforced through cultural rules associated with the institution of marriage, which make it more difficult for women to break out of this bond (and which also reduces male competition for access to females). Then, when a woman is older, menopause, or an evolved incapacity for further reproduction, ensures that her own reproductive effort ceases, so that she begins to aid the reproductive success of the next generation family to which she is most closely related (Davis and Daly, 1997; Hawkes and Coxworth, 2013). All of these mechanisms serve to tightly regulate female reproductive potential and ensure the isolation of reproduction from other functions in society, as one would expect of a superorganism.

With reproductive function isolated in these ways, other social structures could arise to serve various kinds of "somatic" functions. Here, too, fitness differences have been suppressed. Ultrasociality appears to have been achieved by transforming the ancestral form of social organization in related ape species of rank-based dominance (which creates significant differences in fitness between individuals within a social group) to one of enforced egalitarianism, as seen in extant human forager societies (Boehm, 1999). The initial trigger for the evolution of large groups in our human ancestors may have been a change in diet toward meat eating, which required group foraging (Tomasello et al., 2012; Tomasello and Vaish, 2013). In these societies, potential fitness differences are suppressed through enforced food sharing, monogamy, and alloparenting (sharing the parental care of particular children among families), all of which served to decrease fitness differences within groups. This first level of social interdependence for achieving basic needs was later reinforced by the need for coordinated defense against competing groups (Tomasello et al., 2012; Tomasello and Vaish, 2013). It appears that there was considerable inter-group conflict among our ancestors (Chagnon, 1983; Keeley, 1997; Peterson and Wrangham, 1997). Such inter-group aggression could have a considerable impact on fitness, because entire groups could be decimated, and their reproductive resources (women) lost to other groups from the frequent wars fought by such groups (Wrangham, 1999; Bowles et al., 2003). Battles and raids would have been more likely to be successful to the extent that war parties were large and well organized.

Although these societies lacked centralized policing authorities, they apparently could nevertheless coordinate large-scale inter-group competition *via* punishment mechanisms. Norms are more likely to be enforced as the degree of social interdependence increases (Horne, 2009), as is the case with large groups. For example, the Turkana, a contemporary African pastoralist society, engage in frequent cattle raids on other ethnic groups with parties of several 100 members drawn from most or all of the clans composing that ethnicity. Because these are risky endeavors, men can desert or otherwise defect, but such cheats are informally judged by the community, and beaten, fined, or socially excluded, even by those not present at a raid (Mathew and Boyd, 2011). It is also more likely that "parochial altruism" (cooperation *against* others) will evolve when there is significant inter-group conflict (Bowles, 2006; Choi and Bowles, 2007). Thus, prestate societies without formal institutions (only informal organizational principles such as age grades, gender roles, and endogamy) are able to organize mechanisms that protect and sustain groups of many tens of thousands of individuals.

This combination of enforced egalitarianism within the group, coupled with significant inter-group competition, was an ideal way to minimize within-group selection and maximize inter-group selection, thus facilitating the formation of organizational adaptations at the ultrasocial group level. If this combination was a general condition for prestate societies during human evolution, it would give considerable depth to the history of ultrasocial organization, consistent with the expectation of psychological

adaptations for such a lifestyle (e.g., shame, guilt, and moral concern). Ethnic groups of this size probably arose tens of thousands of years ago—plenty of time for psychological adaptations to have been established. Then, more recently, the agricultural revolution brought the first economic surpluses, and therefore allowed for social specialization (Gowdy and Krall, 2016)—the sort of differentiation of roles within the larger structure expected in the second step of a major transition.

Gene-culture coevolutionary theory provides evidence that specific psychological machinery to support morality—in the form of perceptual biases, social learning abilities, and neurological mechanisms—could have evolved in human populations through the kind of intensive cultural group selection just discussed, accounting for the appearance of moral systems in our recent history (Bowles et al., 2014). According to this theory, humans should have evolved expectations to find some behaviors are sanctioned in the groups to which they belong, to have innate abilities to recognize which behaviors are moralized, to rapidly learn how to perform these behaviors from others in their vicinity through sophisticated forms of social learning, and to experience both conformity with norms and punishment of norm violators as rewarding, so that moral experiences are reinforced (Chudek et al., 2013). Further, this machinery should produce defaults that lead to automatic conformity with social norms, such that cognitive efforts are necessary to override this tendency, so that the prosocial psychology has been internalized in the form of internal sanctions like feelings of shame or guilt. These suppositions are consistent with experiments using behavioral economic games, social psychological studies, and neuro-economic evidence.

Recent evolutionary biological theory also suggests that societies which become truly large share two characteristics: a multi-individual (group-level) production system involving a division of labor and defensible outputs (e.g., nest), and a multi-individual reproductive system in which the young require significant nurturing (e.g., a colony) (Wilson, 2012). These features ensure high levels of social interdependence and form the organizational principles on which natural selection at the level of the group can act to secure cohesive ultrasocial groupings. In primate societies, raising an offspring successfully went from the job of the mother to the family, and in some cases, a troupe of alloparents (Mitani et al., 2012). Simultaneously, the basic economic production unit has gone from individuals to families to professional organizations (such as guilds and companies), as more and more resources have been dedicated to these processes, and a wider variety of interdependent roles have been required for the production of increasingly specialized outputs. Social groups with these two features thus developed both productive and reproductive stores of resources—in the form of domesticated animals and women—that could be contested by other groups, leading them to become desirable targets for takeover through inter-group conflict, as suggested above. Like eusocial insect groups, these human populations could be considered a "factory inside a fortress," which constituted the foundation of selection for superorganisms (Wilson, 2012). This self-reinforcing process of social differentiation leading to increasing resource accumulation allowed further increases in group size, culminating in the very large groups one can find in both social insects and humans.

These large groups of humans with a low average level of relatedness require novel mechanisms for maintaining social cohesion. The suggestion made here is that morality is a control mechanism punishing social defection in human ultrasocial groups (Haidt, 2007, 2012).

The primary question, then, is what set of functions is needed to organize cooperation in ultrasocial groups. Such a set should define the domain of moral action. To answer this question, we will treat human ultrasocial groups as superorganisms. A superorganism can be defined as a "collection of single creatures that together possess the functional organization implicit in the formal definition of organism" (Wilson and Sober, 1989). This is a perspective that has been adopted in evolutionary biology to describe the social organization of a variety of species, ranging from insects like ants, termites, and bees, to mole rats (Wheeler, 1911; Seeley, 1989; Hölldobler and Wilson, 2008). The idea has also been extended to human social groups (Spencer, 1896; Campbell, 1958; Richerson and Boyd, 1999; Foster and Ratnieks, 2005; Wilson et al., 2007; Christakis and Fowler, 2009; Szathmáry, 2015). However, the extent to which human societies cohere is less than that seen in eusocial insects (which have the advantage of high levels of average kinship). Thus, we might say that the human superorganism is relatively "crude" (Richerson and Boyd, 1999; Kesebir, 2012). It might also be the case that our major transition to superorganism status is still in progress and that future developments (e.g., the world wide web) will increase the degree of social cohesion, leading to an even greater relative importance of group selection pressures (Stearns, 2007). In any case, my contention will be that not only is human society organized in a fashion worthy to be called a (crude) superorganism but also that the function of morality is to *police defections from cooperative activities among individuals in their obligations as "cells" in a human superorganism.*

A THEORY OF SUPERORGANISM FUNCTIONALITY

Our task then is to find a way to rigorously define these obligations. This is a step that has not previously been taken in discussions of morality, partly because selection at a particular level of organization will produce novel adaptations to fulfill functions operating only at that level. Thus, one cannot assume that a superorganism (a level of coherent organization above the level of the individual organism) will exhibit all of the systems of a multicellular organism—that is, simply transferring "organs" from one level to the other by analogy is unlikely to work (Turner, 2002). Rather, selection in groups could result in the evolution of novel features to support the development and maintenance of membership in groups (Brewer and Caporael, 2006). What we require then is a foundation for making inferences about the "shape" of a superorganism from a more general theory (Mitchell, 2003).

We can find that theory by recognizing that organisms at any level of organization are complex adaptive systems—that is, they can be described in terms of dynamic processes involving flows of material, energy, and information (Bowles, 2006; Miller and Page, 2007; Sumpter, 2010). As living systems, they are complex

adaptive systems that are open (i.e., exchange material, energy, and information with their environments) and self-organizing (by means of interactions among their elements), with emergent properties. Relevant kinds of organizational functions at the superorganism level can therefore be extracted from "living systems" theory (Miller, 1978), "minimal life" theory (Gánti, 2003; Rasmussen et al., 2007; Bedau, 2011), collective animal behavior (Sumpter, 2010), and eusocial insect ecology (where these societies are treated as superorganisms) (Wheeler, 1911; Seeley, 1989; Hölldobler and Wilson, 2008; Gadau and Fewell, 2009) as well by comparison to the set of organs in multicellular organisms like mammals, and the Indian caste system (taken as an example of a human superorganism) (**Table 1**).[1]

From these sources, we can extract a set of subsystems that all living systems must depend upon to survive and reproduce, regardless of their complexity.[2] Any function nominated by two or more of these sources will be taken to qualify as a subsystem of a human superorganism (i.e., single nominations will be considered to be eccentric). There is a high level of agreement among these sources, with only a few nominations not qualifying by this rule (at the bottom of **Table 1**).

However, there are reasons to include two components, despite their not meeting this criterion, and to exclude another, even though it does. (Remember, one has to take into consideration the special features of superorganisms, rather than simply transfer all potential functions.) Enforcement can be included because it is the primary function of moral action and occurs at the organism level in humans—a close evolutionary analog to the human superorganism—in the form of an immune system. Signaling should also be included because it happens in eusocial insects (although not mentioned explicitly by our source on eusociality)—another close analog of the human superorganism. Including signaling also makes for a complete set of information-based functions to mirror those working on material and energetic aspects.

Another decision has been taken to exclude one function even though it has been nominated by several sources: the motor/movement function, which does not apply to human superorganisms, as they do not move territory or migrate as large units of unrelated people (although eusocial insect colonies can move location when local resources are depleted). (Human superorganisms are more akin to plants than animals in this respect, being concentrated on growth rather than movement.) Movements of superorganism "body" parts (e.g., displays of strength in the form of military parades, or signals of trustworthiness in the shape of free market institutions) can be considered expressive or communicative; hence, the most relevant kinds of superorganism "movement" are encapsulated in the notion of signaling in any case.

Thirteen functions of a human superorganism can be identified in this way (**Table 2**). For example, living systems need boundaries to hold their elements together and protect them from threats in the external environment. These boundaries can manifest as specific structures at different levels of organization, such as a specialized membrane at cell level, skin at organism level, or an army (to patrol territorial encapsulation) at superorganism level.

This list of components can be adequately summarized by a phrase used to describe a eusocial superorganism: "a factory inside a fortress" (Oster and Wilson, 1978). This phrase highlights the sometimes competing goals of defense and production (the basic political and economic functions of any social group). For example, the biological fitness of a eusocial insect colony depends on the ratio of soldiers to workers: if there are more soldiers, the colony is more secure, but this means fewer workers, and hence reduced food production for the colony (i.e., the workers' "factory" job), and hence a smaller colony size. Such colonies have mechanisms to regulate the soldier/worker ratio such that it optimizes colony survival (Hölldobler and Wilson, 2008). Similarly, the number of plumbers, doctors, and lawyers in a human group is loosely regulated by market forces (i.e., how much these professions can earn).

