

Fair Social Contracts and the Foundations of Large-Scale Collaboration

Eric Beinhocker

22nd August 2022

INET Oxford Working Paper No. 2022-26



Fair Social Contracts and the Foundations of Large-Scale Collaboration

Eric D. Beinhocker*

Institute for New Economic Thinking, Oxford Martin School, University of Oxford Manor Road Building, Oxford OX1 3UQ, United Kingdom

Draft version: 22 August 2022

Abstract

Large-scale collaborations with non-kin are a unique feature of human societies and foundational to human civilization. Individual relationships with collectives can be thought of as "social contracts". This article argues that perceptions of social contract fairness are essential for effective large-scale collaboration and that factors likely to create perceptions of fairness are subject to empirical analysis. Drawing on empirical behavioral and social science literature, the article proposes nine dimensions of social contract fairness. It further proposes that each dimension is distinct, imperfectly substitutable, and universal, although with individual and cultural variations in interpretations and preference weightings. Finally, the article applies the nine dimensions to the breakdown in political collaboration in the U.S. and argues that for large segments of the population, all nine dimensions of social contract fairness were broken during the mid-1970s-2010s. The article concludes with thoughts on social contract repair and further research into perceptions of social contract fairness.

Keywords: Collaboration; fairness; social contracts; moral psychology; political populism.

Introduction

Modern society is built on large-scale, complex, collaborations amongst "strangers", i.e., people who are neither kin nor with whom one has thick personal bonds (Seabright 2010). Firms, global supply chains, governmental bodies, scientific research collaborations, religious communities, cultural organizations, and many other institutions provide examples of thousands to millions of people collaborating toward some shared ends. Most people in these networks will have never met, let alone be related or known to each other personally. Such large-scale collaboration amongst strangers appears to be a uniquely human capability that developed during the Neolithic period (Henrich et al. 2010; Bowles and Gintis 2011; Gintis 2011). In traditional, pre-agricultural societies, group sizes typically ranged from a dozen to 150 individuals (although larger collections of groups into societies could number in the thousands) (Hamilton et al. 2007; Diamond 1997; Bird et al. 2019; Dunbar 1993; 1992). These groups were comprised mostly of related individuals and individuals with thick, personal bonds. In contrast, in most present-day societies, individuals not only have kinship and personal relationships, but also have abstract relationships with large collectives of strangers, for example, their employer, their government, or their nation.

^{*} Professor of Public Policy Practice, Blavatnik School of Government and Executive Director, Institute for New Economic Thinking, Oxford Martin School, University of Oxford; External Professor, Santa Fe Institute. Email: <u>eric.beinhocker@inet.ox.ac.uk</u>.

We can refer to these abstract relationships between an individual and a collective as a "social contract". This paper will argue that our ability to collaborate with such large groups of strangers depends on social contracts being perceived to be "fair" by the individuals in them. If individuals perceive a social contract to be fair, then they are more likely to engage in high functioning collaborative behaviors; in contrast, if individuals perceive the arrangements to be unfair, then they are more likely to withdraw their collaboration or engage in destructive behaviors. In this sense, the perceived fairness or unfairness of social contracts is foundational to establishing and maintaining large-scale collaborations.

The key contribution of this paper is to further argue that whether a social contract is likely to be perceived as fair or unfair is a question subject to empirical analysis. Drawing on a growing literature in moral psychology, social psychology, neuroscience, anthropology, organizational studies, and behavioral economics, this paper advances an empirically informed hypothesis that judgements about the fairness of social contracts are based on nine dimensions arising from underlying moral instincts (reinforced by cultural norms) for relational fairness, process fairness, and distributive fairness, that evolved in humans to support cooperation with non-kin. I will further claim that while there may be individual preferences and cultural variations in interpretations and weightings across the nine dimensions, they are highly universal, distinct, and only imperfectly substitutable (i.e., they don't collapse to a unidimensional notion of utility). I will further argue that when a social contract satisfies these nine dimensions, it enables participants to trust the collective they are interacting with, and make pro-social collaborative choices, such as making costly investments in collective action with uncertain future payoffs and engage in altruistic behaviors. But when the nine dimensions of fairness are violated, not only may this result in a loss of trust and withdrawal of collaboration but may also trigger anti-social behaviors that undermine capacities for large-scale collaboration or stoke conflict between groups.

