

Cooperation in large-scale human societies—What, if anything, makes it unique, and how did it evolve?

Simon T. Powers¹  | Carel P. van Schaik² | Laurent Lehmann³

¹School of Computing, Edinburgh Napier University, Edinburgh, UK

²Department of Anthropology, University of Zürich, Zürich, Switzerland

³Department of Ecology & Evolution, University of Lausanne, Lausanne, Switzerland

Correspondence

Simon T. Powers, School of Computing, Edinburgh Napier University, Edinburgh, EH10 5DT, UK.

Email: s.powers@napier.ac.uk

Abstract

To resolve the major controversy about why prosocial behaviors persist in large-scale human societies, we propose that two questions need to be answered. First, how do social interactions in small-scale and large-scale societies differ? By reviewing the exchange and collective-action dilemmas in both small-scale and large-scale societies, we show they are not different. Second, are individual decision-making mechanisms driven by self-interest? We extract from the literature three types of individual decision-making mechanism, which differ in their social influence and sensitivity to self-interest, to conclude that humans interacting with non-relatives are largely driven by self-interest. We then ask: what was the key mechanism that allowed prosocial behaviors to continue as societies grew? We show the key role played by new social interaction mechanisms—change in the rules of exchange and collective-action dilemmas—devised by the interacting individuals, which allow for self-interested individuals to remain prosocial as societies grow.

KEYWORDS

cooperation, cultural group selection, evolutionary psychology, human social evolution, institutions, large-scale societies

1 | INTRODUCTION

Understanding how large-scale human societies arose from small-scale ones and continue to function is a central challenge in science. It raises the question of how far the behavior of individuals in this major transition can ultimately be explained by individual self-interested motivations. These motivations are the backbone of economic models, where individuals pursue their material payoff, for example, energy, resources, time (Fudenberg & Tirole, 1991; Mas-Colell et al., 1995; North, 1990). Because material payoff tends to increase survival and reproduction it correlates with fitness (McNamara et al., 2001), and so individuals pursuing their genetic interests should tend to be payoff self-interested in interactions with non-relatives. This provides an ultimate evolutionary explanation (Alexander, 1979; Alexander, 2014; Davies et al., 2012) for the self-interested

motivations of economic models. Understanding how far human behavior in large-scale societies can be explained by self-interested motivations is crucial to improving our ability to engineer solutions to societal challenges, from climate change to genocides (Alexander, 2014; Milinski et al., 2008; Tavoni et al., 2011).

The behaviors that allow human societies to function and cohere are fundamentally prosocial in nature, by which we broadly mean behaviors that benefit the interacting group and thus other individuals. This is exemplified by the exchange of resources between individuals, and contribution to joint activities such as the construction of public goods. Unlike any other species, humans today rely on exchange of resources for nearly all of their vital needs, from food to shelter to medical care. Individuals exchanging resources are often unrelated and unfamiliar strangers, but are nevertheless involved in massive specialization and reliance on others.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. Evolutionary Anthropology published by Wiley Periodicals LLC.

The human species has spent most of its existence living in small-scale hunter-gatherer societies (Boehm, 1999; Kelly, 2013; Marlowe, 2005). There are also many prosocial behaviors in these societies, from food sharing to cooperative hunting and the construction of dams (Hooper et al., 2015; Jaeggi & Gurven, 2013; Kaplan et al., 2009). But this cooperation occurs in small groups where individuals are either kin, or personally know each other directly, or indirectly by word of mouth from other group members. It is well established that kin selection and direct and indirect reciprocity can explain the evolution of prosocial behaviors under these conditions (Alexander, 1979; Alexander, 2014). Following the origin of agriculture around 10,000 years ago, humans started to live in larger and larger societies, eventually culminating in the modern states of today. These larger groups are only functioning because of the exchange and collective action occurring between their constitutive individuals.

But why would individuals keep expressing prosocial behaviors in societies larger than the small-scale? Do the same motivations that were selected for in small-scale societies, and which are ultimately based on genetic self-interest, provide a sufficient explanation for the presence of human prosociality in large-scale societies (Alexander, 1979; Alexander, 2014; Pinker, 2015; Tooby & Cosmides, 2016), or is a different and novel explanation needed? (Richerson et al., 2016; Richerson & Boyd, 2005; Turchin, 2015) In other words, can prosocial behavior in large-scale societies be an equilibrium among self-interested individuals (Greif, 2006; North, 1990; Powers et al., 2016), or are individuals no longer acting in their own self-interest? Individuals would not be acting in their own self-interest if prosociality relies, for instance, extensively on acts of *altruistic* punishment towards non-relatives (Boyd et al., 2003; Henrich & Boyd, 2001), as in the view of “moralistic” punishment (Richerson & Boyd, 2005: p. 200). Individuals would also not necessarily be acting in their self-interest if their prosocial behaviors need to be maintained by social learning rules that can be insensitive to the actor’s material self-interest, such as conformist (Henrich & Boyd, 2001) or prestige-biased (Henrich et al., 2015) social learning.

Despite much debate, there has been little movement towards a resolution of these questions, and hence little progress in understanding the key factors that allow prosocial behaviors to be maintained in large-scale societies. To resolve this impasse, we suggest that two mechanisms need to be disentangled theoretically and then analyzed empirically. First, there is a need to determine whether and how the interactions between individuals in large-scale societies are different from those in small-scale societies. Are *social interaction mechanisms*—the constraint relationships between individual behavior and material outcomes (Figure 1)—in large-scale societies *qualitatively* different from those in small-scale societies, or is the scaling up of interaction mechanisms merely *quantitative*, and so does not involve any fundamentally new kind of behavior? Second, and regardless of the scale of society, there is a need to determine whether the *decision-making mechanisms*—the constraint relationships between cognitive state and behavior (Figure 1)—used by individuals in interactions are driven solely, partly, or minimally by the material incentives to individuals. In other words, to what extent is decision-making

dependent on the individual rewards of prosociality? Because the decision-making mechanisms are likely to be complex cognitive traits with a genetic basis, they are biologically evolved. They are thus the outcome of an *evolutionary mechanism* transforming the genetic composition of the population in successive generations (Figure 1). Consequently, they are unlikely to have changed significantly during the few thousand years since the origin of large-scale societies. This means that the same broad evolved decision-making mechanisms are presumably used by individuals in both small-scale and large-scale societies. However, different models addressing interactions between individuals in large-scale societies have made very different assumptions about the decision-making mechanisms that individuals use, and hence very different assumptions about sensitivity to incentives and self-interest.

This paper aims to contribute to the theoretical foundations of human evolution by demonstrating the conceptual clarification that can be gained by disentangling social interaction mechanisms from decision-making mechanisms to better understand the transition to large-scale societies. We do so by addressing these mechanisms in a stylized way in three steps. (1) We start by providing an operational definition of social interaction mechanisms that allows us to delineate different types of prosocial behaviors, and we synthesize what the empirical literature tells us about the similarities and differences between prosociality across human societal scales. (2) We then describe three broad decision-making mechanisms that have been widely used to describe human behavior, and that make different assumptions about individual self-interest. We synthesize what the empirical literature tells us about sensitivity to incentives, and use this to evaluate the fit of the different decision-making mechanisms to observed human behavior. (3) Finally, we ask which of the following two factors explains the evolution of large-scale human societies. Is a special decision-making mechanism, where choice of behavior is less sensitive to incentives, necessary to allow prosocial traits to spread in large-scale societies? This is postulated by the cultural group selection hypothesis (Richerson et al., 2016; Richerson & Boyd, 2005). Or alternatively, did groups create new rules that changed their social interaction mechanisms, such that prosocial behaviors are still favored by self-interested individuals that choose their behaviors largely according to reproductive compatible incentives? This is postulated by the institutional path hypothesis (Powers et al., 2016; Powers & Lehmann, 2017).

2 | SOCIAL INTERACTION MECHANISMS IN SMALL-SCALE VERSUS LARGE-SCALE SOCIETIES

Over the past 2 million years, our ancestors developed the lifestyle of nomadic hunting and gathering, and lived in small-scale societies. Following sedentarization and the subsequent Neolithic Demographic Transition around 10,000 years ago (Bocquet-Appel, 2011), large-scale societies arose. These societies are large in terms of number of individuals, and tend to have hierarchical organization, that is,