Of particular interest for us is the fact that group cohesion requires agents at sub-levels in the organization to act for the good of the group rather than in their own genetic interest. As a consequence, punishment is needed to keep individual elements in line. In cells, punishment is undertaken by lysosomes, cellular organelles that contain acidic enzymes to break down waste materials and cellular debris. They digest excess or worn-out organelles, food particles, and pathogens. In multicellular organisms, punishment is meted out by the immune system, which identifies pathogens and kills them. In superorganisms, we argue that this policing function is performed by individuals motivated by moral sentiments. These motivations inspire people to punish infractions by other members of their group. Opportunities to serve a moral function arise when individuals fail to perform their functions as components of a superorganism. (Note that we suggest *all* tasks are subject to policing, including failures to serve informational or other control functions—even policing itself.) In such situations, other individuals are required to serve as enforcing agents, bringing the offenders back into line, so that the human superorganism continues to cohere. It is this domain of behavior we wish to investigate here.

Functions in the Human Superorganism

Thus, we suggest that people living in a superorganism serve functions analogous in many cases to the cells in a multicellular organism (Richerson and Boyd, 1999). The human superorganism thus needs to work properly in many ways. We now discuss each of these functions in turn. Remember that we argue moral actions are those which punish those who defect from their obligations. Whereas cells tend to serve a single function—e.g., as muscle or fat cells, neurons, or sensory receptors—individuals do not exhibit the same level of specialization due to the more recent, looser organization of ultrasocial groups. Hence, a particular individual may serve a variety of superorganism functions during their lifetime.

---

[1]Other scholars are working in this area. Heylighen (2007) has previously argued that the functions of a human superorganism can be identified using living systems theory, but not in the context of morality. Kesebir (2012) has suggested that morality helps regulate human superorganisms but does not identify the specific kinds of obligations that would define a moral domain using this insight.

[2]Other sources that are relevant, but that do not provide listings of system components, because they are largely about principles of self-organisation, include cybernetics and general systems theory (von Neumann, 1966; von Bertalanffy, 1968) and biological self-organisation theory (Camazine et al., 2001).

TABLE 1 | Potential components of the human superorganism.[a]

Component	Living systems theory	Mammalian organ systems	Eusocial systems	Indian caste system	Minimal life	Collective animal behavior
Boundary	Boundary	Integumentary system (e.g., skin)	Defensive caste (e.g., soldiers), colony recognition labels, and alarm-defense communication	Kshatriyas (warriors)	Holism (indivisibility)	Predation defense
Control (decision-making)	Decider	Nervous system (brain)		Brahmins (priests)	Control	Collective decision-making and regulation
Production	Producer		Worker caste	Shudras (artisans)	Metabolism	Enhanced foraging
Structure	Supporter	Skeletal system	Nest		Stability	Structures (e.g., nests)
Communication (info-distribution)	Channel and net	Nervous/endocrine systems	Communication interactions		Information carrying	Information transfer
Distribution	Distributor	Circulatory system (i.e., cardiovascular and lymphatic)	Food distribution system	Vaisyas (merchants)		
Reproduction	Reproducer	Reproductive system (e.g., gonads)	Reproductive caste (e.g., queen)			Reproductive opportunities
Perception (info-production)	Input transducer	Sensory systems (e.g., eyes and ears)	Combined sensory organs of members			
Excretion	Extruder	Excretory system (e.g., kidneys and bladder)		Harijans (outcastes)		
Storage	Matter-energy storage	Adipose tissue (i.e., body fat)				
Memory (info-storage)	Memory	Endocannabinoid system[b]				
Signaling (info-excretion)	Output transducer					
Enforcement		Immune systems				
	Motor	Muscular system				Locomotion efficiency
	Ingestor					
	Converter					
	Internal transducer					
	Decoder					
	Associator					
	Encoder					
		Respiratory system				
		Vestibular system (e.g., cochlea)				

[a]Derived from Miller (1978), Hölldobler and Wilson (2008) (Figure 5.1), traditional social distinctions in India (Bhattacharya, 1995), the set of mammalian organ systems, Gánti (2003), and Sumpter (2010), respectively. (Note: collective animal behavior lists the evolutionary benefits of group living, which can be tied to specific functions.)
[b]The recently identified endocannabinoid system (composed of a class of lipids and specialized brain receptors) plays a central role in the regulation of learning and memory. It acts essentially as a negative feedback mechanism within the central nervous system to dampen the release of classic neurotransmitters. By so doing, it helps control energy balance (eating, digestion, adipose storage, and related metabolic processes), but is also actively involved in the formation and storage of long-term memory in the hippocampus (De Oliveira Alvares et al., 2008; Campolongo et al., 2009), and plays a particular role in regulating emotional responses based on memory of fearful events (Ruehle et al., 2012).

Boundary

The boundary serves two natural functions: it keeps the outside from being inside (i.e., establishes a physical line between self and non-self) and helps to defend what is held inside from threats originating in the environment (e.g., the role of cell membranes and multicellular skin). In the case of the human superorganism, the primary job is to keep unwanted people and hazardous substances out of the social "body." The Border Patrol and Coast Guard services in the United States, for example, regulate immigration, exercise customs (i.e., prevent entry of illegal goods and hazardous materials), and control the spread of disease into the country with their quarantining authority. Sociopolitical

TABLE 2 | The human superorganism as a living system.

Subsystem	Function	Examples		
		Cell	Organism	Human superorganism
Systemic functions				
Boundary	Hold components together; regulate entry of elements from environment	Membrane	Skin	Army and border patrol
Enforcement	Internal defense	Lysosome	Immune system	Police
Structure	Maintain proper (spatial) relationships among units	Cytoskeleton (microfilaments/tubules, vacuoles)	Skeleton	Physical infrastructure (e.g., road systems, WWW)
Reproduction	Create similar offspring	Miosis	Sexual reproduction	Family-based reproduction
Control (decision-making)	Coordinate/regulate the system as a whole	Chromosomes	Brain	Government bureaus
Material/energy functions				
Production	Transform materials or provide services for use within system	Ribosome and Golgi complex	Digestive system	Factory
Storage	Retain material/energy within system for later use	Mitochondria (energy) and endoplasmic reticulum (proteins)	Adipose tissue	Warehouse
Distribution	Transport material/energy between system components	Cytoplasm	Circulatory system	Economy
Excretion	Remove wastes from system	Membrane vesicle	Execratory system	Sewer system
Information functions				
Perception (info-production)	Update information on external and internal conditions	Chemical exchange	Sensory organs	Media organizations
Memory (info-storage)	Retain information for later use	Chemical states	Endocannabinoid system	Archives
Communication (info-distribution)	Transmit information between internal components	Chemical signaling (internal)	Peripheral nervous system	Gossip networks
Signaling (info-excretion)	Indicate state/express identity; send messages into external environment	Chemical signaling (external)	Phenotypic markers and speech	Diplomatic corps and public relation organizations

boundaries serve as markers of geographic territory, and as the site at which defensive agents can be deployed.

Boundaries can also move: human superorganisms can grow during their life spans, by increasing the number of families within it, either internally (through demographic "success") or externally, by taking over members of other groups (e.g., through immigration). Growth can also be achieved by conquering entire groups (e.g., by invading a country with a military force) and thus incorporating surrounding territory, changing the location of boundaries.

Enforcement

Sometimes the boundary is ineffectual in protecting the group from hazards (e.g., at organism level, letting in pathogens, or at superorganism level, social parasites). This requires a second line of defense against threats that are (or have become) internal to the group. As stated above, this is the category of primary interest here because it includes moral action.

There are several kinds of threats in fact (**Table 3**). First, there can be problems with those who do not perform their social roles—social parasites, whether they are born into the group, or infiltrate it—who need punishing. Sometimes, even those who

TABLE 3 | Types of enforcement.

Type	Who	Reason for punishment
Informal	Everyone	Non-performance of social role
Second-order	Everyone	Non-performance of informal enforcement role
Formal	Police/judges	"Criminal" enforcement (e.g., murder and assault)

are "insiders" do not follow through on their responsibilities, or break the rules; these individuals need punishing. This is a job that, in the human superorganism, every citizen is informally expected to perform. Types of informal punishment are typically somewhat limited: shunning, verbal abuse, refusal to cooperate, or mild pain inducement.

Further, what can be called "second-order" enforcement, or the punishment of non-punishers, is directed at those who did not perform their function as informal punishers of primary offenses. Forcing non-enforcers to toe the line, or punishing those who engage in non-performance of the enforcement role, helps to produce group cohesion (Gintis, 2000; Hauert et al., 2007). However, perceived responsibility to perform this function can be dissipated by the presence of others in equal

position to take on the task (the so-called "bystander effect") (Darley and Latané, 1968).

Others (e.g., the police and legal system) are also weighed with the professional responsibility of punishing norm violators in the group. This formal enforcement is usually left to the police because enforcement in such cases is considered dangerous (e.g., targets violent offenders), is more serious (e.g., incarceration or major fines), or requires the authority of the state to "stick." What gets punished in this case is often "criminal" malfeasance, which can include over-zealous punishment of within-group members without due cause (e.g., murder and assault). These acts of criminal enforcement are typically motivated by moral sentiments—that is, the individual who perpetrates the crime thinks they have a moral reason that justifies the act (e.g., she is my sexual partner and has violated our relationship; the "condemned" are members of an irreligious community), but the logic is not generally recognized in the cultural group. This includes the category of "crimes against the person" (in particular, their body—in the form of suicide, drug use, homicide, or assault and battery—or liberty, *via* slavery or abduction).

Note that, from a superorganism perspective, "unlawfully" eliminating a member of the group is also an offense against production (due to lost economic benefits) and reproduction (particularly to the family to whom the member belonged). These ramifications make unwarranted violence against insiders emotionally reprehensible. Superorganism theory thus suggests that it is legitimate to injure or take the life of others if they are outsiders (especially in the context of inter-group conflict), and insiders too, if they are not being a "good citizen," but that punishment of "good citizens" is itself punishable.

Structure

Structures like a vertebrate's skeleton maintain spatial and functional relationships among an organism's components. In a superorganism, structures include physical infrastructure such as city buildings, electricity, transportation, and other public service systems. (Note: structure is restricted to physical structures and processes; social processes are under the control function below.) Thus, hampering this function is associated with vandalism of public facilities (including religious buildings, cemeteries, monuments, historic sites, and military installations), sabotage (major disruption to government functions, the use of public services, and potentially major economic loss), and mischief with respect to services (including transportation, water supply, postal services, and voting). These actions can destroy or damage public property, render it dangerous or inoperative, or obstruct its proper use or enjoyment. Such practices are often exhibited by juveniles, perhaps because adolescence is a period during which individuals reflect on whether or not to follow social norms and conformist social roles (Waterman, 1985; Steinberg, 2007). Worker strikes can render well-functioning public services such as transport systems inoperative. Dependency on the infrastructures supporting the internet for many of the other superorganism functions is becoming significant. Hate crimes (e.g., ransacking a church or synagogue, abortion clinic or vivisectionist research lab) are a related phenomenon, which interferes with working relationships among subgroups of the superorganism (even graffiti is often

sexist or racist or directed against religious groups). These kinds of offenses are often considered to be relatively minor but can involve substantial expense to the public purse, and be disruptive to (the quality of) public life. Vandalism can also risk public safety (e.g., by taking place in a context of rioting, where destruction of public property is used as a tactic of political protest).

Reproduction

Reproduction is the most crucial function for any living system, which must survive and duplicate itself (when it cannot persist indefinitely), or simply cease to exist. Insect superorganisms often reproduced by going through a "reproductive funnel" (all reproduction is handed over to a single individual, the hive queen), but in the human superorganism [as in mammalian societies generally (MacLean, 1990)], the unit of reproduction is the family, which produces copies of itself through a complex process of producing component parts—offspring—who then combine with members of other families to form new families. These cooperative breeding units are necessary in humans because it is advantageous to recombine genetic adaptations. This requires sexual unions between unrelated individuals [so there is an optimal outbreeding distance (Bateson, 1983)]. Because offspring are born highly immature and dependent, they also require investment from multiple individuals (ideally), so mothers pair-bond long-term with fathers to help provide that care and the necessary resources, both physical and cultural (since children have a lot to learn as well, being born altricial), for proper growth and development of the new member of society.