After describing this empirically informed hypothesis, I will then apply it to the specific problem of political discord in the U.S. I will argue that a critical reason why the U.S. has experienced a widespread loss of institutional trust, breakdown in political collaboration, and rise of political populism, is that all nine dimensions of social contract fairness were degraded for large segments of the population during the mid-1970s-2010s. This then implies that restoring social contract fairness is an essential step to restoring trust and functional politics. The nine dimensions provide a useful guideline for such a program of social contract renewal.

The Problem of Complex Collaboration at Scale

To see why fair social contracts are a necessary condition for collaboration at scale, it is helpful to clarify some of the specific challenges of initiating and sustaining such collaborations. The term "collaboration" is used in varying ways in behavioral and social science (and in this volume), so I will briefly define how I am using the term. Collaboration involves agents aligning their behaviors to achieve some mutual end. Yet, one can think of a spectrum of such behavioral alignment, from simple and mechanical, to complex and cognitively demanding. I will use the term "collaboration" to refer to more complex, cognitively demanding forms of behavioral alignment and distinguish it from "coordination" and "cooperation" as follows:

• *Coordination* occurs when agents align their behaviors to achieve some collective end. Processes of coordination may be quite mechanical and not require complex cognitive capacities. For example, honeybees regulate their hive temperature by generating heat from their muscles when the temperature is too low and beating their wings when it is too hot. Individual bees are genetically programmed to engage in thermoregulation at varying temperature points such that when the temperature deviates from the target by a small amount, a small number of bees thermoregulate; and when the temperature deviation is greater, more bees join in. This feedback mechanism coordinates the bees to smoothly respond to temperature fluctuations, maintaining the hive at roughly 35°C.

- Cooperation occurs when agents align behaviors in mutually beneficial ways, anticipating or • understanding the behavior of other agents. Imagine a dog and a human playing a game of fetch. It is a cooperative game where if both behave in certain ways, both get pleasure (although perhaps with asymmetric payoffs for the canine). This setting is more complex and cognitively demanding than the simple coordination example (Moll and Tomasello 2007; Tomasello and Carpenter 2007). There is *shared intentionality* in that the players need to understand the rules of the game and payoffs and then voluntarily choose to enter the game (e.g., the dog initiates the game by dropping a ball at the human's feet and wagging its tail). Furthermore, such cooperative games require that each player has a theory of mind about the other (e.g., the dog has a theory or expectation about how the human will behave when he or she drops the ball at the human's feet). Furthermore, each agent must understand their own causal role and that of other agents, enabling them to see, amongst the large set of possible actions, the sequence of actions that will yield the cooperative payoff (e.g., the human understands if they pick up the ball and toss it, the dog will fetch it). Other examples of cooperative behaviors in both humans and non-humans include group hunting, grooming, mutual defense, shared tool use, and shared care for the young.
- *Collaboration* can then be thought of as a sub-set of cooperation that occurs when agents align behaviors in mutually beneficial ways, but where the structure of the game is not given and static; instead, *the players themselves are inventing, co-creating, and evolving the structure of the game over time.* For example, imagine a group of people who come together to start-up a new business. While such an activity certainly requires both coordination and cooperation, many of the rules and future payoffs of this game are unknown and perhaps unknowable (Knightian uncertainty). In fact, the rules of the game will be at least partially co-created by the players themselves as the game progresses and are likely to evolve over time. And an activity such as building a business is not a single game, but involves *multiple, interlinked, repeated games and sub-games*, both within the entrepreneurial group itself, but also dynamically interacting with other games and players in the environment. Furthermore, the *causal links between player actions and future payoffs may be significantly separated by time and space, be noisy and complex, and it may be difficult or impossible to disentangle individual contributions to collective results.*