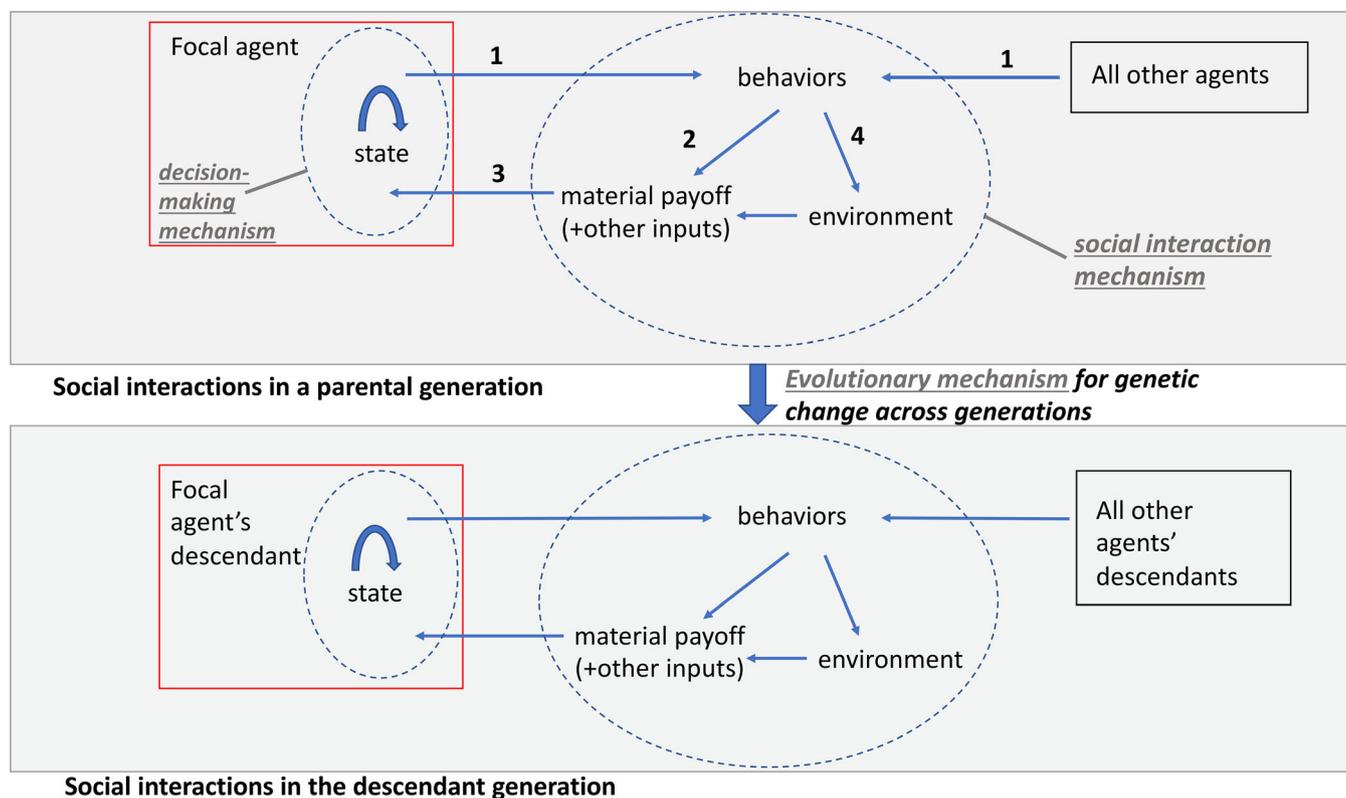


FIGURE 1 Three mechanisms describing a population of interacting agents in a parental and descendant generation. 1. The *decision-making mechanism* describes an individual as an open system exchanging matter, energy, and information with its surroundings, where the internal state of the individual and the inputs from the outside world transform into new internal states and behavioral output. This follows standard animal behavior. 2. The *social interaction mechanism* describes how the *behaviors* of all individuals in the population transform into outcomes, given the current exogenous and endogenous constraints of the *environment* in which the population resides. This follows standard game theory. The decision-making mechanism together with the social-interaction mechanism—the whole rectangle in a given generation—defines the *social interaction*. 3. The outcomes of the social interaction mechanism in turn feedback as inputs into the decision-making mechanism. 4. The behavior output by the decision-making mechanism can affect the relationship between behaviors and payoff through endogenous modification of the environment, which underlies the idea of *institutions* (Box 1). The *evolutionary mechanism* describes how the frequency of genes affecting the decision-making mechanism changes between generations, and which is ultimately induced by the forces of natural selection, mutation, and drift (Alexander, 1979). We note that the environment could also change exogenously between generations, and this would follow the laws of physics. The figure provides an exhaustive description of an evolving population of interacting individuals. Importantly, many rounds of interactions can occur within a generation, hence in Figure 1 there is an explicit separation of time scale between behavioral time (a single gray box) and demographic time (changing from the upper gray box to the lower gray box). If there has not been time for evolution to change gene frequencies during the Neolithic, then the evolutionary mechanism no longer has any effect and only the decision-making and social interaction mechanisms are operating

chiefdoms and states (Earle, 1997; Johnson & Earle, 2000). To discuss the differences and similarities of prosocial behavior in small-scale and large-scale societies, we start by introducing a model of social interaction that underpins our analysis and that is independent of both societal scale and of decision-making mechanisms. We thus now zoom in on the social interaction mechanism in Figure 1.

2.1 | The rules of the game and changes of these rules

We take our model of social interaction between individuals from game theory (Fudenberg & Tirole, 1991; Hurwicz, 1996; Mas-Colell et al., 1995), as this provides a common descriptive model that is fully

endorsed across evolutionary biology, the social and the computational sciences, and over which there is no debate. Rather, debates are about decision-making mechanisms, which we discuss in the section “Three broad decision-making mechanisms.” We consider that the individuals in a group can express feasible alternative behaviors (see Glossary, Table 1). A combination of behaviors, one for each individual in a group, determines a behavioral profile. We assume that to each behavioral profile there is an associated material outcome for each individual. This relationship between behavioral profiles and outcomes, that is, the rule assigning consequences to behaviors, is what we call a *social interaction mechanism* (dotted circle in Figure 1). This is formally known as a *game form* in game theory (Fudenberg & Tirole, 1991; Hurwicz, 1996), whereas a “game” consists of this together with individual preferences over material outcomes.

TABLE 1 Glossary

Term	Description
Behavior	Single action or stream of actions, where an action is an elementary behavioral unit.
Prosocial behavior	Behavior that results in (a) an increase in the payoff to recipients (other group members) and (b) a decrease in the actor's payoff relative to that of recipients. A prosocial behavior so defined involves a social dilemma and can be <i>altruistic</i> or <i>cooperative</i> (since it can still result in a net increase in the actor's payoff).
Social interaction mechanism	Constraint relationships between individual behavior in a group and outcomes. This specifies all the behavioral options available to individuals and the concomitant material outcomes to each individual, given the behavior of other individuals. The constraints can be either exogenous (e.g., the law of physics) or endogenous.
Decision-making mechanism	Constraint relationships between cognitive state and behavior. The cognitive state of an individual itself depends on past (internal) state and the environmental input. Hence, a decision-making mechanism is an input-output system.
Game form	Technical term used in game theory for a social interaction mechanism. Colloquially, this is called "the rules of the game."
Institution	Two-level social interaction mechanism. The outcome of the first level (the "political" interaction) yields the rules of the interaction leading to material outcomes (or the "economic" interaction, or "institutional rules"). This subsumes the idea that the rules of this interaction are humanly devised, arising either "spontaneously" or by deliberate design.
Payoff	One dimensional numerical value associated to outcomes of behavior under a social interaction mechanism. Payoffs are always considered material in this paper and usually involve multi-shot interactions (or a stream of behaviors).
Altruistic behavior	Behavior that results in a net decrease in the actor's payoff (and survival and reproduction) and a net increase in the payoff to recipients.
Cooperative behavior	Behavior that results in a net increase in the actor's payoff and an increase in the payoff of recipients.

Colloquially a social interaction mechanism is thus the "rules of the game," which specify the relationship between behavior and outcomes. Material outcomes are usually multidimensional, for example, calorie intake, nutrient type, or size of shelter. But we assume that all such outcomes can be summarized by a single number, the material payoff to an individual, from which average group material payoff can also be evaluated. Material payoff is therefore a unifying currency,

and has been useful in capturing incentives in the social sciences (Greif, 2006), but also in evolutionary biology because material payoff is expected to correlate with reproduction and survival, so it is a workable and useful proxy of fitness (McNamara et al., 2001). And even if behavior may be guided by other incentives than material payoff, these often correlate with it: status and reputation may correlate with power, which may correlate with wealth, which may in turn correlate with access to material resources (Alexander, 1990). Hence our focus on material payoff in this paper.

Finally, by a *social interaction* (gray outer rectangle in Figure 1), we mean a social interaction mechanism augmented with a decision-making mechanism (red inner rectangle in Figure 1), as together this specifies both the rules of the interaction and the cognition leading to behavioral choice within these rules (and thus subsumes the notion of "game" from game theory). For now we make no assumption about the nature of the cognition that leads to behavior; in the forthcoming section "Three broad decision-making mechanisms" we will specify concrete alternative decision-making mechanisms that make different assumptions about sensitivity to payoff. We focus on social interaction mechanisms involving a social dilemma, so that the interaction faced by individuals involves a tension between individual and group material payoff (Kollock, 1998). We will specifically refer to the behavior of an individual as *prosocial* if it (a) not only increases the average payoff to interacting group neighbors but also (b) decreases the individual's own payoff relative to the average among those group neighbors. In other words, a prosocial behavior so defined increases the payoff to others more than to self. This means that an individual has a larger payoff if everyone else performs the prosocial act apart from itself, hence the "dilemma," which is ubiquitous in human interactions (Kollock, 1998). However, a prosocial behavior so defined does not say whether the behavior results in a net decrease or increase in the actor's own payoff. In order to be able to distinguish between these two cases, we call a prosocial behavior *altruistic* if it results in a net decrease in the actor's payoff, and assume that this then also decreases reproduction and/or survival. By contrast, a prosocial behavior will be called *cooperative* if it increases the net payoff of both actor and recipients, and hence mutually increases reproduction and/or survival. As such, our definition of altruism and cooperation follows those of evolutionary biology (Bshary & Bergmüller, 2007; Hamilton, 1964) yet our currency of interest is material payoff rather than individual fitness, and these two behavioral categories have been applied to different fitness relevant currencies before (Bshary & Bergmüller, 2007).