There are several ways for a family to fail to reproduce itself: poor choice of mate (e.g., same-sex coitus), or misallocation of reproductive effort (e.g., incest, as a failure of family members to mate outside the family, and thus reproduce the family unit). So in addition to the usual faults of stealing the reproductive resources of others (e.g., rape and child molestation), and inappropriate allocation of mating effort (e.g., bestiality), there is the potentially immoral action of not investing sufficiently in being a parent by failure to raise independent, prosocial offspring (e.g., neglect or abandonment of a child).

Control

Control, as conceived here, is essentially about subsystems that contribute to the guidance, coordination, and regulation of the internal system. Group-level control in a superorganism begins when some individuals restrict access of others to resources; this happens when social dominance or status conflicts arise. Alternatively, groups of low-status individuals can gain power by involving themselves in political coalitions (de Waal, 1982). Where boundary functions are directed toward outsiders, control functions are about regulation of internal processes, particularly social processes (unlike the structure function). These processes occur at several levels of organization. Human superorganisms contain organizations like government, businesses, religious groups, and fan clubs—one of the unique features associated with human ultrasociality. These organizations help to coordinate actions among the members that belong to them and to define relationships between those who belong to different organizations as well. In this way, they help give structure to social relationships.

Local institutions (both public and private) and national governments serve to regulate many of the other functions of those within their jurisdiction. These tend to be dominant forces in HSoTs in "making the rules" by which those in the group live.

Production

Production in the form of economic work is one of the primary requirements of any person living in a human superorganism. The individual's obligation is therefore to find their most economically valuable role in their human superorganism, produce their output as efficiently as possible (i.e., to the best of their ability), and then make it available to the group as a whole. Failure to engage in any of these aspects of production should be culpable. Hence, engaging in subterfuge with respect to the rights of property, including intellectual property (so-called "white collar crime"), and thus gaining an unfair advantage in business, should be morally reprehensible.

Storage

As soon as people began to modify their environments in relatively durable ways through their own labor, there was a need to defend these modifications from theft or appropriation. This includes shelters and tools, which had to be guarded by corporate groups (beginning with families). With the rise of agriculture, social systems began to accumulate surplus material resources, which required storage (e.g., seeds for subsequent planting). Safeguarding such stores became an important function for survival of the group. Destruction of these agricultural stores (e.g., destroying domesticated animals) should therefore be considered harmful. Failure to conserve resources (i.e., waste) or to set aside some portion of surpluses for a "rainy day" can also be seen as failures to optimally allocate resources over time, as well as reckless consumption or loss of resources (e.g., through gambling). Finally, there are criminal failures associated with "storing" people against their will—i.e., so-called "offenses against liberty" such as confinement, detention, and slavery—and not recognizing the value of "human capital," or mistreating people as economic goods or chattel (e.g., slavery).

Distribution

Distribution is one of the primary needs once a social system involves a division of labor, as no one then produces all the means necessary to survive themselves and must acquire some of what they need from others in the group (Smith, 1904; Ridley, 2010). This is obviously the function of an economy (whether regulated or free market). Failures to abide by the rules of exchange (i.e., theft, breaches of contract obligations, and debasement of the means of exchange) should therefore be important targets of moral retribution.

Excretion

Animals of any kind should not surround themselves with their own wastes as this exposes group members to infection. Thus, there are always (implicit) rules in place about how to increase the distance between group members and their own waste (e.g., defecation grounds). This category has been extended in humans to include rules for controlling the social effects of any

of an individual's metabolic processes, such as eating, flatulence, or sneezing (i.e., manners). At the superorganism level, people should also be punished for not contributing to the production or maintenance of clean public spaces as well as sanitation systems such as sewers.

Perception

By perceiving its external and internal environment, an organism maintains and updates its awareness of current conditions. A vital task is to continually update information about any threats and opportunities arising, either from the external environment or from activities within the group itself. Scouting and reconnaissance efforts (e.g., spying) can uncover such information. Those who manage to gain access to such information have an obligation to share it with others in their human superorganism. Failure to do so, or to convey false information, can cause damage or result in lost opportunities, and hence represent moral failures by the responsible parties from the perspective of the human superorganism.

Memory

A living system learns through experience. If the information acquired in this way is not stored, it is lost. Whatever good information individuals have, and whatever useful learning that those in the group acquire through experience, should be stored for later use by themselves or others in the group should similar situations recur. Human groups depend significantly on information stored in linguistic form, which enables many new kinds of more specific information (e.g., social reputations) that can be distributed diffusely throughout the group (as in the nervous system of a multicellular organism). [Reputations begin even earlier than mammals; however, individual contributions to social production can be remembered and recalled even by fish (Alfieri and Dugatkin, 2009).] Groups often use older members of the group as stores of such information, as they are most likely to have had a wide range of experience, and to have survived those experiences. "Forgetfulness" should therefore be a "sin" (e.g., letting old people die without "downloading" their wisdom). Group or cultural memory of the group's history should also be seen as valuable as a means to preserve group identity.

Contemporary human groups also exhibit novel means of intergenerational information inheritance: cultural transmission (including, now, written and digital transmission) (Richerson et al., 2016) and significant niche construction (e.g., rebuilding the physical environment into forms such as cities) (Odling-Smee et al., 2003; Laland et al., 2015). These both lead to an increased ability to store information between generations through non-genetic mechanisms.

Communication

Coordination of functions often requires that different subsystems get inputs about what is happening elsewhere in the system. This is the job of the informational equivalent of the circulation system that transfers material between components. In HSoT, this is accomplished at the interindividual level by linguistic exchanges (e.g., gossip about reputations). This function is facilitated by technological infrastructure such as the telecommunications

networks and world wide web, which allow communication between individuals who are not geographically proximate to one another. Misrepresentations of one's own status or that of others in the social group (i.e., crimes related to reputation) are the primary cause of moral censure with respect to this function.

Signaling

Sometimes it is also necessary to exchange information with the outside environment, including interacting with other superorganisms. Signaling typically concerns the (symbolic) power of the group as a whole (e.g., in inter-group conflicts), not the functioning of particular processes within the group (which affects control). Signaling can be cooperative or combative in intent, depending on the context and target. This "supersocial" communication between groups may involve translation from one language to another, or simple use of displays. For example, it is likely that the evolution of eusociality depended on a high level of group selection, which could have been associated with inter-group conflicts that began as territorial displays by large groups of soldiers (larger groups of soldiers indicate larger group size and hence greater colony strength) (Thorne et al., 2003; Wilson and Hölldobler, 2005). The HSoT equivalent is obviously military parades and displays of weaponry, exemplified by the concept of mutually assured destruction—the display of weapons so powerful it is not necessary to actually use the weaponry in an attack. [Symbolic displays can also be important in the context of mating (e.g., leks and bowerbird nests) (Kirkpatrick and Ryan, 1991); however, this context seems irrelevant to relations between superorganisms.] Defecting on participation in such aggressive displays could have been a significant cost to the group.

Signaling can also be important in cooperative situations, to indicate that a superorganism is a trustworthy partner for international trade or exchange, for example. Signals can also serve as markers of membership in cultural groups—e.g., political coalitions of superorganisms such as NATO or the Allied Powers in recent Western history. It is the job of diplomatic ministers to maintain good relations between governments, and public relations departments between organizations belonging to different superorganisms. The crime of treason or sedition (expression of sentiments in favor of or inciting insurrection against the established order) obviously undermines both cooperative and combative relations between human superorganisms by making the group look weak or uncoordinated.

THE EVOLUTION OF MORALITY IN THE HUMAN LINEAGE

We have already provided (in the introduction) an argument for how human society evolved into a (rough) superorganism; however, now that we have identified the particular functions required for superorganism status, we need to elaborate what developments in moral psychology accompanied this social innovation. As with any complex adaptation, it is likely that moral capacities emerged gradually, through a number of evolutionary innovations, so not all of these functions are crucial to the evolution of ultrasociality and hence human morality, as they were

present prior to the arrival of large-scale human groups. Here, we are interested in identifying which aspects of superorganism function were crucial to making this major transition, and which therefore constitute the essential elements of *human* morality.

Comparative biology can provide clues as to which of the functions appeared in the human lineage prior to ultrasocial living. For example, reproduction through families is characteristic of many mammals, so reproduction could not have been crucial. Similarly, xenophobia and preference for insiders (the psychological underpinnings of boundary) are evident in many mammalian species (Hammond and Axelrod, 2006; Efferson et al., 2008; Fu et al., 2012) so boundary is not key either.

Basic storage, in the form of housing, is present even in insects, but also in reptiles and mammals. Humans expanded the types of structures and devised new means of construction to allow larger structures (e.g., weight-bearing arches) (Aunger, 2010) but largely after the transition to large social groups, so this function does not appear crucial either.

Excretion was largely a personal, private matter until the evolution of cities, which required novel innovations for dealing with human waste (e.g., sewerage systems); these are of very recent origin (i.e., the last thousand years), so this function could not have held up the transition to ultrasociality.

The basic function of memory (long-term memory/recall in brains) is evolutionarily old and may have sufficed to coordinate ultrasocial life (e.g., memory among elders of earlier experiences), but ultrasociality may also have required more reliable intergenerational recall of important information—which was made possible first through writing, possibly about 5,000 years ago [although the first markings on bone date to 30,000 years ago (Davis, 1974)]. The earliest writings tend to document economic exchanges, suggesting a link with markets, and hence large-scale societies. The contemporaneity of writing and large-scale social organization suggests that this "augmented" memory may have been important in helping superorganisms to cohere. Memory continues to evolve new modalities today, as significantly new forms of non-genetic intergenerational inheritance evolved *after* the transition to ultrasociality (e.g., digital communications).

Perception is about monitoring external threats and opportunities—that is, keeping an accurate, up-to-date account of what is going on in the external environment. For example, chimps have been known to patrol their territorial limits as groups to investigate what is going on in other groups (Williams et al., 2004) and to identify ripe fruit sources (Janmaat et al., 2013). Human hunter-gatherers do the same (Cashdan, 1983). Specialists at performing this function are media organizations that keep the population up with any relevant news, but these obviously arose late in history, suggesting that innovations in perception did not constrain ultrasociality.

While signaling outsiders can be done with simple iconic messages (e.g., threats of aggressive intent), Communication within large unrelated groups likely required syntactic language to maintain personal reputations, and thus avoid punishment for being antisocial. Maintaining reputations in large groups requires complex language abilities—the ability to gossip about others (Dunbar, 1998). In particular, recent modeling efforts suggest that the ability to maintain reputations (and thus avoid

punishment) is crucial to the formation of large unrelated groups, as indirect fitness benefits, relatedness, and memory for past dyadic interactions quickly diminish in larger groups (Powers and Lehmann, 2016). When this innovation in communication occurred remains highly contentious, however, although most would argue it happened sometime during hominid or hominoid evolution (Hauser et al., 2002; Corballis, 2012; Morgan et al., 2015). Communication is thus probably uniquely human and may have been required to coordinate ultrasocial cooperation.