What I call collaboration is thus clearly more cognitively demanding than either coordination or playing cooperative games with fixed rules and clear payoffs. There is ample evidence of what I have called coordinating and cooperative behavior in many species (Musgrave 2023). But what I have called collaborative behavior appears at least unique to primates and possibly unique to humans (some might argue that non-human primate behavior is more correctly viewed as proto collaborative rather than fully collaborative in the sense I have described). But what is clearly uniquely human is our ability to engage in collaborative behaviors at scale with strangers. Chimpanzees, for example, cooperate (or possibly collaborate) in small groups of kin, near-kin, and known individuals in troupe sizes of 20 to 30. But humans can collaborate to build an Airbus A380 aircraft, assembled from 4 million parts, and manufactured by tens of thousands of people in 1,500 companies from around 30 different countries.

Social Contracts as the Foundation of Large-Scale Collaboration

How then can agents align their actions in complex, dynamic settings, with large groups of strangers, evolving, co-created rules, with imperfect information on the goals, contributions, and abilities of those strangers? Collaboration at scale requires agents to make a mental leap and see

Eric Beinhocker

themselves as having collaborative relationships not just with other agents *individually*, but with a *collective* of agents as an *entity*. As those entities may contain many agents not known to them, and the composition of agents may change over time, agents must be able to abstract the entities from the individuals who form them, i.e., the agent sees themself in a relationship with a collective entity such as a tribe, firm, government, school, or a sports team.

We can define that relationship between the individual and the collective entity as a *social contract*. The idea of a social contract goes back to the Ancient Greeks, but modern discussions have their roots in concepts introduced by figures such as Thomas Hobbes, John Locke, and Jean-Jacques Rousseau, and then further developed in the 20th century by John Rawls. Describing the relationship between an individual and a collective as a "contract" implies a mutuality of commitments: the individual voluntarily aligns their behaviors with the interests of the collective, agrees to contribute effort and resources towards collective goals, and submits to being governed by collectively enforced social arrangements, in exchange for some set of future benefits. A core claim of social contract theory is that if an individual voluntarily submits to being governed by such collectively enforced social arrangements, then they must *ipso facto* view those arrangements as fair and legitimate, or at least "fair enough" (D'Agostino, Gaus, and Thrasher 2021). In contrast, if people do not view the arrangements as fair, then they will either withdraw their cooperation or only submit to collective governance involuntarily.

To make our use of the term "social contract" more precise, we can think of it as follows: There is a "game" where if a group of players collaborates, they will potentially generate some nonzero-sum gains. An individual offers their collaboration to the collective group of players, conditioned on the following terms:

- I consent to play the game,
- I agree to play by the rules of the game, and
- I promise to play the game to the best of my abilities,
- If the game is fair

The *social contract* then defines the set of arrangements for that conditional offer of collaboration between the individual and the collective. The individual then makes judgements on the fairness of that contract based on their moral intuitions and cultural norms. Those feelings of fairness or unfairness then influence the agent's collaborative or non-collaborative behaviors.

As discussed above, collaborations involve a significant degree of uncertainty and imperfect information. In economic terms, this would imply that we cannot write a complete contract between the individual and the collective for the collaboration. Thus, an agent's doxastic representation will be incomplete. As such, the agent cannot simply make a self-interested rational choice as in, say, a cooperative game setting where the rules and payoffs are known in advance, probabilities can be assessed, a complete contract written, and a rational choice can be made (Binmore 1994). I am thus hypothesizing that the evaluative criteria for the individual's commitment to a collaboration is *fairness*, which may incorporate aspects of self-interested rationality (e.g., I might not think the game is fair if the costs of my contributions outweigh the benefits), but involves a broader set of evaluative criteria (e.g., even if I receive net-positive payoffs, I might not think the game is fair if I am treated less well than others).