A striking observation of human social interactions is that individuals can change the rules of the game, that is, individuals can change their social interaction mechanism (Greif, 2006; Hayek, 1988; Hurwicz, 1996; Kaplan et al., 2005; North, 1990; Ostrom, 1990), by devising constraints on interactions. This can occur consciously by deliberate design, or unconsciously as a result of trial-and-error or lucky accidents. The constraints can be informal, such as customs, taboos or norms, or they can be formal, such as laws and constitutions (North, 1990). They may affect the behaviors available to individuals in the interaction, or the relationship between behaviors and material

payoff. An *institution* is defined as a set of social interaction mechanisms that individuals can potentially choose between. This set of social interaction mechanisms that individuals can possibly implement is necessarily constrained by their physical environment and the current state of their technology (Hurwicz, 1996). Colloquially, this can be thought of as a set of possible “rules of the game,” or “institutional rules,” that individuals can choose between. The hallmark of an institution is the presence of two types of social interaction (Powers et al., 2016) p. 3: (i) active genesis of institutional rules through trial-and-error, communication and/or bargaining by the individuals in a group—this is a political interaction; (ii) social interactions whose outcomes are material and which are affected by the institutional rules from (i)—this is the economic interaction. The key idea behind an institution is the formalization of the point that humans, unlike other animals, can deliberately self-modify the material-payoff structure (rules) of their social interactions, and hence they “play” political as well as economic games (see also Box 1). The political interaction generates the rules defining the economic interaction, and may potentially involve many informal and spontaneous communication events between subsets of the group of interacting individuals.

2.2 | Two broad types of social dilemmas

Any prosocial behavior in the economic interaction can be placed on a scale representing the excludability of the economic good that it involves. At the one end of this scale is the voluntary exchange of private goods between individuals. These are goods that the actor controls or otherwise has property rights over, meaning that other individuals can readily be excluded from using them. Exchange of private goods allows individuals to obtain resources that they want but do not currently have. It also allows gains in efficiency from division of labor and specialization (Greif, 2006; Mas-Colell et al., 1995; North, 1990). But it is not obvious that individuals will choose to engage in exchange, for two reasons. First, one individual must part with its goods before it receives anything in return. This means that

an individual risks being cheated and receiving nothing in return (Greif, 2000). Second, the individual offering a good inherently knows more about its quality than the receiver, creating an information asymmetry that can be used to exploit the other party (North, 1990). This means that exchange involves a social dilemma.

At the other end of the excludability scale are collective-action forms of economic interaction, which involve the production and consumption of public goods such as village fortifications, or common-pool resources such as fish stocks, grazing land, or irrigation water (Ostrom, 1990). It is costly to monitor the behavior of individuals producing and using this type of good, and hence to exclude those that do not contribute to it. This means that there is a social dilemma because of the temptation to free-ride and enjoy the benefits without expressing the behaviors underpinning their production (Olson, 1965). In the following discussion we categorize prosocial behaviors based upon whether they are closer to private exchange or public goods/common-pool resources on the excludability scale.

2.3 | The world until yesterday: Cooperation in small-scale societies

2.3.1 | Exchange

Hunter-gatherers engage in several types of exchange between individuals. The most well studied is the exchange of meat between large-game hunters, that is, food sharing. Male hunters selectively donate food to other hunters in their group when they make a kill, and in turn receive food when the latter are successful. The exchange occurs repeatedly—essentially for an indefinite number of times—between members of the camp, which would number around 30 individuals (Kelly, 2013; Marlowe, 2005). The exchange of meat is *personal*, that is, between individuals that know each other. People obtain information about the behavior of other specific individuals in their group. This information is either direct—they remember exactly who they have given food to in the past, or is indirect—they obtain verbal

BOX 1 The concept of institutions

The conception of institution described in the main text follows closely that of Hurwicz (1996: p. 127) (see also Powers et al., 2016) who discusses relationships with the perhaps more widespread conception of institution by North (1990) as humanly devised rules of the game. In addition, we note that the term institution is often used to refer to either organizations or to social norms or even used as a buzzword that everybody understands differently. But considering an institution as a two-level social interaction mechanism (Hurwicz, 1996; Powers et al., 2016), where individuals affect the choice of the rules of their economic interaction, makes explicit how the rules of the game can potentially change over an individual's lifetime as a result of the individual's actions, and disentangles “politics” from “economics.” This is conceptually useful, at least from an evolutionary biology perspective, because it makes explicit that human behavior is structured by rules that are themselves humanly-devised and not exogenously imposed, which thus requires a process to devise the rules of interactions. Further, the separation between “politics” and “economics” pinpoints a possible key difference between humans and other animals. In other animals, no distinctive example of systems of interaction for changing the rules of social interactions in behavioral time has been observed so far.

reports of an individual's behavior from others. This is supported by systems of rules that regulate the conditions under which individuals should give food to others, and which apply to the whole group (Kaplan et al., 2009). Group members enforce these rules upon each other through a variety of sanctions ranging from gossip and public ridicule through to shunning and ostracism (Boehm, 1999). Hunter-gatherers also exchange one type of commodity for another. A sexual division of labor is evident, particularly between men that specialize in hunting and women that specialize in gathering plant materials (Marlowe, 2007). Among horticulturalists, we see the exchange of horticultural produce for meat, and the exchange of childcare for labor and sick care (Jaeggi et al., 2016).

2.3.2 | Collective action

Hunter-gatherers also engage in a variety of collective actions related to subsistence. A prime case is the so-called “cooperative” hunting, where the actions of several individuals are necessary to prevent a prey animal from escaping (Alvard & Nolin, 2002; Kaplan et al., 2009). Hunter-gatherers also engage in various collective construction projects, such as burning habitat, and building dams to trap fish (Kaplan et al., 2009). Because the number of individuals taking part in the collective action is relatively small, the payoff benefits are immediately and directly felt by the participants, including the actor. In a small group of five hunters, if one individual does not pull its weight then it will directly feel the payoff impact through a markedly reduced probability of catching prey. This means that the prosocial acts tend to be cooperative rather than altruistic. The benefits to an individual of an act of prosociality of this kind are also returned with a small time delay, that is, on the order of hours to weeks. The cost that an individual pays to receive this benefit is measured in terms of the opportunity cost of time and labor invested, or the direct contribution of material resources. Humanly devised rules regulate how exactly the benefits of collective action are distributed. For example, in the !Kung Bushmen, the owner of the first arrow that penetrates the animal controls distribution after a cooperative hunt (Testart, 1987).

2.4 | The world today: Cooperation in large-scale societies

2.4.1 | Exchange

In large-scale societies, we see the specialization and division of labor that already existed in hunter-gatherer societies become much more pronounced and no longer based on gender. Individuals now largely specialize in one occupation (Demps & Winterhalder, 2019), and obtain essentially all of their vital resources through exchange with others. And this exchange is impersonal—it is often with unfamiliar strangers who may never meet again (Greif, 2006; North, 1990). But crucially, institutional rules of the exchange game have changed to account for this (Greif, 2006; Greif et al., 1994; Hayek, 1988; Milgrom

et al., 1990; Vanberg, 1994). These include both private order (enforced without the state) and public order (enforced by the state) social interaction mechanisms.

For instance, face-to-face repeated interactions have been replaced with repeated interactions with the same “entity” or player with which trust can be built up over time in essentially the same way as observed in hunter-gatherers, by relying on first-hand information. Examples are repeated interactions with the same vendor or company. Other situations may involve no first-hand information about how exchange partners have behaved in the past but trust may nevertheless be possible between agents. For example, the Law Merchant was a system of private order courts in medieval Europe that used reputation to incentivize honest behavior between merchants in long-distance trade. The rules of the Law Merchant game were such that it was beneficial for self-interested individuals to register occurrences of cheating with the courts, and to pay to query the courts to find out whether potential partners had any disputes against them, even though the court had no coercive authority from the state to sanction traders (Milgrom et al., 1990). In modern societies, credit reference agencies act as an elaboration of this kind of interaction mechanism, with state enforced public order sanctioning of individuals that default on credit agreements being a final resort. Other private order interaction mechanisms in large-scale societies work by using reputation to facilitate partner choice, as is widely used in online auction sites such as eBay (Houser & Wooders, 2006). This is essentially an elaboration of the spreading of reputation by gossip seen in hunter-gatherers. In fact, any decentralized exchange system—a market—is always a system of social interaction characterized by a specific set of rules defining certain restrictions on the behavior of the market participants, whether these rules are enforced by private orders or the state (Vanberg, 1994).