Enforcement, in a superorganism context, requires third-party punishment, or the willingness and ability to punish others in the group for antisocial activity, even when the individual themselves has not been involved or injured (Gintis et al., 2008). Further, in major transition theory, it is the enforcement function, which is always crucial to the achievement of higher levels of organization (Corning and Szathmáry, 2015; Szathmáry, 2015; West et al., 2015). The evolution of social complexity thus depends strongly on the existence of effective within-group suppression mechanisms—that is, enforcement that minimizes antisocial activity among group members (Frank, 1995; Bowles et al., 2003; Joyce, 2007). In particular, sanctioning institutions fare better than non-sanctioning ones (Gürerk et al., 2006), and fair institutions (i.e., which reduce free riding) are also more likely to evolve (Kosfeld et al., 2009; Kube et al., 2015) probably because they suppress within-group variability in selective outcomes. It is evident that individuals will actually perform third-party norm enforcement occurs in real-world situations in a variety of contexts (e.g., Balafoutas and Nikiforakis, 2012; Yoeli et al., 2013; Tsusaka et al., 2015; Przepiorka and Berger, 2016).

However, this important mechanism, needed to drive cooperation in large-scale groups, appears lacking in non-human primates. Although they do reject lower-valued rewards than fellow experimental subject animals (Brosnan and de Waal, 2014) and thus care about *relative* outcomes, apes have thus far not been demonstrated to reject unfair offers in ultimatum games (Proctor et al., 2013), suggesting that they do not have a general principle of fairness. Third-party punishment in non-human primates has thus yet to be conclusively demonstrated (Hall and Brosnan, 2016; Talbot et al., 2016; Schmidt and Rakoczy, in press), making enforcement a candidate for the crucial advance. A recent review of experimental studies on responses to unequal rewards in primates suggests that the factor which best correlates (across 10 species) with aggressive responses to unfair offers is the frequency with which members of that species cooperate with non-kin; that is, species which routinely engage in cooperative hunting, coalitions, alliances, and/or food sharing (Hall and Brosnan, 2016). Of course, the primary constraint on the evolution of ultrasociality is cooperation with non-kin, so this correlation is an important confirmation of the centrality of this aspect of moral functionality.

The date at which this innovation occurred is difficult to pinpoint, but it must have been in the last six million years, since that is when our most recent common ancestor diverged from other primates. Presumably, large-scale social groups simply would not form and function without an enforcement mechanism being in place, so we can look to the archeological record for evidence of such groups. (Any sufficiently large human group will have low average relatedness, and hence require third-party enforcement.)

Such groups appear to have been a fairly recent innovation, as the earliest appeared no more than about 10,000 years ago (Bellwood, 2004).

Pool punishment (the game theoretic equivalent of paying a tax toward a police force) can emerge in situations where second-order free riders are punished—that is, where those who fail to punish are themselves punished (Zhang et al., 2014). People prefer to use policing to maintain common-pool resources (Andreoni and Gee, 2012; Traulsen et al., 2012). The evolution of specialized social institutions for internal sanctioning and punishment came surprisingly late. The ancient world (China, Rome, and Greece) did not have specialized domestic police forces, but rather used slaves, local magistrates, or government administrators to enforce laws and regulations, but were notoriously lax in doing so (Hunter, 1994; Jones, 2014). Modern policing began only in the 1600s (Critchley, 1978).

Control of large social groups probably first took place through implicit rule making and reference to social norms. Examples include rules of food sharing in hunter-gatherers, rules for the using irrigation systems in agriculturalists, informal property rights, and systems for sharing reputation between traders (Johnson and Earle, 2000). Use of norms was probably made stronger through sacralization—investing normative options with extra meaning and social force through reference to absolute values—to help ensure prosocial behavior when individual- and group-level incentives conflict (Marshall, 2010).

Later, control became a specialized activity, eventually manifesting itself in the development of a professional class, bureaucrats and/or priests, and new forms of organization such as corporations (Wirth, 1938). A crucial innovation was therefore institutions, or self-created rules of social organization where cooperation can be individually advantageous even in large groups of unrelated individuals. Forming institutions requires shared intentionality, language, and other cognitive abilities largely absent in other primates (Powers et al., 2016). So organizations are quintessentially human, and coevolved with superorganism status.

In their ultrasocial form, all the economic functions—production, distribution, and structures—are related to a division of labor. However, they appear to be consequences, rather than instigators, of ultrasociality—that is, they are only necessary once ultrasociality has arisen. In particular, per capita production increased significantly in human society as social roles differentiated with economic specialization, but this only occurred once cities had developed.

Distribution is not necessary until there are significant social inter-dependencies, due to economic specialization, a recent feature of human society. Distribution can occur "naturally" through the unsupervised diffusion of goods and services *via* multifarious interactions. However, some distributions could lead to conflict over ownership or payment, and therefore require enforcement functions. Distribution is therefore likely to be associated with a concern with fairness, which only arises after cooperative ventures, or when exchanges take place, but again probably only post-facto, not as an initiator, of enforcement. Superorganismal distribution essentially takes advantage of an adaptation that had already arisen for other reasons. In this case, an elaboration of fairness took place: chimps sometimes engage

in group hunting and exhibit a concern for fair allocation or sharing (Boesch, 2002). However, in humans, fairness was extended to cover concern for fair *exchange* of goods produced by separate, non-collaborative activities, once human societies had developed economic specialization.[3]

Similarly, structures such as road and electricity systems are of very recent origin on an evolutionary timescale, and related to city life, but again are derivative of the essential transition point, and probably without needing specific psychological innovations for their construction.

The late development of these "economic" functions suggests that the major transition to superorganism status constitutes what is called a "fraternal" transition (Queller, 2000). In fraternal transitions, similar units join together, reaping the benefits first from new economies of scale, and then evolving a division of labor through differentiation (similar to the development of complex multicellular organisms with multiple organ systems from simple multicellular organisms without internal differentiation). On the other hand, in "egalitarian" transitions, dissimilar units come together, so that their functions are complemented through the formation of a higher-level unit—e.g., the eukaryotic cell, which combines elements from multiple, more primitive cell types. Obviously, in fraternal transitions, the differentiation takes place as the final step of the transition and can continue to increase once the transition has successfully established a higher-order unit (Szathmáry, 2015).

The crucial innovations required to support the transition to ultrasocial living therefore seem to be associated with control, enforcement, and communication—essentially, social institutions, exchange-based punishment, and syntactic language. These developments are in fact interrelated, in that social reputations are maintained by language (gossip) (Joyce, 2007; Tse, 2008), and stories about miscreant activity can lead to third-party punishment. Further, punishment has been institutionalized into police forces in modern societies.

The account produced *via* this logic is similar in many respects to that devised by Krebs (2008) from a somewhat different evolutionary foundation. He suggested that the first development in human morality was social life, which introduced differences of rank (requiring new social emotions such as deference and awe), as well as conflicts of interest. This in turn required the evolution of resolution mechanisms such as social sanctions, followed by the internalization of moral norms *via* fear of punishment, and anticipated regret and remorse, all of which were facilitated in turn by the ability (through so-called "theory of mind") to adopt the perspective of others. Such internalized motivations for moral action have lower biological and social costs, as actual conflict can often be avoided when individuals prejudge the consequences of their actions and curb their antisocial proclivities. A fourth development was then the development of language and symbolic thinking, which facilitated the formation of moral judgments (the foundation of a sense of fairness). These judgments were finally themselves

internalized into moral intuitions (i.e., subconscious predilections to find particular acts immoral or wrong). Such a sequence is consistent with the overall picture painted above, identifying several of the same mechanisms (e.g., complex language and third-party sanctioning) while making some nice theoretical additions in terms of explaining the origin of particular psychological adaptations such as moral intuitions.

Another prominent theorist of the evolution of morality, Franz De Waal, also admits that "objectified" concern for fairness, expressed positively as a generalized empathy, and facilitated by language, is a uniquely human aspect of morality (de Waal, 2009). He says, "We have no evidence that other animals judge the appropriateness of actions that do not affect themselves. The great pioneer of morality research, the Finn Edward Westermarck, explained what makes the moral emotions special: "Moral emotions are disconnected from one's immediate situation: they deal with good and bad at a more abstract, disinterested level." This is what sets human morality apart: a move toward universal standards combined with an elaborate system of justification, monitoring and punishment." (De Waal, 2010)

So punishment of antisocial behavior by anyone, facilitated by the development of instinctive moral tendencies, the maintenance of reputations through language, and the eventual institutionalization of punishment and norms in military and police were the human elaborations of morality that were central to the transition of human society to a superorganism.

COMPARISON TO OTHER APPROACHES

Having laid out the basic claims of HSoT, and provided an evolutionary story for its development, we can now compare it to alternative explanations of the moral domain.

Moral Foundations Theory (MFT)

Perhaps the most advanced and widely accepted contemporary approach to understanding the moral universe is MFT (Haidt and Joseph, 2007). This theory is based on an insight from cross-cultural work that moral considerations exist beyond the typical Western concerns of harm and fairness (which are both associated with individual rights). [4] MFT asserts that there are also moral concerns about spiritual purity and degradation, about fulfilling roles in a social hierarchy, and about expectations of loyalty to the local or national group. It is based in earlier work on three "ethics" or "moral codes" (autonomy, community, and divinity), identified by Richard Shweder during cross-cultural work (Shweder et al., 1997; Rozin et al., 1999). To this base, authority [from relational models theory (Fiske, 1991)] and harm/care [from attachment theory (Bowlby, 1979)] were added, based on a cross-cultural review of research on virtues (Haidt and Joseph,

---

[3]There is one specialized form of exchange in non-human primates, of food for sex (de Waal, 1996), but this reflects a sexual division of labor that exists in many sexual species.

[4]MFT and HSoT share some intellectual foundations, as Haidt (2012) has argued that major transitions in human evolution led to superorganisms, and that morality is an evolved solution to the free rider problem. Nevertheless, I argue that the current paper makes a number of advances over MFT: it provides a derivation of the functions' characteristic of human superorganisms, then links those functions to categories of moral concern, and also demonstrates a strong association between categories of moral concern and types of offense in a consensual criminal code.

2004). In its most recent formulation (Haidt, 2012; Graham et al., 2013; Koleva et al., in press), MFT suggests that morality has six dimensions (what I will call the "Six Dimension Hypothesis")[5]:

- Harm/care (attachment, empathy, and nurture)
- Fairness/cheating (based on reciprocal altruism, with idea of justice, property rights and autonomy)
- Loyalty/betrayal (coalition formation and self-sacrifice for group)
- Authority/subversion (recognition of status differentials; respect for tradition)
- Sanctity/desecration (based in the sense of disgust and contamination, body as a temple, underlies attempts to live elevated, and less carnal life)
- Liberty/oppression (reactance/resentment at dominance; often in tension with authority dimension)

Comparing the domain of moral concern defined by HSoT to that resulting from MFT is quite difficult to do, given that the two theories have somewhat different ambitions: MFT looks at the psychological foundations or dimensions of support for morality, while HSoT identifies kinds of cheating that affect functional problems for social groups. Nevertheless, we can attempt to draw rough parallels between their categories (**Table 4**).[6]

As there is nothing in MFT to which HSoT does not have an equivalent, this comparison suggests that HSoT covers the entire range of MFT, plus some additional territory. This indicates that, although MFT has considerably broadened the notion of morality beyond individual rights and obligations (i.e., harm and fairness concerns), it still has not identified all the ways in which people can "fail their group" and hence be morally punished.

In particular, HSoT identifies several entirely new categories of moral concern. The first is associated with group-level information processing.[7] According to HSoT, one has obligations to correctly transmit information about others through the social group (the communication function), not to violate group memory about its own past and to ensure that others perceive the group correctly. If we make the assumption that legal sanctions reflect a community's moral concern with various types of action, then the fact that there are "information-based" crimes would support HSoT. In fact, failures to deal with social information truthfully are legalized in many criminal codes (as crimes of slander, perjury, and treason, for example).