Process Fairness, Equality, and Deservedness

If fair social contracts are a necessary condition for effective, complex, large-scale collaborations, then the next question is what do we mean by "fair"? This question has been widely explored philosophically, for example, by asking what kinds of social contracts would lead to a morally just distribution of resources, power, and rights in society (e.g., (Rawls 1971)), or asking what social

contracts individuals would rationally choose to enter (e.g., (Gauthier 1986; Binmore 1994). But we will take a different approach (more Hume than Plato) and start from an empirical question: what are the characteristics of social contracts that most people are likely to perceive as fair?

We should note that this empirical perspective does *not* imply that people's individual moral intuitions about fairness will necessarily lead to social contracts that are just from a societal or philosophical perspective. Nor does it imply that empirically observed moral intuitions about social contract fairness will be logically consistent, non-contradictory, or economically rational. Instead, I am making a simpler claim: if individuals view their social contract arrangements as fair, then they are more likely to engage in effective collaborative behaviors. Therefore, in designing policies and institutions to maximize collaboration, it is useful to know what kind of arrangements are likely to be viewed as fair.

Human instincts about fairness appear to have deep evolutionary roots and likely evolved to facilitate cooperation and collaboration (Gintis 2003; 2004; Gintis et al. 2008; Gintis 2011; Bowles and Gintis 2011). Fairness instincts are found in non-human primates (Brosnan 2011; 2013) and appear early in child development (Shaw and Olson 2012; Shaw, DeScioli, and Olson 2012; Gredebäck et al. 2015). Feelings of fairness or unfairness also evoke distinctive neuro-physiological responses, including the production of hormones associated with trust, pleasure, stress, or anger, and heightened activity in the amygdala brain region (Molly J. Crockett 2009; Haruno and Frith 2010; Chang et al. 2015; Tanaka, Yamamoto, and Haruno 2017). Additionally, there appears to be some genetic heritability in cooperative norms (Cesarini et al. 2008). Certain fairness norms appear to be highly universal, although their specifics may be more culturally variable. For example, an experimental study of resource sharing by children in seven societies showed a universality of preferences for equal outcomes and rejection of unequal outcomes that disadvantaged individuals (Blake et al. 2015). However, rejection of unequal outcomes that advantaged individuals was more culturally variable. Similarly, a large cross-cultural study of reciprocity norms showed high universality in the structure of those norms but with cultural variability in parameters for what specifically was considered fair, reciprocal behavior (Henrich et al. 2004).

One specific finding that is important for our purposes is that judgements about fairness are, to a significant extent, judgements about process fairness, and assessments of distributive outcomes are used as signals of whether a process is fair or unfair, based on a priori expectations of process outcomes (Starmans, Sheskin, and Bloom 2017). Two examples illustrate this point: Imagine a group playing a coin-flipping game. Our *a priori* expectation would be that if the game was played fairly, the outcome would be a roughly equal distribution of heads versus tails amongst participants. If, however, the outcome was significantly unequal, with, say, a number of people flipping highly unlikely, long streaks of heads, we would then suspect the game was not played fairly, that they somehow cheated. So based on participants' understanding of the process, the expectation is an equal outcome and unequal outcomes are a signal of potential process unfairness. Now imagine a second game, a 100-meter running race between a random group of people and Usain Bolt, the world record holder. Our a priori expectation would be that a fair race would yield an unequal outcome, with Bolt winning by a lot. If, on the other hand, the race yielded an equal outcome, with everyone crossing the line at the same time, we would suspect that something about the race was unfair-it was rigged. So, an equal or unequal outcome is not inherently fair or unfair but instead may be a signal as to whether a given process is fair or not.

In games involving distributions of resources, however, people express strong preferences for equal outcomes as a kind of default setting (Shaw and Olson 2012; Blake and McAuliffe 2011; Mccrink, Bloom, and Santos 2010). But then exceptions are made to the equality default rule based on perceptions of *deservedness* or *merit* and unequal distributions may then be regarded as fair (and equal outcomes may then be viewed as unfair). For example, imagine a group of friends sitting at a table, and one puts a large cookie in the middle; how do they divide it up? The default answer