2.4.2 | Collective action

Large-scale societies also engage in numerous collective actions, from building roads and fortifications through to the use of irrigation systems and fishing waters. These goods are produced and used by many more individuals, which means that the effect that any one individual feels as a result of its own effort will be negligible. The result of the collective action can be delayed by a very long time, often on the order of years. As such, there are many more temptations to not act prosocially than in small-scale societies. But crucially, the institutional rules regulating collective action have also changed, and include again private and public order mechanisms. For example, online communities from Wikipedia through to Minecraft game servers contain hundreds or thousands of individuals that are unrelated and that do not personally know each other, but that must contribute through their actions to the maintenance of the community. In the communities that persist over time, institutional rules prescribing accepted behaviors, and including sanctions for their violation, regulate behavior without state enforcement (Frey & Sumner, 2019). On the other hand, public order mechanisms include collecting contributions from

taxes, whose payment can be easily monitored. The ability to permanently levy tax and monopolize the use of violence for sanctioning is actually a defining feature of the state (Hoffman, 2015), whose public order social interaction mechanisms are based on formal rules of laws and constitutions, enforced by third-party sanctioning.

2.5 | Comparing cooperation across scales

As we move from small-scale to large-scale societies, a change in societal structure must occur to account for the change in size (Bonner, 2011). A much higher degree of specialization and division of labor is observed in large-scale societies (North, 1990), a feature predicted by the size-complexity rule: bigger social units have greater division of labor (Bonner, 2011). At the same time, the exchange upon which this division of labor depends becomes more impersonal, with individuals less likely to have first-hand knowledge about the past behavior of their exchange partners. In collective action, we see the number of participants become so large that the marginal effect of any one individual's contribution is negligible, and the delay between making one's contribution and harvesting the benefits can be quite long.

Qualitatively, though, both small-scale and large-scale societies face the same *types* of exchange and collective-action problems. The organizational problems, however, become much more difficult in large-scale societies, and as we increase scale, we see more institutional rules that spread out into new domains such as long-distance trade and large-scale construction projects. The increase in the number of rules with the scale of a society is striking. For example, the small-scale Kapauku Papuan society has around 120 rules regulating areas from property rights through to punishment for murder, whereas 40,000 new laws took effect in the United States in 2014 alone (Singh et al., 2017). From this we can infer that as societies of any scale engage in new economic activities, the number of institutional rules that the society generates increases. It is no surprise that economists have long emphasized and recognized the fundamental role played by humanly devised rules of the game in structuring human interactions (Greif, 2006; Hayek, 1988; North, 1990).

3 | DECISION-MAKING MECHANISMS AND THE MAINTENANCE OF COOPERATION IN SOCIETIES OF ANY SCALE

3.1 | Three broad decision-making mechanisms (“minds”)

So far, we focused on the social interaction mechanisms individuals face in small-scale and large-scale societies, and how these are at least in part humanly devised. But as Figure 1 shows, in itself this description does not specify how or why individuals choose cooperative and/or altruistic prosocial behaviors in these interactions. To address

this, we now zoom in on the decision-making mechanism in Figure 1 and present three main types of decision-making mechanisms that have been proposed to explain the behavior of individuals. We refer to these decision-making mechanisms as types of agents or “minds.” For each hypothesized agent type, we indicate what assumptions it makes about how individuals choose behavior in social interactions, and the conditions that must hold for prosocial behavior to be stable given this agent type.

(1) **The Rational Strategizing Mind (hereafter RSM).** Broadly, individuals are assumed to have free choice of behaviors that are guided by a striving to maximize their own payoff, given that other individuals are also exhibiting payoff-maximizing behavior. This is the standard model of human behavior assumed in economics: the “rational man” (Fudenberg & Tirole, 1991; Mas-Colell et al., 1995; North, 1990). In this paper, we take the “payoff” as being the agent's own material payoff, and so we conceive the RSM agent as being self-interested. Basing behavior choice on their own material payoff is how RSM agents are traditionally used to make predictions about human behavior in exchange and collective-action situations (Fukuyama, 2011; North, 1990).

There is a whole range of RSMs depending on how they react to payoff, ranging from fully forward-looking agents that strive to maximize long-term payoff, to boundedly-rational agents that have to make a decision under constraints of limited time and information processing, to myopic agents that choose actions that maximize short-term payoff. But all are self-interested. Hence, an RSM agent can only express a cooperative prosocial behavior, and not an altruistic behavior. For cooperative behaviors to be maintained in exchange and collective-action situations by RSMs, the social interaction must either result in a net benefit in single-move games, or involve repeated interactions with known or unknown individuals (Fudenberg & Tirole, 1991), or the behaviors must be exogenously enforced. Both the conditioning of behaviors on the past behavior of group members, and exogenous enforcement of property rights and contracts, can create incentives for cooperative behavior in large-scale societies (Binmore, 2005a; Fudenberg & Tirole, 1991; Milgrom et al., 1990). It is important to realize that this can be the case regardless of societal scale (Milgrom et al., 1990).

(2) **The Fitness Maximizing Mind (hereafter FMM).** Broadly, individuals are assumed to express behavioral rules that serve, over their lifetime, their genetic interests. This is the standard model of human behavior of evolutionary biology, where individuals are expected to treat interaction partners according to their degree of genetic relationship towards them (Alexander, 1979; Alexander, 2014), and should thus appear to behave as if they strive to maximize a measure of inclusive fitness. The behavioral rules of an FMM agent are evolutionarily acquired. Since genetic relatedness will be greater than zero in small groups under limited genetic mixing, an FMM agent may express both cooperative and altruistic prosocial behaviors, depending on the relatedness to the interaction partner(s). Because reproduction and survival correlate with material payoff, an FMM agent is expected to evolve to be payoff-sensitive although this depends on the exact nature of the agent type, which we discuss next.

There is a range of conceivable FMM agent types that vary in how they are payoff-sensitive. One could make no more assumptions than stipulating that fitness relevant payoff guides behavior in any circumstances, no matter how novel the circumstance. This is the approach taken by *human behavioral ecology* (Nettle et al., 2013). In this case, FMM, would behave quite like an RSM agent, albeit taking relatedness with the recipient(s) into account. Alternatively, one can assume that FMM agents are designed to solve specifically the survival and reproductive puzzles of hunter-gatherer societies. This is the approach taken by *evolutionary psychology* (Barkow et al., 1992). We call this subtype of FMM agent the “Pleistocene Adapted Mind” (or PAM), since it is assumed to behave according to domain-specific decision-making algorithms, specialized to solve particular adaptive problems of the Pleistocene, for example, language acquisition, mate selection, or cooperative exchange. These problems are posited to make up the Environment of Evolutionary Adaptedness, or EEA, for human exchange and collective-action behaviors (Barkow et al., 1992; Cosmides & Tooby, 2013). The algorithms may do varying amounts of computation, ranging from a complex assessment down to the use of simple heuristic rules (Gigerenzer & Brighton, 2009). In this perspective, the EEA for exchange and collective-action in small-scale societies would have selected for a *payoff-sensitive* psychology that initiates and monitors reciprocal exchanges, including specialized algorithms for detecting cheaters and calculating the probability that an exchange partner will reciprocate (Cosmides & Tooby, 1992). A PAM agent would cooperate in large-scale societies whenever these algorithms were activated with inputs that resembled situations where it would have been incentivized to cooperate in their EEA.

(3) **The Social Learning Mind (SLM).** Broadly, individuals are assumed to acquire their prosocial behavior mostly from others through various forms of social learning. Hence, behavior is mostly transmitted between individuals, and not freely chosen by individuals as with RSM agents. SLM is the standard model of decision-making assumed in much of the literature on cultural evolution of prosocial behavior (Richerson et al., 2016; Richerson & Boyd, 2005).

There is a range of conceivable SLM agents that vary in their payoff-sensitivity, according to the bias by which it chooses other individuals to learn from. Three key biases have been proposed in the cultural evolution literature (Richerson et al., 2016), which vary in their payoff-sensitivity from high-to-low as follows: 1. payoff bias; 2. prestige bias; 3. conformity bias. SLMs using a payoff bias copy a prosocial behavior if they observe other individuals doing the prosocial behavior to be receiving a higher payoff than the population average. SLMs using prestige-bias copy multiple behaviors, potentially including prosocial ones, from high status individuals who overall exhibit a high material payoff. Prestige-biased SLMs copy a range of behaviors from high status individuals to avoid the costs of calculating whether each particular behavior enhances payoff—they assess the aggregate payoff to the suite of behaviors instead. Importantly, this means they may copy prosocial behaviors that do not lead to an increase in payoff. Finally, conformity-biased SLMs make no assessment of payoff when choosing other individuals to copy from. Instead, they copy the most frequent variant of a behavior exhibited in the population.

Because SLMs are less sensitive to the material payoff of a prosocial behavior if they use prestige or conformity biases, the social interaction mechanism becomes less important in their choice of behavior. For example, under a pure conformity bias, prosocial acts may be maintained as an equilibrium behavior regardless of the social interaction mechanism. Consequently, prestige and conformity biased SLMs may exhibit altruistic behaviors that they copy from other individuals, even though these reduce their own material payoff. For these types of SLM, how the type of exchange and collective-action social interaction mechanism, and hence institutions, incentivize behavior therefore does not matter much. On the other hand, if SLMs are more payoff biased then they become more self-interested with respect to choice of prosocial behavior. The social interaction mechanism then matters more in explaining the sustainability of prosocial behaviors, and prosocial behaviors between non-relatives can no longer be altruistic.