TABLE 4 | Domains of Human Superorganism Theory (HSoT) versus Moral Foundations Theory (MFT).

MFT	HSoT	Concern
Purity/sanctity	Excretion	Disease avoidance (disgust)
Harm/care	Reproduction	Violence to self/kin
Loyalty/betrayal	Boundary	Insiders versus outsiders
Fairness/cheating	Distribution	Resource allocation issues arising from the division of labor
Authority/subversion	Structure	Status differentials (MFT); general organizational issues (HSoT)
Liberty/oppression	Control	Maintenance of group cohesion and coordination of within-group functions
–	Perception, memory, and communication	Information processing functions
–	Production and storage	Fulfilment of social roles
–	Enforcement	Internal defense

The second novel area concerns the fulfillment of social roles (i.e., the production and storage of material produced). An important aspect of modern human groups is the existence of organizations like businesses, governmental bureaucracies, and local clubs. MFT has little concern for social role fulfillment except for recognizing the authority of others (i.e., being subservient in social relations). HSoT suggests one has an obligation to the group to fulfill one's potential in one or more socially productive roles (contrary to the notion that one is free to express oneself and choose how to spend one's life without concern for others). Again, if failure to fulfill the missions of the various organizations to which one belongs is moralized, that would support HSoT. In fact, various kinds of failures to perform official or private duties are criminalized in many societies.[8]

Further, there is nothing in MFT about a third issue: the moralization of failures to punish or contribute to social welfare, or "second-order morality," associated with the enforcement domain. Again, criminal codes suggest that these areas are in fact moralized: people in some countries can be fined or imprisoned for not helping during emergencies or for taking their punishments of others to an extreme, as in injuring or killing innocent others.

---

[5]The sixth dimension, liberty/oppression, was added in 2011, and several others (including honesty, ownership, and self-control) are currently being investigated as well (Haidt, 2012).

[6]Kesebir (2012) argues that each of the foundations of MFT has a specific role in the human superorganism (working from the earlier version with five dimensions). She suggests that purity/sanctity is about preservation of the "natural order" and of the "meaning system," thus distinguishing moral acts from the mundane. The harm/care foundation derives from the mammalian system of attachment (between mother and offspring). It underlines the basic moral capacity to dislike the pain of others. The in-group loyalty foundation involves hostility toward free riders and traitors and is concerned with "sharpening group boundaries," while fairness/reciprocity underlies the human concern with egalitarianism (i.e., economic exchanges). Hierarchy/duty calls for respectful submission to legitimate authorities, institutions, and cultural traditions.

[7]The MFT questionnaire includes items on cultural history or tradition related to memory but categorizes them differently.

[8]However, one of the dimensions "currently under investigation" by those associated with MFT (Haidt, 2012) is "ownership," which arguably is a psychological motivation underlying production and storage of goods and services, while another, "honesty," could be said to be associated with moral performance of the information functions of perception, memory, and communication (e.g., reputation management). In this way, several of the "holes" in the comparison in Table 8 could be filled, making MFT more similar in range to HSoT, although still without the simple theoretical definition of the moral domain as cheater suppression mechanisms in human superorganisms. Haidt's definition (Haidt, 2012), by contrast, is this: "moral systems are interlocking sets of values, virtues, norms, practices, identities, institutions, technologies, and evolved psychological mechanisms that work together to suppress or regulate self-interest and make cooperative societies possible." This definition does not limit morality to humans, as many species have evolved psychological mechanisms to support cooperation, or if restricted to humans by the focus on institutions and technologies, does not describe why we need these extra mechanisms in order to cooperate effectively.

Finally, MFT has difficulty explaining "victimless crimes" (http://faculty.virginia.edu/haidtlab/Current_Research.html). Why do many people moralize issues related to intake of food, "deviant" sex, and self-abuse? MFT suggests these moral intuitions are related to ideas of purity or sacredness (e.g., suicide is bad because we have a soul; bestiality because it is "unnatural"). On the other hand, HSoT suggests that these so-called "vices" are not moral add-ons, or difficult to explain, but are also fundamentally antisocial in the sense of subverting important functions of the superorganism. These actions are not in fact without harm, but rather cause harm to the "body politic": they involve failures to perform the literally vital functions of reproducing the superorganism, either because individuals are withholding their own productivity as components of the superorganism, or fail to reproduce themselves effectively, thereby not helping to create the next generation of superorganism components (i.e., families). Thus, there are many kinds of actions which MFT does not (currently) cover but which HSoT suggests should be moralized.

Model of Moral Motives (MMM) Approach

Recently, Janoff-Bulman and Carnes (2013) presented a novel approach to describing the moral domain, represented by a diagram with two dimensions: one dimension distinguishes between prescriptive and proscriptive regulation ("shoulds" and "should nots") while the second identifies the target of the regulation: self, other, or group (Janoff-Bulman and Carnes, 2013). These two dimensions create a matrix of six cells covering different kinds of infringements and virtues. They call this the MMM approach. So both inhibiting and activating actions can be directed either at the self (to explain "victimless" transgressions), at others (the domain of standard harm/fairness concerns) or at the group, through rules such as minimizing social inequality and protecting in-group members from external threats.

As the authors of this approach note, MFT covers only three of the six cells in their matrix: ingroup/loyalty, authority/respect, and purity/sanctity are group-oriented proscriptive while harm/care and fairness/reciprocity cover the other-regarding moral rules, with fairness being considered a kind of positive action. This leaves off the self-focused codes and prescriptive group-oriented ones out of the mix. So MMM covers more ground than MFT.

Dyadic Morality Approach

A third approach, called Dyadic Morality, is grounded in the cognitive psychology of concepts, and suggests that morality is understood through the template of a dyad—an intentional moral agent and victim—linked by the causal action taken by the agent on the victim (e.g., physical abuse and cheating) (Gray et al., 2012, 2014). Where each of these components is not explicit, there is a mental operation that completes the dyadic relationship. For example, even "victimless" moral injunctions (e.g., against masturbation) should be considered to have the self as a victim, while flag burning should be considered to victimize the group the flag represents. This victim—whether a person, God, social institution, or future generations—may suffer physically, emotionally, or spiritually. Moral judgment is *always* about the perception of intended harm, which is the single "currency" of moral calculus. Essentially, the approach equates wrongness with harm.

In general, then, HSoT is inconsistent with Dyadic Morality theory (but not MFT) in arguing that morality has a number of foundations or functions (each with a somewhat different psychology), but consistent with Dyadic Morality theory's emphasis on harm as the central principle underlying moral judgments (Gray et al., 2012). HSoT is also technically consistent with the orientation of MMM (Janoff-Bulman and Carnes, 2013), with its emphasis on a wide range of moral infringements, but not with its theoretical argument nor with the content of the moral domain.

Overall Comparison

I argue that HSoT has a number of theoretical advantages over these alternatives. The first is parsimony: all of the claims from HSoT derive from a single proposition: that morality evolved as a system of social obligations designed to "solve" cooperation problems in large groups of unrelated individuals. This is more parsimonious than MFT, which amalgamates six foundations from various sources to explain the domain of moral concerns. Indeed, it is difficult to know where MFT should stop in terms of its identification of foundations, as it has already expanded from three to six dimensions and there are grounds to believe that other foundations remain to be identified (Haidt, 2012). In contrast, the number of domains or foundations for HSoT is fixed. MMM and Dyadic Morality theory can also be said to be founded on single principles (approach/avoidance and the dyad concept, respectively), so they also seem to have an advantage over MFT.

Second is the issue of theoretical foundations. HSoT has its foundation in evolutionary biology (specifically multilevel selection theoretical explanations of major transitions), rather than in the psychology of motivation and self-regulation [for MMM (Janoff-Bulman and Carnes, 2013)] or the cognitive psychology of concepts [for Dyadic Morality (Gray et al., 2012)]. Evolutionary biology is generally thought to be a stronger scientific foundation than specific principles in psychology.

Third is the issue of the extent of the moral domain. Like the alternatives, HSoT includes moral infringements against the self (e.g., suicide, drug abuse, and bestiality), against others (e.g., burglary and murder) and against the group (e.g., destruction of public property and treason). But MMM also includes the possibility of "positive" morality. Although not emphasized here, it is possible that prosocial acts, and rewards for performing the same (e.g., commendations and reputational enhancement), could support superorganismal functionality, and so have been selected for over time, such that psychological predilections for performing such behaviors would be part of moral judgment. This would mean that HSoT covers all the same ground as MMM, but also the purely information-based superorganism functions (perception, memory, communication, and signaling)—a larger overall territory. On the other hand, Dyadic Morality does not provide clear indications as to what the limits of the moral domain might be (i.e., if anything can be a harm of some sort to some sort of victim, what sorts of relations are excluded?). So it is difficult to say that Dyadic Morality theory makes a clear empirical prediction about what can be moral and what non-moral. Further, none of the alternative approaches includes "second-order morality" or punishability of a failure to punish first-order miscreants. HSoT

thus (at least potentially) explains the largest domain among these alternatives. Evidence from criminal law thus suggests social groups do find a very broad range of actions morally objectionable and worthy of punishment, including treason, homicide, theft, adultery, arson, corruption, kidnapping, forgery, slander, fraud, public drunkenness, denigration of a corpse, divulging of secrets, trespass, and failure to disclose knowledge of conspiracy. It therefore seems that all of the things considered by HSoT are what is actually moralized in practice by people. Hence, bringing the domain that can be explained theoretically into closer conformity with practice would seem advantageous. For these reasons, HSoT has the best suite of features and therefore should be preferred.

DISCUSSION

Human Superorganism Theory suggests that a group's morality reflects how that society is organized. Morality is designed to produce "good" behavior, but what is "good" is a function of the ways in which a given superorganism has been organized (including *via* cultural rules). "Good" people are therefore those who fulfill their roles in their superorganism; those who do not are "bad." In this way, specific actions may be required by one superorganism, but not another.

Moral action feeds off failed demonstrations of particular practices, and the sanctioning thereof. It achieves this by moralizing specific cultural practices, to which individuals must conform or risk punishment from their fellows. How these behaviors are selected for moralization is an interesting question. HSoT suggests this process is related to the action's ability to demonstrate prosocial tendencies, or to the action's relationship to some superorganism function. While there may be some arbitrariness in the initial choice of practices used to demonstrate prosocial intentions and trustworthiness, there is likely to be some selection of those which serve a social function. This is certainly the case for manners, for example, where European rules of etiquette that had hygienic functions (e.g., urinating in private) were more likely to persist over several centuries than those which were simply local conventions (such as where to place one's hands during a meal) (Nichols, 2002). Inter-group variation in moral codes can thus be arbitrary and non-functional in their origin, but probably become functional once they become sanctionable moral practices, as they reinforce social cohesion and limit cheating.

Similarly, there can be justifiable variation in moral codes *within* societies. This is seen quite effectively in the degree of concern about the rule of primogeniture—the need for a male heir to continue the family's ability to hold onto wealth. The strength of this need differs between heads of state, and poor people. In the line of English kings, not only wealth but also continued political power depends on a male offspring. The consequences are that any woman who marries into such a family must bear a son, and therefore is treated largely as chattel. She can be beaten, raped, and humiliated by her husband without compunction, and must not have any lovers (because then the fruit of her womb becomes suspect), while the husband can philander to his heart's content—all while entirely different morals hold in lower classes of the same society around the institution of marriage (e.g., men cannot rape

wives without compunction). Only a theory that allows morality to be a function of social organization can explain such variation in what is considered to be "right" behavior *within* societies.