would be equally; if one grabbed more of the cookie they would be viewed as greedy and unfair. We have strong instincts for relational fairness, or more specifically, moral equality-the idea that we are each of equal worth and moral standing (Cook and Hegtvedt 1983; Killen et al. 2001; Konstantareas and Desbois 2001). Intuitions about moral equality arguably appear in primates (Brosnan 2011; 2013), develop in young children (LoBue et al. 2011), appear across cultures (Kim and Leung 2007) and are a central idea in many religions (i.e., we are all equal before God). The intuition for moral equality tells us that, in the absence of any other information, we each deserve an equal share of the cookie. Now imagine it turns out that one person just returned from a long run; the others might view her as deserving of more cookie. Or one of the group members had lost his job and so might need more cookie to cheer him up. Or someone starts acting in an objectionable way and the group expresses its displeasure by saying, "you don't deserve any cookie!" While equal distribution may be the default rule, we also have instincts for deservedness, for merit-based exceptions to the equality default. A fair process also takes account of information on differences in circumstance, merit, luck, and the nature of the game being played, to adjust the outcome based on such factors, ending in a result where everyone "gets what they deserve". Denying these merit-based exceptions to the equality default rule would then itself likely be viewed as unfair. In a small group of people personally known to each other (such as our cookie example), such a fair process may be quite informal, although still with potential for contested views as to what counts as "deserving" and what that in turn means for allocation. In a large group where people are not personally known to each other, information is imperfect, and behavioral monitoring is limited, the challenges become significant and confidence in process fairness becomes critical.

Nine Dimensions of Fair Social Contracts

We can use these findings to construct a simple framework for assessing social contract fairness. We have preferences for *relational fairness* that includes the principle of moral equality as a precondition (e.g., it is hard to have a fair process with unfair power relations); we then have preferences for *procedural fairness* that includes the principle of deservedness or merit; and we finally have preferences for *distributional fairness* that relates perceived outcomes to expected outcomes to make assessments about the fairness of the game. Building on these general underlying preferences we can then ask what are the specific attributes of social contracts that are likely to be viewed as fair?

In this section I will argue that there are nine attributes that contribute to perceptions of social contract fairness, summarized in Table 1.

Underlying moral	Dimensions of fair	Description
preferences	social contracts	
Relational fairness	1. Agency	I have a choice whether to play the game and choices within the
		game
	2. Inclusion	I have an opportunity to play the game, I am not excluded
	3. Dignity	If I play by the rules and contribute to the best of my abilities, I
		will be valued, respected, and have status
Procedural fairness	4. Rules-based	I know the rules of the game and they are applied equally to
		everyone
	5. Meritocratic	I, and everyone else, will receive rewards and punishments in the
		game based on merit
	6. Security	If I play by the rules and contribute to the game, but suffer
		misfortune through no fault of my own, I will be protected
Distributional fairness	7. Capabilities	I have the capabilities to play the game or the opportunity to
		acquire them
	8. Reciprocity	If I play by the rules and contribute, others will reciprocate, and
		I will share in the game's rewards
	9. Progress	If I play by the rules and contribute to the game, my life, and the
		lives of those I care about will improve

Table 1. An empirically informed hypothesis: nine dimensions of fair social contracts.

I will briefly elaborate on each dimension and the evidence supporting them. I have phrased these as "I" statements as they are from the perspective of the individual agent facing the collective.