The three agent types are summarized in Table 2. We wish to stress two points about them. First, the agents just described are *independent* of the scale of the society, and so apply to behavior choice in both small-scale and large-scale societies. Second, while each agent type covers a spectrum of variation, they are nevertheless necessarily abstract caricatures of human behavior, since they correspond to the general theoretical assumptions that researchers in different fields have made about how human agents choose behaviors. While few, if any, researchers would argue that the human mind literally functions as any of the specific agent types described, and in reality is likely to involve some mix of several of them, these caricatures are widely used in the theoretical literature to model how individuals choose behaviors. Moreover, they are used to make empirical predictions about how humans will behave. It is thus useful to draw caricatures in order to delineate clearly different hypotheses. To evaluate how well each model can explain human cooperative behavior, we therefore need to look at the weight of evidence that pulls humans towards and away from each of the agent types.

3.2 | Empirical assessment of agent types

While all three agent types may be compared on various attributes, we here compare the agent types specifically on the payoff-sensitivity

TABLE 2 Types of agent (“mind”)

Agent type	Acronym	Variants
Rational strategizing mind	RSM	Myopic or forward-looking
Fitness maximizing mind	FMM	Modules specific to the pleistocene (Pleistocene adapted mind) or domain general (human behavioral ecology)
Social learning mind	SLM	Conformist-biased, payoff-biased, prestige-biased

scale as this is the focus of our paper. We stress that this is payoff-sensitivity with respect to prosocial behaviors. We can then rank them from the most to the least payoff-sensitive, when assuming no interactions with relatives, as follows:

1. RSM and the behavioral ecology-based subtype of FMM,
2. payoff-biased SLMs,
3. the PAM subtype of FMM, and.
4. prestige and conformity biased SLMs.

RSMs, behavioral ecology-based FMMs, and payoff-biased SLMs are expected to always try to increase their material payoff given the information available to them. PAMs, on the other hand, are not expected to always increase their own material payoff; for instance, they will not when environmental cues trigger the evolved algorithms in inappropriate circumstances (Burnham & Johnson, 2005; Hagen & Hammerstein, 2006; Johnson et al., 2003; Raihani & Bshary, 2015). Prestige or conformity biased SLMs would choose prosocial behaviors that are not payoff-sensitive if these are exhibited by high status individuals, or the majority of other individuals, respectively.

Experimental economics has long demonstrated in the laboratory that individuals are markedly sensitive to material payoffs in exchanges that resemble the types of exchange problems that occur in small-scale and large-scale societies (Smith, 1962). Additional evidence for humans making payoff-sensitive decisions in these situations is provided by the fact that levels of cooperation observed in repeated Prisoner's Dilemma experiments are affected by whether the end point of the game is known (Roth & Murnighan, 1978).

A different line of research involves experiments where prosocial behaviors are *not* payoff-sensitive, particularly public goods games without incentives to act prosocially, but argues that individuals nevertheless still behave prosocially (Fehr & Gächter, 2002). This has been suggested as evidence that humans are less self-interested when choosing prosocial behaviors (Boyd et al., 2003; Fehr et al., 2002). However, as in exchange games, when individuals play the public goods game for a longer period of time then they often start to behave in a payoff-sensitive manner (Binmore, 2005b; Sefton et al., 2007). Further, analysis of multiple experiments and dissection of decision making reveals that such behavior is likely to be consistent with responsiveness to payoff (Burton-Chellew et al., 2015; Burton-Chellew et al., 2017; Ledyard, 1995; Thomas et al., 2016). This suggests that humans are not perfectly rational and do not always immediately make optimal decisions, and so are not the caricature rational agent assumed in economics. They do, however, seem to learn over time in a way that is sensitive to payoff.

Some experiments have also reported quite substantial cross-cultural variation in behavior in these games (Gerkey, 2013; Henrich et al., 2006; Herrmann et al., 2008). This can be interpreted as support for SLM, since it suggests that localized prestige or conformity biased social learning may be more important than maximizing individual material payoff in determining whether individuals behave prosocially. However, these differences could also reflect RSM or FMM agents acting in different economic environments (Baumard, 2013). Pertinent

environmental differences include the extent to which the interaction is repeated, value of long-term relationships given the institutional rules of the local market (North, 1990), and the fidelity with which reputational information is transmitted (Delton et al., 2010; Greif, 2006). Variation in these features between cultures would cause RSM agents to correspondingly vary their levels of prosociality.

We can also ask whether humans routinely perform conformity-biased social learning. Conformity is very common in children (Haun et al., 2014). Several experiments with adults, however, have demonstrated a lack of conformity, especially in situations where conforming would result in a reduction in material payoff (Burton-Chellew et al., 2015; Burton-Chellew et al., 2017; Lamba, 2014; Lamba & Mace, 2011). In general conformity is reduced if the actor being copied is not very successful at performing the task (Schillaci & Kelemen, 2014; Scofield et al., 2013), or if conformity would conflict with the actor's existing knowledge (Sobel & Kushnir, 2013). Crucially, different individuals tend to use different social learning rules (van den Berg et al., 2015).

Taken as a whole, the experimental literature demonstrates that payoff-sensitivity is a key driver of individual decision making in social interactions. Humans also seem to have mental capabilities for abstraction that allow them to create models of causality, and thus potentially conceive rules of interactions to regulate prosocial behaviors (Fukuyama, 2011), which pushes them away from prestige and conformity biased SLMs on the payoff-sensitivity scale. However, there is a pressing need for continued empirical work in both experimental and field studies to identify the conditions in which peoples' decisions are more or less payoff-sensitive in exchange and collective-action situations, and more or less conformity-biased or prestige-biased in these situations. A particular focus should be placed on the different time frames of net positive payoffs, that is, whether the benefit is immediate or delayed, and whether it is conditioned on past behavior towards the same partner, or based on reputation effects.

3.3 | Evolutionary foundation for agent types

From the perspective of evolutionary biology, the decision-making mechanism has evolved and the underlying genetic material coding for the decision-making mechanism ultimately changes according to the fundamental evolutionary forces, with selection being the principal guiding force that leads to adaptation (Alexander, 1979; Alexander, 2014) (see also Figure 1). As such, evolution would favor cognition that advances the actor's genetic interest. This is consistent with the predictions of FMM, and in particular its subtype PAM, if no massive change has occurred in the decision-making mechanism(s) since the Neolithic. The version of FMM that evolution favored could also involve strong planning and forward-looking abilities, and thus lean on RSM. Under certain circumstances evolution indeed favors literal RSM agent types maximizing long-term material payoff and taking into account when interaction occurs between relatives (Alger et al., 2020). Individuals may also evolve to have strong social learning skills (and thus lean on SLM). The extent to which individuals will

express planning and forward-looking abilities is not directly deducible from evolutionary theory, without considering specific details of human development and ancestral ecological conditions. What is generally deducible, however, is that such agents would only express altruistic behaviors towards genetically related individuals, whereas cooperative behaviors would be expressed towards unrelated individuals if the right incentives were in place. As such, any population of agents systematically expressing payoff and thus fitness incompatible altruistic prosocial behaviors would not be favored by evolution and would thus be exhibiting a biologically *maladapted* decision-making mechanism.

4 | PROSOCIALITY IN LARGE-SCALE SOCIETIES: DECISION-MAKING MECHANISMS VS SOCIAL INTERACTION MECHANISMS

As discussed in Section 2, large-scale societies are dependent on prosocial behaviors for their existence. Yet, evolutionary theory shows that, everything else being equal, the selection pressures favoring prosocial behaviors decrease drastically as the number of interacting individuals increases, and hence prosocial behavior is unlikely to be favored in large groups (Powers & Lehmann, 2017). So what was the key mechanism that allowed for prosocial behavior to be sustained in the transition from small-scale to large-scale societies? Did prosocial behavior start to go against the actor's genetic self-interest as the scale of society increased? If so, it would need to be maintained by a decision-making mechanism that is less sensitive to material payoff compared to the decision-making mechanism of evolutionary biology (FMM) described in the previous section. Or alternatively, was prosocial behavior still payoff-sensitive because of changes in the social interaction mechanism?

4.1 | Hypothesis 1: A special decision-making mechanism is the key driver of prosocial behavior in large-scale societies

The first hypothesis is the *cultural group selection hypothesis* (Richerson et al., 2016; Richerson & Boyd, 2005), which posits that the main driver maintaining the expression of prosocial behavior in the transition to large-scale societies is that humans are largely SLM agents with a high degree of prestige and conformity bias. These biases can maintain prosocial behavior *within* a group, even if the prosocial behavior is not payoff-sensitive and hence not an equilibrium behavior for self-interested agents. Prosocial acts in large-scale societies can therefore be altruistic under this hypothesis. If different groups reach different patterns of behavior, some with a greater frequency of prosocial behaviors, and some with less, then competition between groups can cause the prosocial behavior to spread throughout the population (Boyd et al., 2003; Henrich & Boyd, 2001). However, explicit formal models investigating these processes (Lehmann &

Feldman, 2008; Molleman et al., 2013; Peña et al., 2009), and holding everything else constant in comparison to payoff-biased transmission, have so far generally failed to show that conformist-biased transmission favors the spread of prosocial behaviors. Prestige-biased transmission fares better, though (Molleman et al., 2013).

Cultural group selection was first proposed to explain the transition to large-scale social societies in economics (Hayek, 1988) but without much detail as to what form of competition between groups would do the job, a point that has been expressed as follows (Sugden, 1993):

Sometimes it seems to be suggested that less successful groups imitate more successful ones; sometimes, that individuals migrate from less successful to more successful groups; sometimes, that more successful groups reproduce more rapidly; and sometimes, that more successful groups exterminate less successful ones. I think we must assume that Hayek has no particular theory of group selection clearly in mind, but has the hunch that there is some common criterion of "success" or "fitness" that would be favored by any plausible theory.