Further, HSoT suggests that *changes* in moral rules over time should be associated with changes in the ways societies are organized. Thus, particular actions can be held to be "right" in a given situation at a given time in a particular superorganism, but not in another because some aspect of social organization has changed. For example, women can only be treated as chattel when they are economically irrelevant. This can be seen in societies that adopted agriculture, where women no longer brought in significant calories (they had been primary provisioners in forager societies). Instead, men were needed to manage beasts of burden to plow the fields, and still did the hunting, so women no longer contributed to primary economic production, allowing them to become just another kind of male property (Sanday, 1973). Thus, HSoT suggests diversity in social rules—whether between societies, within a society, or in the same society over time—can be explained using a single function that is implemented using a universal psychological mechanism. It is simply that substantively variable rules for organizing social life apply in different cultural groups.

All of this depends on the legitimacy of the suggestion that human societies can profitably be considered (crude) superorganisms. As indicated earlier, the truth of this assertion depends on human society having undergone a major transition toward strong group-level regulation of social life. An example of regulation at the social level can be seen in eusocial insect colonies, where the lifespan, growth rate, and rate of reproduction of these colonies are nearly indistinguishable from those of individual organisms (scaled to their difference in sizes), suggesting that superorganisms obey the same kinds of metabolic rules as units at the lower level of organization (Gillooly et al., 2010). Similarly, human social networks regulate metabolic features of the individuals within them. In particular, a person is 10% more likely to be obese if the friends of their friends' friends are obese, even if they do not know those people directly (Christakis and Fowler, 2009). People are not aware of the structure of their own networks and do not control them. Yet how people are connected determines what kinds of functions the network can perform, as well as its qualities (e.g., people tend to be happier in denser networks). Human social networks thus have emergent properties—the primary quality of a complex adaptive system. For these reasons, the collective of all social networks in a population can be called a "human superorganism" (Christakis and Fowler, 2009).[9]

---

[9]Although I have focused here exclusively on individuals, I note that superorganisms are complex superstructures. Just as the human body is not composed simply of cells (but also liquids, minerals, and numerous commensal species), superorganisms are not composed simply of people (Epstein, 2015). Supporting ecological structures help superorganisms to function. Cities, for example, contain human-built structures and systems (e.g., roads, buildings, and electricity grids). So a proper social ontology must also contain physical structures and think of social phenomena as being instantiated *via* practices taking place in particular spaces and times (Schatzki, 2003). Indeed, one can think of the human superorganism as an architecture of instantiated social practices that function together like a living thing (Kemmis et al., 2012).

Of course, the mechanisms through which social control occurs remain important. Human superorganisms differ from eusocial insect colonies in not being able to depend on the genes in such groups sharing a common fate—genetic diversity is much higher in human societies than among haplodiploid social insects. Nor do individuals acquire all their goods directly from those who produce them. Hence, kin selection and reciprocal altruism cannot explain the level of cooperation observed in modern human societies. Here, we have argued that the function of moral concerns is to induce behaviors that help such ultrasocial groups cohere, both through fear of punishment for antisocial behavior and the promise of rewards for prosocial behavior (Janoff-Bulman and Carnes, 2013). This need for regulation arose because of the incredible degree of interdependence between people living in societies characterized by a sophisticated division of labor. We have also sought to delineate the dimensions of the moral domain by outlining the kinds of obligations individuals can owe to a human superorganism.

A major advantage of this approach is that it brings a society's legal systems much closer to its moral systems—they can be shown to overlap in content to a much greater degree than when moral justification is based on traditional approaches limited to fairness and harm. In particular, the categories of moral concern identified in this way also map closely onto the types of offenses in a consensus-based criminal code derived from legal codes from around the world, suggesting that HSoT is consistent with the kinds of problems that social groups actually need to solve through sanctioning systems. HSoT can thus provide a parsimonious yet powerful way of delineating the domain of real-world moral concerns. This could present a significant boon to legal scholars and others seeking to legitimate the criminalization of particular kinds of antisocial behavior.

CONCLUSION

Hume (1998) thought that the need for justice is a function of a society's size: in small societies, members can rely on kinship to align interests and thus minimize conflict; he thought that only when society extends beyond the narrow circle of kin does the need arise for rules to regulate human affairs (Hume, 1998). This is essentially equivalent to the claim (without an evolutionary context) of HSoT that morality is necessary when human societies become composed of large groups of unrelated individuals. HSoT claims that moral concerns arose from the need to police social actions in human "superorganisms" (i.e., large groups without shared kinship). It is grounded in major transitions theory (Maynard Smith and Szathmary, 1995), which explains how human groups organized into crude "superorganisms." HSoT identifies a broader range of moral concerns than competing approaches such as MFT (Haidt, 2007), the MMM approach (Janoff-Bulman and Carnes, 2013), or Dyadic Morality theory (Gray et al., 2012, 2014). It also allows one to develop a sophisticated story about how human morality differentiated itself from simpler forms in our primate relatives. For these reasons, HSoT appears to be an advance in understanding the nature of human morality.

REFERENCES

Alexander, R. (1987). *The Biology of Moral Systems*. Hawthorne, NY: Aldine de Gruyter.

Alfieri, M. S., and Dugatkin, L. A. (2009). Increased survival as a function of interactions with experienced partners. *Ethol. Ecol. Evol.* 21, 101–109. doi:10.1080/08927014.2009.9522499

Andreoni, J., and Gee, L. K. (2012). Gun for hire: delegated enforcement and peer punishment in public goods provision. *J. Public Econ.* 96, 1036–1046. doi:10.1016/j.jpubeco.2012.08.003

Aunger, R. (2010). Types of technology. *Technol. Forecast. Soc. Change* 77, 762–782. doi:10.1016/j.techfore.2010.01.008

Balafoutas, L., and Nikiforakis, N. (2012). Norm enforcement in the city: a natural field experiment. *Eur. Econ. Rev.* 56, 1773–1785. doi:10.1016/j.euroecorev.2012.09.008

Bateson, P. (1983). "Optimal outbreeding," in *Mate Choice*, ed. P. Bateson (Cambridge: Cambridge University Press), 257–277.

Bedau, M. A. (2011). A functional account of degrees of minimal chemical life. *Synthese* 185, 73–88. doi:10.1007/s11229-011-9876-x

Bekoff, M., and Pierce, J. (2009). *Wild Justice: The Moral Lives of Animals*. Chicago, IL: University of Chicago Press.

Bellwood, P. (2004). *First Farmers: The Origins of Agricultural Societies*. New York, NY: Wiley.

Bhattacharya, J. N. (1995). *Hindu Castes and Sects*. Delhi: Munshiram Manoharlal.

Boehm, C. (1999). *Hierarchy in the Forest: The Evolution of Egalitarian Behavior*. Cambridge, MA: Harvard University Press.

Boesch, C. (2002). Cooperative hunting roles among Tai chimpanzees. *Hum. Nat.* 13, 27–46. doi:10.1007/s12110-002-1013-6

Bourke, A. F. (2011). *Principles of Social Evolution*. Oxford: Oxford University Press.

Bowden, R., MacFie, T. S., Myers, S., Hellenthal, G., Nerrienet, E., Bontrop, R. E., et al. (2012). Genomic tools for evolution and conservation in the chimpanzee: *Pan troglodytes ellioti* is a genetically distinct population. *PLoS Genet.* 8:e1002504. doi:10.1371/journal.pgen.1002504

Bowlby, J. (1979). *The Making and Breaking of Affectional Bonds*. London: Tavistock.

Bowles, S. (2006). Group competition, reproductive leveling and the evolution of human altruism. *Science* 314, 1569–1572. doi:10.1126/science.1134829

Bowles, S., Choi, J.-K., and Hopfensitz, A. (2003). The co-evolution of individual behaviors and social institutions. *J. Theor. Biol.* 223, 135–147. doi:10.1016/s0022-5193(03)00060-2

Bowles, S., Loury, G. C., and Sethi, R. (2014). Group inequality. *J Eur. Econ. Assoc.* 12, 129–152. doi:10.1111/jeea.12037

Boyd, R., Gintis, H., Bowles, S., and Richerson, P. J. (2003). The evolution of altruistic punishment. *Proc. Natl. Acad. Sci. U.S.A.* 100, 3531–3535. doi:10.1073/pnas.0630443100

Boyd, R., and Richerson, P. (2006). "Culture and the evolution of human social instincts," in *Roots of Human Sociality: Culture, Cognition, and Interaction*, eds N. Enfield and S. Levinson (New York, NY: Berg). p. 453–67.

Brewer, M. B., and Caporael, L. R. (2006). "An evolutionary perspective on social identity: revisiting groups," in *Evolution and Social Psychology*, eds M. Schaller, J. Simpson, and D. Kenrick (New York, NY: Psychology Press), 143–161.

Brosnan, S. F., and de Waal, F. B. (2014). Evolution of responses to (un)fairness. *Science* 346, 1251776. doi:10.1126/science.1251776

Buss, L. W. (1987). *The Evolution of Individuality*. Princeton, NJ: Princeton University.

Camazine, S., Deneubourg, J. L., Franks, N. R., Sneyd, J., Theraulaz, G., and Bonabeau, E. (2001). *Self-Organization in Biological Systems*. Princeton, NJ: Princeton University Press.

Campbell, D. T. (1958). Common fate, similarity and other indices of the status of aggregates of persons as social entities. *Behav. Sci.* 3, 14–25. doi:10.1002/bs.3830030103

Campolongo, P., Roozendaal, B., Trezza, V., Hauer, D., Schelling, G., McGaugh, J., et al. (2009). Endocannabinoids in the rat basolateral amygdala enhance memory consolidation and enable glucocorticoid modulation of memory. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4888–4893. doi:10.1073/pnas.0900835106

Cashdan, E. (1983). Territoriality among human foragers: ecological models and an application to four Bushman groups. *Curr. Anthropol.* 24, 47–66. doi:10.1086/202934

Chagnon, N. A. (1983). *Yanomamo: The Fierce People*. New York, NY: CBS College Publishing.

Chapais, B. (2013). Monogamy, strongly bonded groups, and the evolution of human social structure. *Evol. Anthropol.* 22, 52–65. doi:10.1002/evan.21345

Choi, J.-K., and Bowles, S. (2007). The coevolution of parochial altruism and war. *Science* 318, 636–640. doi:10.1126/science.1144237

Christakis, N. A., and Fowler, J. H. (2009). *Connected: The Surprising Power of Our Social Networks and How They Shape Our Lives*. New York, NY: Little, Brown & Company.

Chudek, M., Zhao, W., and Henrich, J. (2013). "Culture-gene coevolution, large-scale cooperation, and the shaping of human social psychology," in *Cooperation and Its Evolution*, eds R. Joyce, K. Sterelny, and B. Calcott (Cambridge, CA: MIT Press), 425.

Corballis, M. (2012). *The Recursive Mind: The Origins of Human Language, Thought, and Civilization*. Princeton: Princeton University Press.

Corning, P. A., and Szathmáry, E. (2015). "Synergistic selection": a Darwinian frame for the evolution of complexity. *J. Theor. Biol.* 371, 45–58. doi:10.1016/j.jtbi.2015.02.002

Critchley, T. A. (1978). *A History of Police in England and Wales*. London: Constable Limited.

Curry, O. S. (2016). "Morality as cooperation: a problem-centred approach," in *The Evolution of Morality*, eds T. K. Shackelford and R. D. Hansen (New York, NY: Springer), 27–51.