Relational fairness

- 1. Agency—I have a choice whether to play the game and choices within the game. If I am forced to play the game (e.g., a slave) I am unlikely to view the social contract as fair. Likewise, if I enter the game but all choices are made for me (particularly if I cannot predict the outcomes from such involuntary choices), I am unlikely to view the contract as fair. One can think of agency as an aspect of relational fairness as it answers the question of who has the power to make decisions that affect an individual. The literature shows that agency is critical to healthy human functioning and sense of identity, motivation and engendering cooperative behaviors (Bandura 1997; 2006; Ryan and Deci 2000; Akbaş, Ariely, and Yuksel 2019). In economic experiments, subjects valued agency to be a key element in determining the fairness of the game (Akbaş, Ariely, and Yuksel 2019; Konow 2000). One may be able to create a kind of large-scale forced coordination with an army of slaves, but it is not possible to create true collaboration capable of solving complex problems (e.g., an army of slaves could not develop a novel vaccine). There are, of course, degrees of agency: for example, people may have choices of employment but still must work to make a living; or people may have democratic political choices, but still must obey the law regardless of who wins an election. Nonetheless, an ability to make choices within a set of options limited by agreed rules, remains a critical component of both fairness and effective collaboration.
- 2. Inclusion—I have an opportunity to play the game, I am not excluded. If one chooses to play the game, but is excluded for unjust (i.e., non-merit based) reasons, one is likely to view the game as unfair. An obvious example is the long history of economic, political, and social exclusion for reasons of race, gender, religion, ethnicity, class, or sexual preference or identity. Inclusion is an aspect of relational fairness in that unjust exclusion violates the principle of moral equality. There is significant social psychology evidence on the detrimental effects of exclusion on subjective well-being (Bellani and D'Ambrosio 2011; Gross-Manos 2017).

Furthermore, unjust exclusion of others appears to trigger people's sense of fairness and prompt them to action (Moor et al. 2012; Williams 2007; Tuscherer et al. 2016; MacDonald and Leary 2005). There is a link to procedural fairness, as it is non-meritbased exclusion that triggers feelings of unfairness. When, for example, Jackie Robinson was excluded from Major League Baseball simply because of the color of his skin that was deeply unfair. But if a middle-aged professor is excluded from Major League Baseball because he is terrible at it, then that would be fair, particularly if (per below) he was previously given access to acquiring capabilities (e.g., opportunities to play Little League), and the process for judging players is meritocratic.

3. Dignity—If I play be the rules and contribute to the best of my abilities, I will be valued, respected, and have status. Humans are status-conscious and status-seeking creatures. Status and dignity evoke strong emotions tied to feelings of fairness (Folger and Cropanzano 2001; Stets 2004). Feeling like a valued contributor to the collective is a powerful motivating force in collaborative behavior and a critical element in forging a common identity with the collective (Axelrod and Hamilton 1981; Fox and Guyer 1978). Violations of dignity (i.e., feeling underappreciated or disrespected) can evoke strong negative emotions and feelings of injustice and lead to behaviors detrimental to collaboration (Greenberg 1988; Milinski, Semmann, and Krambeck 2002). Affording people dignity recognizes their worth and standing and thus is an aspect of relational fairness.

Procedural fairness

- 4. *Rules-based—I know the rules of the game and they are equally applied.* This attribute bridges preferences for relational fairness and process fairness: If everyone is of equal moral worth, then the rules of the game must apply equally to everyone, and a fair process is one in which the rules are known, followed, and equally enforced. There is evidence from cognitive science and social psychology that people have strong preferences for such procedural fairness (Greenberg 1987; 1990; Folger 1986; Folger and Cropanzano 2001; Henrich et al. 2010; Marwell and Ames 1981; Engel 2005). People's degree of association between respecting the rules and fairness varies by culture (e.g., cross-cultural studies of tax compliance)(Cummings et al. 2005), but the idea that the same set of rules should apply to everyone (even if not always complied with) does appear to be widely held—and situations where rules are unevenly applied, manipulated or ignored by privileged groups, or opaque and subject to arbitrary interpretation or enforcement are widely viewed as unfair.
- 5. Meritocratic— I, and everyone else, will receive rewards and punishments in the game based on merit. People appear to have intuitive notions of merit and deservedness, that rewards should go to those who contribute to the collective effort, engage in reciprocal behaviors, have relevant capabilities, play by the rules, are of good character, and so on (Adams 1965; Kulik and Ambrose 1992; Cohn et al. 2011; Baumard, Mascaro, and Chevallier 2012). While (Sandel 2020) argues that meritocracy can lead to excessive individualism, reinforce inequities, and harm collective endeavors, nevertheless, people tend to intuitively see meritocratic processes as providing a basis for distributive justice. For example, most people would see a university admission process based on some notion of merit (e.g., student academic achievement, potential to contribute to the student body, etc.) as fairer than one based on non-meritorious criteria (e.g., a parent's donations to the university). But what constitutes "merit" is highly contestable and context-dependent (e.g., some might argue that conventional measures of academic merit favor students born to wealthy parents who can afford private schools and tutors). But even though the

criteria for merit may be contested, most people would view a meritorious process as more likely to lead to distributional fairness.