Much the same variations of group competition have been proposed in the more recent literature (Richerson et al., 2016). It has been argued that cultural group selection will be reinforced if competition between groups involves the physical displacement of less prosocial groups by their more prosocial neighbors, for example, through warfare (Bowles et al., 2003; Boyd et al., 2003; Turchin et al., 2013). This type of group competition could cause payoff-insensitive prosocial behaviors to spread if SLMs use prestige or conformity bias, including altruistic behaviors. It has also been proposed that competition between groups could take the form of individuals either migrating to more successful groups, or imitating individuals in more successful groups (Richerson et al., 2016). This could cause payoff-insensitive prosocial behaviors to spread if SLMs use prestige bias when choosing the group to migrate to or imitate. In both of these cases, the social interaction mechanism does not matter much because of the agent's decision-making mechanism.

On the other hand, both types of group competition could also function with payoff-biased SLMs. However, in this case there would need to be a mechanism of social interaction which ensures that prosocial behaviors give a higher payoff than non-prosocial behaviors *within* a single group, for example, non-altruistic forms of sanctioning (Ostrom, 1990). And with payoff-biased SLMs, prosocial behaviors could only be cooperative in large groups of genetically unrelated individuals, and not altruistic. This idea has been proposed in some "weaker" versions of the cultural group selection hypothesis, which argue that prosocial behaviors in social dilemmas are actually cooperative equilibria within a *single* group (Richerson et al., 2016). However, this is typically an assumption of cultural group selection models (Bowles et al., 2003; Boyd & Richerson, 1990), rather than the models demonstrating the evolution of a social interaction mechanism that makes

prosocial behaviors an equilibrium for self-interested individuals in a single group. In summary, there is no single theory of cultural group selection and the different variants make different assumptions on the payoff-sensitivity of individual behavior.

4.2 | Hypothesis 2: A refinement of social interaction mechanisms is the key driver of prosocial behavior in large-scale societies

The second hypothesis is the *institutional path hypothesis* (Powers et al., 2016; Powers & Lehmann, 2017), which posits that the driver maintaining the expression of prosocial behavior in the transition to large-scale societies is a refinement of social interaction mechanisms (see Glossary, Table 1); namely, people changed the rules of their economic games. Individuals are thus assumed to have formal and/or informal political interactions that affect their economic interactions. The hypothesis is that as groups grew in size, individuals have refined and created new institutional rules supporting exchange and/or have changed systems of monitoring and sanctioning to handle larger numbers of individuals in collective-action problems. Institutional rules may also have reduced the effective number of individuals that interact through the creation of nested group structures (Ostrom, 1990). These new mechanisms of social interaction (not necessarily created by “deliberate design,” see more on this in the next section) would lead to prosocial behaviors increasing material payoff to the actor, and hence can be generally favored even by self-interested individuals. Prosociality among non-relatives in large-scale societies is thus always cooperative, rather than altruistic, and so individual behavior is always payoff-sensitive under this hypothesis.

Due to this payoff-sensitivity, the institutional-path hypothesis is compatible with RSM agents, behavioral ecology-biased FMM agents, and with SLM agents that use payoff-biased social learning when choosing prosocial behavior. It is also compatible with the PAM subtype of FMM to the extent that the institutional rules recreate the conditions where cooperative prosocial behaviors were payoff-sensitive in small-scale societies, for example, effective sharing of reputational information. Moreover, PAMs would be expected to create institutional rules similar to those found in small-scale societies in circumstances that are ecologically similar (Boyer & Petersen, 2012; Petersen et al., 2013), for example, to create rules of uniform sharing in periods of high resource variance (Cosmides & Tooby, 1992).

The form of the institutional rules a group ends up with will be influenced by proximate factors such as asymmetries in power, influence, and information (Singh et al., 2017), which determine the outcome of political interactions. Furthermore, only a subset of the individuals affected by the institutional rules may take part in the political interactions, and the interests of those taking part may not be representative of the interests of the group as a whole. Consequently, conflicts of interests between segments of the group may result in institutional rules not being optimal for all group members, as exemplified by the rise of highly despotic states such as Ancient Egypt, where despotic leaders biased institutional rules in favor of

themselves. As such, the institutional path hypothesis is compatible with the widespread existence of inefficient institutions (North, 1990). On the other hand, when the interests of group members are aligned, or bargaining strengths are equal, then efficient institutions that increase average material-payoff are more likely to arise, a point that has been repeatedly stressed in the (political) economics literature (Greif, 2006; North, 1990; Ostrom, 1990).

The ability to create and enforce rules by self-interested individuals, especially over food sharing and property rights, would have been necessary to support the hunter-gatherer lifestyle (Hill, 2009). If hunter-gatherers did not have political interactions, then the institutional path hypothesis cannot explain the origin of large-scale societies. But there is evidence that hunter-gatherers do indeed have political interactions that affect their economic interactions, even though they lack the bureaucratic elements of large-scale societies. For example, when the extant Ache hunter-gatherer society transitioned from foraging to horticulture, they advocated and voted in local meetings to transfer fields from public to private ownership (Kaplan et al., 2005).

4.3 | A combination of a special decision-making mechanism and a change in social interaction mechanisms?

Since both the cultural group selection and the institutional-path hypotheses assume a quantitative scaling up of the same kinds of exchange and collective-action problems, the above analysis shows that a key question in determining the driving factors in the transition to large-scale societies is what decision-making mechanism determines the expression of prosocial behavior. If individuals are sensitive to material payoff when choosing prosocial behavior, then there *must* have been a change in their social interaction mechanisms, as proposed by the institutional-path hypothesis. Without such a change, individuals should stop acting prosocially as they took part in exchanges and collective actions with more individuals, because when everything else is constant the pressures favoring prosocial behaviors decrease rapidly as the number of interacting individuals increases. Conversely, if there was no change in the social interaction mechanisms then individuals *must* be less sensitive to payoff. If so, a special decision-making mechanism *must* operate, whereby some form of cultural group selection does the work in explaining why prosocial behaviors, be they cooperative or altruistic, are stable in large-scale societies.

Because humans undoubtedly experiment with many behaviors by trial-and-error and do considerably rely on social learning (Legare, 2017), the rules constraining behavior in economic interactions to which a society converge must to some extent at least partly be the outcome of some “spontaneous order” (Hayek, 1988; Sugden, 1993; Vanberg, 1994) and not the outcome of fully deliberate design. A case in point is the advent of the usage of money, which is a typical rule-based change in economic organization that is both in the interest of individuals using it and that is likely to have spread

gradually by payoff-biased social learning (Vanberg, 1994). As such, some ingredients of the cultural group selection hypothesis may be complementary with the institutional-path hypothesis, with competition between groups spreading different “spontaneous orders.” However, this depends critically on the exact version of “cultural group selection” that has been operating. If the version of cultural group selection involves altruistic behaviors or altruistic punishment, then it is not complementary as it assumes individuals that are not self-interested (André, 2011; Pinker, 2015).

Competition between groups resulting from warfare, differential migration, or environmentally induced extinctions acts as an equilibrium selection device (Binmore, 2005a; Boyd & Richerson, 1990; Harsanyi & Selton, 1988), favoring equilibria that lead to a higher average payoff for group members. Cultural group selection advocates traditionally stressed that high-payoff equilibria resulted from prestige and conformity biased SLMs causing behaviors to spread within groups even if they were not payoff-sensitive in the underlying social interaction mechanism, rather than being equilibria because they were payoff-sensitive (Henrich et al., 2015; Richerson & Boyd, 2005). But prosocial equilibria can also exist within groups under payoff-biased social learning, or under RSM agents that rationally choose their actions, if the right mechanisms of social interaction are in place. In this case, between-group competition can again act as an equilibrium selection device, spreading by cultural transmission mechanisms of interaction that lead to cooperation, without individuals acting against their self-interest (Binmore, 2005a; Boyd & Richerson, 1990; Harsanyi & Selton, 1988). This can act alongside the creation of mechanisms of interaction by bargaining and negotiation, helping to fill in where individuals are less than fully rational, that is, boundedly-rational RSMs or PAMs. On the other side, an explicit consideration of political interactions for changing institutional rules can complement cultural group selection models, which typically leave unspecified how a group arrives at a particular equilibrium in the first place. Much formal work remains to be done to ascertain the conditions under which such equilibrium selection processes at the level of rules of the game (instead of economic behavior under given rules) may work. There are essentially no models of this to date.

Despite ingredients of the cultural group selection and institutional path hypotheses not being necessarily mutually exclusive, there is a crucial need to understand whether the main driver of the evolution of prosocial behavior in large-scale societies is a special decision-making mechanism that can cause agents to perform prosocial behaviors that are not payoff-sensitive, or the creation of new mechanisms of social interaction that maintain the expression of payoff-sensitive behavior as group size increases. Without clarification, the perennial question of the extent to which prosocial behaviors in large-scale societies are compatible with (genetically) self-interested individuals will remain. Fully elucidating the evolved decision-making mechanism that humans use is extremely challenging. Determining, however, whether observed prosocial behaviors are payoff-sensitive is less challenging. For example, we can more easily determine whether systems of monitoring and sanctioning involve altruistic behaviors, or whether they directly benefit the individuals doing the monitoring and

sanctioning by increasing their material payoff (Guala, 2012; Ostrom, 1990). If it is the former, then this suggests that a special decision-making mechanism was key to their spread and maintenance. If it is the latter, then the creation of new mechanisms of social interaction is likely to have been the key driver. Empirical work should thus pay more attention to the payoff sensitivity of monitoring and sanctioning behaviors.