Darley, J. M., and Latané, B. (1968). Bystander intervention in emergencies: diffusion of responsibility. *J. Pers. Soc. Psychol.* 8, 377–383. doi:10.1037/h0025589

Davis, J. N., and Daly, M. (1997). Evolutionary theory and the human family. *Q. Rev. Biol.* 72, 407–435. doi:10.1086/419953

Davis, S. (1974). Incised bones from the Mousterian of Kebara cave (Mount Carmel) and the Aurignacian of Ha-Yonim cave (Western Galilee), Israel. *Paléorient* 2, 181–182. doi:10.3406/paleo.1974.4181

De Oliveira Alvares, L., Genro, B., Diehl, F., and Quillfeldt, J. (2008). Differential role of the hippocampal endocannabinoid system in the memory consolidation and retrieval mechanisms. *Neurobiol. Learn. Mem.* 90, 1–9. doi:10.1016/j.nlm.2008.01.009

de Waal, F. (1996). *Good Natured: the Origins of Right and Wrong in Humans and Other Animals*. Cambridge, MA: Harvard University Press.

De Waal, F. (2010). Morals without God? *New York Times*, 17th October.

de Waal, F. B. M. (1982). *Chimpanzee Politics: Power and Sex among Apes*. Baltimore, MD: Johns Hopkins University Press.

de Waal, F. B. M. (2009). *Primates and Philosophers: How Morality Evolved*. Princeton, NJ: Princeton University Press.

Dunbar, R. (1998). *Grooming, Gossip, and the Evolution of Language*. Cambridge, MA: Harvard University Press.

Efferson, C., Lalive, R., and Fehr, E. (2008). The coevolution of cultural groups and ingroup favoritism. *Science* 321, 1844–1849. doi:10.1126/science.1155805

Ellemers, N., and Van der Toorn, J. (2015). Groups as moral anchors. *Curr. Opin. Psychol.* 6, 189–194. doi:10.1016/j.copsyc.2015.08.018

Emlen, S. T. (1995). An evolutionary theory of the family. *Proc. Natl. Acad. Sci. U.S.A.* 92, 8092–8099. doi:10.1073/pnas.92.18.8092

Epstein, B. (2015). *The Ant Trap: Rebuilding the Foundations of the Social Sciences*. New York, NY: Oxford University Press.

Fiske, A. (1991). *Structures of Social Life: The Four Elementary Forms of Human Relations*. New York, NY: Free Press.

Foster, K. R., and Ratnieks, F. L. W. (2005). A new eusocial vertebrate? *Trends Ecol. Evol.* 20, 363–364. doi:10.1016/j.tree.2005.05.005

Frank, S. A. (1995). Mutual policing and repression of competition in the evolution of cooperative groups. *Nature* 377, 520–522. doi:10.1038/377520a0

Fu, F., Tarnita, C. E., Christakis, N. A., Wang, L., Rand, D. G., and Nowak, M. A. (2012). Evolution of in-group favoritism. *Nat. Sci. Rep.* 2, 460.

Gadau, J., and Fewell, J. (2009). *Organization of Insect Societies: From Genome to Sociocomplexity*. Boston, MA: Harvard University Press.

Gánti, T. (2003). *The Principles of Life*. New York, NY: Oxford University Press.

Gavrilets, S. (2010). Rapid transition towards the division of labor via evolution of developmental plasticity. *PLoS Comput. Biol.* 6:e1000805. doi:10.1371/journal.pcbi.1000805

Geary, D. C., and Flinn, M. V. (2001). Evolution of human parental behavior and the human family. *Parent. Sci. Prac.* 1, 5–61. doi:10.1080/15295192.2001.9681209

Gillooly, J., Hou, C., and Kaspar, I. M. (2010). Eusocial insects as superorganisms: insights from metabolic theory. *Community Integr. Biol.* 3, 360–362. doi:10.4161/cib.3.4.11887

Gintis, H. (2000). Strong reciprocity and human sociality. *J. Theor. Biol.* 206, 169–179. doi:10.1006/jtbi.2000.2111

Gintis, H., Henrich, J., Bowles, S., Boyd, R., and Fehr, E. (2008). Strong reciprocity and the roots of human morality. *Soc. Justice Res.* 21, 241–253. doi:10.1007/s11211-008-0067-y

Gowdy, J., and Krall, L. (2016). The economic origins of ultrasociality. *Behav. Brain Sci.* 39, e92. doi:10.1017/S0140525X1500059X

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., et al. (2013). Moral foundations theory: the pragmatic validity of moral pluralism. *Adv. Exp. Soc. Psychol.* 47, 55–130. doi:10.1016/B978-0-12-407236-7.00002-4

Gray, K., Schein, C., and Ward, A. F. (2014). The myth of harmless wrongs in moral cognition: automatic dyadic completion from sin to suffering. *J. Exp. Psychol.* 143, 1600–1615. doi:10.1037/a0036149

Gray, K., Waytz, A., and Young, L. (2012). The moral dyad: a fundamental template unifying moral judgment. *Psychol. Inq.* 23, 206–215. doi:10.1080/1047840X.2012.686247

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., and Cohen, J. D. (2001). An fMRI study of emotional engagement in moral judgment. *Science* 293, 2105–2108. doi:10.1126/science.1062872

Gürerk, Ö, Irlenbusch, B., and Rockenbach, B. (2006). The competitive advantage of sanctioning institutions. *Science* 312, 108–111. doi:10.1126/science.1123633

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychol. Rev.* 108, 814–834. doi:10.1037/0033-295X.108.4.814

Haidt, J. (2003). "The moral emotions," in *Handbook of Affective Sciences*, eds R. J. Davidson, K. R. Scherer, and H. H. Goldsmith (Oxford: Oxford University Press), 852–870.

Haidt, J. (2007). The new synthesis in moral psychology. *Science* 316, 998–1002. doi:10.1126/science.1137651

Haidt, J. (2012). *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. New York, NY: Pantheon.

Haidt, J., and Joseph, C. (2004). Intuitive ethics: how innately prepared intuitions generate culturally variable virtues. *Daedalus* 133, 55–66. doi:10.1162/0011526042365555

Haidt, J., and Joseph, C. (2007). "The moral mind: how 5 sets of innate intuitions guide the development of many culture-specific virtues, and perhaps even modules," in *The Innate Mind*, Vol. 3, eds P. Carruthers, S. Laurence, and S. Stich (New York, NY: Oxford), 367–391.

Haidt, J., and Kesebir, S. (2010). "Morality," in *Handbook of Social Psychology*, 5th Edn, eds S. T. Fiske, D. Gilbert, and G. Lindzey (Hoboken, NJ: Wiley), 797–832.

Hall, K., and Brosnan, S. F. (2016). "A comparative perspective on the evolution of moral behavior," in *The Evolution of Morality*, eds T. K. Shackelford and R. D. Hansen (New York, NY: Springer), 157–176.

Hammond, R. A., and Axelrod, R. (2006). The evolution of ethnocentrism. *J. Conflict Resolut.* 50, 926–936. doi:10.1177/0022002706293470

Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A., and Sigmund, K. (2007). Via freedom to coercion: the emergence of costly punishment. *Science* 316, 1905–1907. doi:10.1126/science.1141588

Hauser, M. (2006). *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong*. New York, NY: Ecco.

Hauser, M., McAuliffe, K., and Blake, P. R. (2009). Evolving the ingredients for reciprocity and spite. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 3255–3266. doi:10.1098/rstb.2009.0116

Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science* 298, 1569–1579. doi:10.1126/science.298.5598.1569

Hawkes, K., and Coxworth, J. E. (2013). Grandmothers and the evolution of human longevity: a review of findings and future directions. *Evol. Anthropol.* 22, 294–302. doi:10.1002/evan.21382

Heylighen, F. (2007). The global superorganism: an evolutionary-cybernetic model of the emerging network society. *Soc. Evol. Hist.* 6, 57–117.

Hölldobler, B., and Wilson, E. O. (2008). *The Superorganism: The Beauty, Elegance, and Strangeness of Insect Societies*. New York, NY: W.W. Norton & Co.

Horne, C. (2009). *The Rewards of Punishment. A Relational Theory of Norm Enforcement*. Stanford, CA: Stanford University Press.

Hume, D. (1998). *An Enquiry Concerning the Principles of Morals*. Oxford: Oxford University Press.

Hunter, V. J. (1994). *Policing Athens: Social Control in the Attic Lawsuits, 420-320 B.C.* Princeton, NJ: Princeton University Press.

Janmaat, K. R., Ban, S. D., and Boesch, C. (2013). Chimpanzees use long-term spatial memory to monitor large fruit trees and remember feeding experiences across seasons. *Anim. Behav.* 86, 1183–1205. doi:10.1016/j.anbehav.2013.09.021

Janoff-Bulman, R., and Carnes, N. C. (2013). Surveying the moral landscape: moral motives and group-based moralities. *Pers. Soc. Psychol. Rev.* 17, 219–236. doi:10.1177/1088868313480274

Johnson, A. W., and Earle, T. (2000). *The Evolution of Human Societies: From Foraging Group to Agrarian State*. Stanford, CA: Stanford University Press.

Jones, A. H. M. (2014). *The Decline of the Ancient World*. Routledge.

Joyce, R. (2007). *The Evolution of Morality*. Boston, MA: MIT Press.

Keeley, L. H. (1997). *War before Civilization: The Myth of the Peaceful Savage*. Oxford: Oxford University Press.

Kemmis, S., Edwards-Groves, C., Wilkinson, J., and Hardy, I. (2012). "Ecologies of practices," in *Practice, Learning and Change*, eds P. Hager, A. Lee, and A. Reich (Dordrecht: Springer), 33–49.

Kesebir, S. (2012). The superorganism account of human sociality: how and when human groups are like beehives. *Pers. Soc. Psychol. Rev.* 16, 233–261. doi:10.1177/1088868311430834

Kirkpatrick, M., and Ryan, M. J. (1991). The evolution of mating preferences and the paradox of the lek. *Nature* 350, 33–38. doi:10.1038/350033a0

Koleva, S., Beall, E., and Graham, J. (in press). "Moral foundations theory: building value through moral pluralism," in *The Handbook of Virtue Ethics in Business and Management*, ed. A. J. G. Sison (New York, NY: Springer), 521–530.

Kosfeld, M., Okada, A., and Riedl, A. (2009). Institution formation in public goods games. *Am. Econ. Rev.* 99, 1335–1355. doi:10.1257/aer.99.4.1335

Krebs, D. L. (2008). Morality: an evolutionary account. *Perspect. Psychol. Sci.* 3, 149–172. doi:10.1111/j.1745-6924.2008.00072.x

Kube, S., Schaube, S., Schildberg-Hörisch, H., and Khachatryan, E. (2015). Institution formation and cooperation with heterogeneous agents. *Eur. Econ. Rev.* 78, 248–268. doi:10.1016/j.euroecorev.2015.06.004

Laland, K. N., Uller, T., Feldman, M. W., Sterelny, K., Muller, G. B., Moczek, A., et al. (2015). The extended evolutionary synthesis: its structure, assumptions and predictions. *Proc. Biol. Sci.* 282, 20151019. doi:10.1098/rspb.2015.1019

MacLean, P. D. (1990). *The Triune Brain in Evolution: Role in Paleocerebral Functions*. New York, NY: Plenum Press.

Marlowe, F. W., and Berbesque, J. C. (2012). The human operational sex ratio: effects of marriage, concealed ovulation, and menopause on mate competition. *J. Hum. Evol.* 63, 834–842. doi:10.1016/j.jhevol.2012.09.004

Marshall, D. A. (2010). Temptation, tradition, and taboo: a theory of sacralization. *Sociol. Theory* 28, 64–90. doi:10.1111/j.1467-9558.2009.01366.x

Mathew, S., and Boyd, R. (2011). Punishment sustains large-scale cooperation in prestate warfare. *Proc. Natl. Acad. Sci. U.S.A.* 108, 11375–11380. doi:10.1073/pnas.1105604108

Maynard Smith, J., and Szathmary, E. (1995). *The Major Transitions in Evolution*. Oxford: Oxford University Press.