5 | CONCLUSION

Is the focus on the self-interested actor of behavioral models in economics, evolutionary biology and human behavioral ecology sufficient to explain the maintenance of prosocial behavior during and after the transition to large-scale societies? To answer this question, we have emphasized that two sub-questions need to be separated and answered. First, are social interaction mechanisms in large-scale societies *qualitatively* different from those in small-scale societies? Our review of the literature suggests that the answer is no. Both scales of society face fundamentally the same type of exchange and collective-action problems, but these problems became more difficult to solve on a *quantitative* scale in large-scale societies. Second, what is the decision-making mechanism by which individuals choose actions in exchange and collective-action scenarios? This is likely to be the same in both small-scale and large-scale societies, as there is unlikely to have been enough time during the origin of large-scale societies for genetic evolution to change the decision-making mechanism to a different agent type. The empirical evidence implies that humans are largely driven by payoff incentives when choosing prosocial behaviors in large groups of unrelated individuals. Thus, individuals express cooperative rather than altruistic prosocial behavior when interacting among non-relatives. This is consistent with RSM, and payoff-biased SLM agents, but where behavioral constraints appear to be set by evolved, genetic-interested preferences and predispositions as expected by FMM, and thus conforms to the general predictions of evolutionary biology applied to human behavior (Alexander, 1979; Alexander, 2014).

In light of these two assessments, we asked the question: what key mechanism allowed prosocial behaviors to be maintained as the size of exchange and collective-action problems increased? On one side, the cultural group selection hypothesis focuses on special decision-making mechanisms of behavior. By stressing the role of prestige and conformity biased social learning above that of material payoff, it places far less importance on self-interest, and considers individuals to be low on the payoff-sensitivity scale when expressing prosocial behaviors and thus does not fully face the constraints of evolutionary biology. On the other side, the institutional-path hypothesis focuses on the role of novel humanly devised mechanisms of social interaction. By stressing the role of changing the rules of the game, it assumes self-interested individuals and thus represents a null hypothesis facing the constraints of the scientific domains on which it bears; namely standard economics and evolutionary biology, and where individuals express cooperative prosociality among

non-relatives. Our analysis across societal scales suggests that it also provides a sufficient explanation for the maintenance of prosociality in the transition to large-scale societies.

Most crucially, we have delineated the two key mechanisms in this transition that need to be resolved (Figure 1): (1) the social interaction mechanism, and (2) the decision-making mechanism. We hope that this contributes to resolving the impasses underlying the conceptual foundations of human social evolution.

ACKNOWLEDGMENTS

We thank Alberto Micheletti, Jorge Peña, Nichola Raihani and two anonymous reviewers for useful comments on the manuscript.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Simon T. Powers  <https://orcid.org/0000-0003-0092-808X>

REFERENCES

- Alexander, R. D. (1979). *Darwinism and human affairs*. University of Washington Press.
- Alexander, R. D. (1990). Epigenetic rules and Darwinian algorithms. *Ethology and Sociobiology*, 11, 241–303.
- Alexander, R. D. (2014). Darwin's challenges and the future of human society. In W. Wayman, P. R. Williamson, B. B. de Mesquita, & S. Polac (Eds.), *Predicting the future in science, economics, and politics* (pp. 55–107). Edward Elgar Publishing.
- Alger, I., Weibull, J. W., & Lehmann, L. (2020). Evolution of preferences in structured populations: Genes, guns, and culture. *Journal of Economic Theory*, 185, 1–45.
- Alvard, M. S., & Nolin, D. A. (2002). Rousseau's whale hunt? Coordination among big-game hunters. *Current Anthropology*, 43, 533–559.
- André, M. O. (2011). Questioning the cultural evolution of altruism. *Journal of Evolutionary Biology*, 24, 2531–2542.
- Barkow, J. H., Cosmides, L., & Tooby, J. (Eds.). (1992). *The adapted mind*. Oxford University Press.
- Baumard, N. (2013). Cultural norms: Transmitted behaviors or adaptive responses? *Current Anthropology*, 54, 144–176.
- Binmore, K. (2005a). *Natural justice*. Oxford University Press.
- Binmore, K. (2005b). Economic man? Or straw man? *The Behavioral and Brain Sciences*, 28, 817–818.
- Bocquet-Appel, J. P. (2011). When the world's population took off: The springboard of the Neolithic demographic transition. *Science*, 333, 560–561.
- Boehm, C. (1999). *Hierarchy in the forest: The evolution of egalitarian behavior*. Harvard University Press.
- Bonner, J. T. (2011). *Why size matters: From bacteria to blue whales*. Princeton University Press.
- Bowles, S., Choi, J. K., & Hopfensitz, A. (2003). The co-evolution of individual behaviors and social institutions. *Journal of Theoretical Biology*, 223, 135–147.
- Boyd, R., Gintis, H., Bowles, S., & Richerson, P. J. (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 3531–3535.
- Boyd, R., & Richerson, P. J. (1990). Group selection among alternative evolutionarily stable strategies. *Journal of Theoretical Biology*, 145, 331–342.
- Boyer, P., & Petersen, M. B. (2012). The naturalness of (many) social institutions: Evolved cognition as their foundation. *Journal of Institutional Economics*, 8, 1–25.
- Bshary, R., & Bergmüller, R. (2007). Distinguishing four fundamental approaches to the evolution of helping. *Journal of Evolutionary Biology*, 21, 405–420.
- Burnham, T. C., & Johnson, D. D. P. (2005). The biological and evolutionary logic of human cooperation. *Analyse & Kritik*, 27, 113–135.
- Burton-Chellew, M. N., Mouden, C. E., & West, S. A. (2017). Social learning and the demise of costly cooperation in humans. *Proceedings of the Royal Society B*, 284, 20170067.
- Burton-Chellew, M. N., Nax, H. H., & West, S. A. (2015). Payoff-based learning explains the decline in cooperation in public goods games. *Proceedings of the Royal Society of London B: Biological Sciences*, 282, 20142678.
- Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163–228). Oxford University Press.
- Cosmides, L., & Tooby, J. (2013). Evolutionary psychology: New perspectives on cognition and motivation. *Annual Review of Psychology*, 64, 201–229.
- Davies, N. B., Krebs, J. R., & West, S. A. (2012). *An introduction to behavioural ecology*. Wiley-Blackwell.
- Delton, A. W., Krasnow, M. M., Cosmides, L., & Tooby, J. (2010). Evolution of fairness: Rereading the data. *Science*, 329, 389–389 author reply 390.
- Demps, K., & Winterhalder, B. (2019). Every tradesman must also be a merchant? Behavioral ecology and household-level production for barter and trade in Premodern economies. *Journal of Archaeological Research*, 27, 49–90.
- Earle, T. (1997). *How chiefs come to power: The political economy in prehistory*. Stanford University Press.
- Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature*, 13, 1–25.
- Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415, 137–140.
- Frey, S., & Sumner, R. W. (2019). Emergence of integrated institutions in a large population of self-governing communities. *PLOS ONE*, 14, 1–18.
- Fudenberg, D., & Tirole, J. (1991). *Game theory*. MIT Press.
- Fukuyama, F. (2011). *The origins of political order*. Profile Books.
- Gerkey, D. (2013). Cooperation in context: Public goods games and post-soviet collectives in Kamchatka, Russia. *Current Anthropology*, 54, 144–176.
- Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, 1, 107–143.
- Greif, A. (2000). The fundamental problem of exchange: A research agenda in historical institutional analysis. *European Review of Economic History*, 4, 251–284.
- Greif, A. (2006). *Institutions and the path to the modern economy: Lessons from medieval trade*. Cambridge University Press.
- Greif, A., Milgrom, P., & Weingast, B. R. (1994). Coordination, commitment, and enforcement: The case of the merchant guild. *Journal of Political Economy*, 102, 745–776.
- Guala, F. (2012). Reciprocity: Weak or strong? What punishment experiments do (and do not) demonstrate. *The Behavioral and Brain Sciences*, 35, 1–15.
- Hagen, E. H., & Hammerstein, P. (2006). Game theory and human evolution: A critique of some recent interpretations of experimental games. *Theoretical Population Biology*, 69, 339–348.
- Hamilton, W. D. (1964). The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*, 7, 1–16.
- Harsanyi, J. C., & Selton, R. (1988). *A general theory of equilibrium selection in games*. MIT Press.