Maynard Smith, J., and Szathmáry, E. (1998). *The Origins of Life*. Oxford: Oxford University Press.

Michod, R. E. (1999). *Darwinian Dynamics: Evolutionary Transitions in Fitness and Individuality*. Princeton, NJ: Princeton University Press.

Mikhail, J. (2011). *Elements of Moral Cognition: Rawls' Linguistic Analogy and the Cognitive Science of Moral and Legal Judgment*. Cambridge: Cambridge University Press.

Miller, J. G. (1978). *Living Systems*. New York, NY: McGraw Hill.

Miller, J. H., and Page, S. E. (2007). *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton, NJ: Princeton University Press.

Mitani, J. C., Call, J., Kappeler, P. M., Palombit, R. A., and Silk, J. B. (2012). *The Evolution of Primate Societies*. Chicago, IL: University Of Chicago Press.

Mitchell, S. (2003). *Biological Complexity and Integrative Pluralism*. Cambridge: Cambridge University Press.

Moll, J., Zahn, R., Oliveira-Souza, R. D., Krueger, F., and Grafman, J. (2005). The neural basis of human moral cognition. *Nat. Neurosci.* 6, 799–809. doi:10.1038/nrn1768

Morgan, T. J., Uomini, N. T., Rendell, L. E., Chouinard-Thuly, L., Street, S. E., Lewis, H. M., et al. (2015). Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nat. Commun.* 6, 6029. doi:10.1038/ncomms7029

Nelissen, R. M. A., and Zeelenberg, M. (2009). Moral emotions as determinants of third-party punishment: anger, guilt, and the functions of altruistic sanctions. *Judgm. Decis. Mak.* 4, 543–553.

Nichols, S. (2002). On the genealogy of norms: a case for the role of emotion in cultural evolution. *Philos. Sci.* 69, 234–255. doi:10.1086/341051

Nichols, S. (2004). *Sentimental Rules: On the Natural Foundation of Moral Judgment*. Oxford: Oxford University Press.

Odling-Smee, F. J., Laland, K. N., and Feldman, M. (2003). *Niche Construction: The Neglected Process in Evolution*. Princeton, NJ: Princeton University Press.

Oster, G., and Wilson, E. (1978). *Caste and Ecology in the Social Insects*. Princeton, NJ: Princeton University Press.

Peterson, D., and Wrangham, R. (1997). *Demonic Males: Apes and the Origins of Human Violence*. New York, NY: Mariner Books.

Powers, S. T., and Lehmann, L. (2016). When is bigger better? The effects of group size on the evolution of helping behaviours. *Biol. Rev.* doi:10.1111/brv.12260

Powers, S. T., van Schaik, C. P., and Lehmann, L. (2016). How institutions shaped the last major evolutionary transition to large-scale human societies. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 371, 20150098. doi:10.1098/rstb.2015.0098

Proctor, D., Williamson, R. A., de Waal, F. B., and Brosnan, S. F. (2013). Chimpanzees play the ultimatum game. *Proc. Natl. Acad. Sci. U.S.A.* 110, 2070–2075. doi:10.1073/pnas.1220806110

Przepiorka, W., and Berger, J. (2016). The sanctioning dilemma: a quasi-experiment on social norm enforcement in the train. *Eur. Sociol. Rev.* 32, 439–451. doi:10.1093/esr/jcw014

Queller, D. C. (2000). Relatedness and the fraternal major transitions. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 355, 1647–1655. doi:10.1098/rstb.2000.0727

Rasmussen, S., Bedau, M. A., Chen, L., Deamer, D., Krakauer, D. C., Packard, N. H., et al. (2007). *Protocells: Bridging Nonliving and Living Matter*. Cambridge, MA: MIT Press.

Rawls, J. (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Richerson, P., Baldini, R., Bell, A., Demps, K., Frost, K., Hillis, V., et al. (2016). Cultural group selection follows Darwin's classic syllogism for the operation of selection. *Behav. Brain Sci.* e58, 1–71. doi:10.1017/S0140525X15000606

Richerson, P., and Boyd, R. (1998). "The evolution of human ultrasociality," in *Indoctrinability, Ideology, and Warfare*, eds I. Eibl-Eibesfeldt and F. Salter (New York, NY: Berghahn Books), 71–95.

Richerson, P. J., and Boyd, R. (1999). Complex societies: the evolutionary origins of a crude superorganism. *Hum. Nat.* 10, 253–289. doi:10.1007/s12110-999-1004-y

Ridley, M. (2010). *The Rational Optimist: How Prosperity Evolves*. London: Fourth Estate.

Rozin, P., Lowery, L., Imada, S., and Haidt, J. (1999). The CAD triad hypothesis: a mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *J. Pers. Soc. Psychol.* 76, 574–586. doi:10.1037/0022-3514.76.4.574

Ruehle, S., Rey, A. A., Remmers, F., and Lutz, B. (2012). The endocannabinoid system in anxiety, fear memory and habituation. *J. Psychopharmacol.* 26, 23–39. doi:10.1177/0269881111408958

Sanday, P. (1973). Toward a theory of the status of women. *Am. Anthropol.* 75, 1682–1700. doi:10.1525/aa.1973.75.5.02a00300

Schatzki, T. R. (2003). A new societist social ontology. *Philos. Soc. Sci.* 33, 174–202. doi:10.1177/0048393103033002002

Schmidt, M. F. H., and Rakoczy, H. (in press). "On the uniqueness of human normative attitudes," in *The Normative Animal? On the Anthropological Significance of Social, Moral and Linguistic Norms*, eds K. Bayertz and N. Roughley (Oxford: Oxford University Press).

Seeley, T. D. (1989). The honey bee colony as a super-organism. *Am. Sci.* 77, 546–553.

Shweder, R. A., Much, N. C., Mahapatra, M., and Park, L. T. (1997). "The "Big Three" of morality (autonomy, community, divinity) and the "Big Three" explanations of suffering," in *Morality and Health*, eds A. Brandt and P. Rozin (New York, NY: Routledge), 119–169.

Smith, A. (1904). *An Inquiry into the Nature and Causes of the Wealth of Nations*, 5th Edn. London: Methuen and Co., Ltd.

Spencer, H. (1896). *Principles of Sociology*. Westport, CT: Greenwood Press.

Stearns, S. C. (2007). Are we stalled part way through a major evolutionary transition from individual to group? *Evolution* 61, 2275–2280. doi:10.1111/j.1558-5646.2007.00202.x

Steinberg, L. (2007). Risk taking in adolescence: new perspectives from brain and behavioral science. *Currr. Dir. Psychol. Sci.* 16, 55–59. doi:10.1111/j.1467-8721.2007.00475.x

Sumpter, D. J. T. (2010). *Collective Animal Behavior*. Princeton, NJ: Princeton University Press.

Sunstein, C. R. (2005). Moral heuristics. *Brain Behav. Sci.* 28, 531–573. doi:10.1017/S0140525X05000099

Szathmáry, E. (2015). Toward major evolutionary transitions theory 2.0. *Proc. Natl. Acad. Sci. U.S.A.* 112, 10104–10111. doi:10.1073/pnas.1421398112

Talbot, C. F., Price, S. A., and Brosnan, S. F. (2016). "Inequity responses in non-human animals," in *Handbook of Social Justice Theory and Research*, eds C. Sabbagh and M. Schmitt (New York, NY: Springer), 387–403.

Thorne, B. L., Breisch, N. L., and Muscedere, M. L. (2003). Evolution of eusociality and the soldier caste in termites: influence of intraspecific competition and accelerated inheritance. *Proc. Natl. Acad. Sci. U.S.A.* 100, 12808–12813. doi:10.1073/pnas.2133530100

Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., and Herrmann, E. (2012). Two key steps in the evolution of human cooperation. *Curr. Anthropol.* 53, 673–692. doi:10.1086/668207

Tomasello, M., and Vaish, A. (2013). Origins of human cooperation and morality. *Annu. Rev. Psychol.* 64, 231–255. doi:10.1146/annurev-psych-113011-143812

Traulsen, A., Röhl, T., and Milinski, M. (2012). An economic experiment reveals that humans prefer pool punishment to maintain the commons. *Proc. R Soc. B Biol. Sci.* 279, 3716–3721. doi:10.1098/rspb.2012.0937

Tse, P. U. (2008). "Symbolic thought and the evolution of human morality," in *Moral Psychology*, ed. W. Sinnott-Armstrong (Cambridge, MA: MIT Press), 269–297.

Tsusaka, T. W., Kajisa, K., Pede, V. O., and Aoyagi, K. (2015). Neighborhood effects and social behavior: the case of irrigated and rainfed farmers in Bohol, the Philippines. *J. Econ. Behav. Organ.* 118, 227–246. doi:10.1016/j.jebo.2015.04.022

Turner, J. S. (2002). *The Extended Organism: The Physiology of Animal-Built Structures*. Cambridge, MA: Harvard University Press.

von Bertalanffy, L. (1968). *General System Theory*. New York, NY: George Braziller.

von Neumann, J. (1966). *The Theory of Self-Reproducing Automata*. Urbana, IL: University of Illinois Press.

Waterman, A. S. (1985). "Identity in the context of adolescent psychology," in *Identity in Adolescence: Processes and Contents*, ed. A. S. Waterman (San Francisco, CA: Jossey-Bass), 5–24.

West, S. A., Fisher, R. M., Gardner, A., and Kiers, E. T. (2015). Major evolutionary transitions in individuality. *Proc. Natl. Acad. Sci. U.S.A.* 112, 10112–10119. doi:10.1073/pnas.1421402112

Wheeler, W. M. (1911). The ant-colony as an organism. *J. Morphol.* 22, 307–325. doi:10.1002/jmor.1050220206

Williams, J. M., Oehlert, G. W., Carlis, J. V., and Pusey, A. E. (2004). Why do male chimpanzees defend a group range? *Anim. Behav.* 68, 523–532. doi:10.1016/j.anbehav.2003.09.015

Wilson, D., and Sober, E. (1989). Reviving the superorganism. *J. Theor. Biol.* 136, 337–356. doi:10.1016/S0022-5193(89)80169-9

Wilson, D. S., Vugt, M. V., and O'Gorman, R. (2007). Multilevel selection theory and major evolutionary transitions: implications for psychological science. *Curr. Dir. Psychol. Sci.* 17, 6–9. doi:10.1111/j.1467-8721.2008.00538.x

Wilson, E. O. (2012). *The Social Conquest of Earth*. New York, NY: Norton.

Wilson, E. O., and Hölldobler, B. (2005). Eusociality: origin and consequences. *Proc. Natl. Acad. Sci. U.S.A.* 102, 13367–13371. doi:10.1073/pnas.0505858102

Wirth, L. (1938). Urbanism as a way of life. *Am. J. Soc.* 44, 1–24. doi:10.1086/217913

Wrangham, R. W. (1999). Evolution of coalitionary killing. *Yearb. Phys. Anthropol.* 42, 1–30. doi:10.1002/(SICI)1096-8644(1999)110:29+<1::AID-AJPA2>3.3.CO;2-5

Yoeli, E., Hoffman, M., Rand, D. G., and Nowak, M. A. (2013). Powering up with indirect reciprocity in a large-scale field experiment. *Proc. Natl. Acad. Sci. U.S.A.* 110(Suppl. 2), 10424–10429. doi:10.1073/pnas.1301210110

Zeitzen, M. K. (2008). *Polygamy: A Cross-Cultural Analysis*. Oxford: Berg.

Zhang, B., Li, C., Silva, H., Bednarik, P., and Sigmund, K. (2014). The evolution of sanctioning institutions: an experimental approach to the social contract. *Exp. Econ.* 2, 285–303.