- Haun, D. B. M., Rekers, Y., & Tomasello, M. (2014). Children conform to the behavior of peers; other great apes stick with what they know. *Psychological Science*, 25, 2160–2167.
- Hayek, F. A. (1988). *The fatal conceit: The errors of socialism*. Routledge.
- Henrich, J., & Boyd, R. (2001). Why people punish defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *Journal of Theoretical Biology*, 208, 79–89.
- Henrich, J., Chudek, M., & Boyd, R. (2015). The big man mechanism: How prestige fosters cooperation and creates prosocial leaders. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 370, 20150013.
- Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., & Ziker, J. (2006). Costly punishment across human societies. *Science (New York, NY)*, 312, 1767–1770.
- Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science*, 319, 1362–1367.
- Hill, K. (2009). Animal “culture”? In K. N. Laland & B. G. Galef (Eds.), *The question of animal culture* (pp. 269–287). Harvard University Press.
- Hoffman, P. T. (2015). What do states do? Politics and economic history. *The Journal of Economic History*, 75, 303–332.
- Hooper, P. L., Demps, K., Gurven, M., Gerkey, D., & Kaplan, H. S. (2015). Skills, division of labour and economies of scale among amazonian hunters and south Indian honey collectors. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 370, 20150008.
- Houser, D., & Wooders, J. (2006). Reputation in auctions: Theory, and evidence from eBay. *Journal of Economics and Management Strategy*, 15, 353–369.
- Hurwicz, L. (1996). Institutions as families of game forms. *Japanese Economic Review (Oxford)*, 47, 113–132.
- Jaeggi, A. V., & Gurven, M. (2013). Natural cooperators: Food sharing in humans and other primates. *Evolutionary Anthropology*, 22, 186–195.
- Jaeggi, A. V., Hooper, P. L., Beheim, B. A., Kaplan, H., & Gurven, M. (2016). Reciprocal exchange patterned by market forces helps explain cooperation in a small-scale society. *Current Biology*, 26, 2180–2187.
- Johnson, A. W., & Earle, T. (2000). *The evolution of human societies: From foraging group to agrarian state* (2nd ed.). Stanford University Press.
- Johnson, D. D. P., Stopka, P., & Knights, S. (2003). Sociology (communication arising): The puzzle of human cooperation. *Nature*, 421, 911–912.
- Kaplan, H., Gurven, M., Hill, K., & Hurtado, A. M. (2005). The natural history of human food sharing and cooperation: A review and a new multi-individual approach to the negotiation of norms. In H. Gintis, S. Bowles, R. Boyd, & E. Fehr (Eds.), *Moral sentiments and material interests: The foundations of cooperation in economic life* (pp. 75–113). MIT Press.
- Kaplan, H. S., Hooper, P. L., & Gurven, M. (2009). The evolutionary and ecological roots of human social organization. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 364, 3289–3299.
- Kelly, R. L. (2013). *The lifeways of hunter-gatherers: The foraging spectrum* (2nd ed.). Cambridge University Press.
- Kollock, P. (1998). Social dilemmas: The anatomy of cooperation. *Annual Review of Sociology*, 24, 183–214.
- Lamba, S. (2014). Social learning in cooperative dilemmas. *Proceedings of the Royal Society B*, 281, 20140417.
- Lamba, S., & Mace, R. (2011). Demography and ecology drive variation in cooperation across human populations. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 14426–14430.
- Ledyard, J. O. (1995). Public goods: A survey of experimental research. In J. H. Kagel & A. E. Roth (Eds.), *The handbook of experimental economics* (pp. 111–181). Princeton University Press.
- Legare, C. H. (2017). Cumulative cultural learning: Development and diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 114, 7877–7883.
- Lehmann, L., & Feldman, M. W. (2008). The co-evolution of culturally inherited altruistic helping and cultural transmission under random group formation. *Theoretical Population Biology*, 73, 506–516.
- Marlowe, F. W. (2005). Hunter-gatherers and human evolution. *Evolutionary Anthropology*, 14, 54–67.
- Marlowe, F. W. (2007). Hunting and gathering: The human sexual division of foraging labor. *Cross-Cultural Research*, 41, 170–195.
- Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory*. Oxford University Press.
- McNamara, J., Houston, A. I., & Collins, E. J. (2001). Optimality models in behavioral ecology. *SIAM Review*, 43, 413–466.
- Milgrom, P. R., North, D. C., & Weingast, B. R. (1990). The role of institutions in the revival of trade: The law merchant, private judges, and the champagne fairs. *Economics and Politics*, 2, 1–23.
- Milinski, M., Sommerfeld, R. D., Krambeck, H. J., Reed, F. A., & Marotzke, J. (2008). The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 2291–2294.
- Molleman, L., Quiñones, A. E., & Weissing, F. J. (2013). Cultural evolution of cooperation: The interplay between forms of social learning and group selection. *Evolution and Human Behavior*, 34, 342–349.
- Nettle, D., Gibson, M. A., Lawson, D. W., & Sear, R. (2013). Human behavioral ecology: Current research and future prospects. *Behavioral Ecology*, 24, 1031–1040.
- North, D. C. (1990). *Institutions, institutional change and economic performance* (political economy of institutions and decisions). Cambridge University Press.
- Olson, M. (1965). *The logic of collective action: Public goods and the theory of groups*. Harvard economic studies, revised edition (Vol. 124). Harvard University Press.
- Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge University Press.
- Peña, J., Volken, H., Pestelacci, E., & Tomassini, M. (2009). Conformity hinders the evolution of cooperation on scale-free networks. *Physical Review E*, 80, 016110.
- Petersen, M. B., Sznycer, D., Sell, A., Cosmides, L., & Tooby, J. (2013). The ancestral logic of politics. *Psychological Science*, 24, 1098–1103.
- Pinker, S. (2015). The false allure of group selection. In *The handbook of evolutionary psychology* (pp. 1–14). John Wiley & Sons, Inc., Chapter 36.
- Powers, S. T., & Lehmann, L. (2017). When is bigger better? The effects of group size on the evolution of helping behaviours. *Biological Reviews*, 92, 902–920.
- Powers, S. T., van Schaik, C. P., & Lehmann, L. (2016). How institutions shaped the last major evolutionary transition to large-scale human societies. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 371.
- Raihani, N. J., & Bshary, R. (2015). Why humans might help strangers. *Frontiers in Behavioral Neuroscience*, 9.
- Richerson, P., Baldini, R., Bell, A. V., Demps, K., Frost, K., Hillis, V., Mathew, S., Newton, E. K., Naar, N., Newson, L., Ross, C., Smaldino, P. E., Waring, T. M., & Zefferman, M. (2016). Cultural group selection plays an essential role in explaining human cooperation: A sketch of the evidence. *The Behavioral and Brain Sciences*, 39, e30.
- Richerson, P. J., & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution*. University of Chicago Press.
- Roth, A. E., & Murnighan, J. K. (1978). Equilibrium behavior and repeated play of the prisoner's dilemma. *Journal of Mathematical Psychology*, 17, 189–198.
- Schillaci, R. S., & Kelemen, D. (2014). Children's conformity when acquiring novel conventions: The case of artifacts. *Journal of Cognition and Development*, 15, 569–583.
- Scofield, J., Gilpin, A. T., Pierucci, J., & Morgan, R. (2013). Matters of accuracy and conventionality: Prior accuracy guides children's evaluations of others' actions. *Developmental Psychology*, 49, 432–438.

- Sefton, M., Shupp, R., & Walker, J. M. (2007). The effect of rewards and sanctions in provision of public goods. *Economic Inquiry*, 45, 671–690.
- Singh, M., Wrangham, R., & Glowacki, L. (2017). Self-interest and the design of rules. *Human Nature*, 28, 457–480.
- Smith, V. L. (1962). An experimental study of competitive market behavior. *Journal of Political Economy*, 70, 111–137.
- Sobel, D. M., & Kushnir, T. (2013). Knowledge matters: How children evaluate the reliability of testimony as a process of rational inference. *Psychological Review*, 120, 779–797.
- Sugden, R. (1993). Normative judgments and spontaneous order: The contractarian element in Hayek's thought. *Constitutional Political Economy*, 4, 393–424.
- Tavoni, A., Dannenberg, A., Kallis, G., & Löschel, A. (2011). Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 11825–11829.
- Testart, A. (1987). Game sharing systems and kinship systems among hunter-gatherers. *Man*, 22, 287–304.
- Thomas, M. G., Næss, M. W., Broardsen, B., & Mace, R. (2016). Smaller Saami herding groups cooperate more in a public goods experiment. *Human Ecology*, 44, 633–642.
- Tooby, J., & Cosmides, L. (2016). Human cooperation shows the distinctive signatures of adaptations to small-scale social life. *The Behavioral and Brain Sciences*, 39, E54. <https://doi.org/10.1017/S0140525X15000266>.
- Turchin, P. (2015). *Ultrasociety: How 10,000 years of war made humans the greatest cooperators on earth*. Beresta Books.
- Turchin, P., Currie, T. E., Turner, E. A. L., & Gavrilets, S. (2013). War, space, and the evolution of Old World complex societies. *Proceedings of the National Academy of Sciences of the United States of America*, 110, 16384–16389.
- van den Berg, P., Molleman, L., & Weissing, F. J. (2015). Focus on the success of others leads to selfish behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 112, 2912–2917.
- Vanberg, V. (1994). *Rules and choice in economics: Essays in constitutional political economy*. Routledge.

How to cite this article: Powers ST, van Schaik CP, Lehmann L. Cooperation in large-scale human societies—What, if anything, makes it unique, and how did it evolve? *Evolutionary Anthropology*. 2021;1–15. <https://doi.org/10.1002/evan.21909>