6. Security-If I play by the rules and contribute to the game, but suffer *misfortune through no fault of my own, I will be protected.* There appears to be widely shared instincts for luck egalitarianism, the recognition that bad luck can strike any of us for reasons not of our own making (Nagel 1979; Dworkin 1981; Anderson 1999; Tinghög, Andersson, and Västfjäll 2017). One might get cancer, be laid off in a recession, or face hunger in a drought. While we cannot protect against all unlucky situations, humans have strong empathetic instincts in such situations and are often willing to act charitably and altruistically (Dovidio 1984; Boyd and Richerson 2005; Fehr, Bernhard, and Rockenbach 2008; Masten et al. 2010; Zak 2011a; Pavey, Greitemeyer, and Sparks 2011). Furthermore, there are strong instincts for mutual protection for fellow members of one's group, particularly if the unlucky individual is seen as a contributor to the group's welfare. However, there are sensitivities to the potential for free riding and abuse of empathetic feelings. Thus, government social safety net programs tend to have higher political support when they insure against bad luck that could strike anyone, require reciprocity, and monitor against abuse (Fehr, Fischbacher, and Gächter 2002; Fong, Bowles, and Gintis 2006; Batson et al. 2007; Sasaki and Uchida 2013; Fehr and Gachter 2017).

Distributional fairness

- 7. Capabilities—I have capabilities, or opportunities to acquire capabilities, to play the game. A fair game requires the capabilities to play. As Amartya Sen has argued, positive freedom requires capabilities to provide the functionings necessary for a fulfilling life (Sen 1985; 2008). This is particularly important for games we play out of necessity, notably the "earn a living game". Yet there is a birth lottery in the distribution of capabilities (e.g., you might be born to a poor family, or where a good education is not available). So distributional fairness requires that people have opportunities to acquire capabilities and fulfill their potential. Likewise, it is unfair to expect people to play a game for which they do not have the capabilities or do not have the opportunities to acquire them. For example, systematic underinvestment in female education violates distributional fairness (Nussbaum 2003; Robeyns 2006; Sen 2008; Nussbaum 2002). While the author is not aware of literature providing direct evidence of people's perception of capabilities as an attribute of fairness, there has been work in psychology connecting capabilities to feelings of well-being (Jayawickreme and Pawelski 2013) and on capabilities as a basis for agency and empowerment (Shinn 2015). One can hypothesize that a social contract that requires certain actions or behaviors but does not provide the capabilities to fulfill those expectations would be widely regarded as unfair.
- 8. Reciprocity—If I play by the rules and contribute, others will reciprocate, and I will share in the game's rewards. I have categorized reciprocity as a form of distributive fairness as reciprocity is an observable and expected outcome in a fair process, i.e., if the process is fair, I will observe reciprocal behaviors in the contributions and the sharing of rewards between players (in ultimatum game experiments, for example, players only have information on their own contributions and observations of distributive outcomes, (Guth and Tietz 1990)). There is a large literature showing that intuitions and norms of reciprocity develop in early childhood (Bos et al. 2010; House et al. 2013; Warneken and Tomasello 2013), appear across cultures (Chen, Chen, and Portnoy 2009; Kuwabara et al. 2007) and are foundational to establishing cooperation and collaboration (Trivers 1971; Adams 1965; Greenberg

1990; Axelrod and Hamilton 1981; Bowles and Gintis 2011). Furthermore, the evidence shows that when reciprocity norms are violated, agents will not only withdraw from cooperation, but even punish the individuals or institutions violating their expectations of fairness (Trivers 1971; Adams 1965; Greenberg 1990; Axelrod and Hamilton 1981; Bowles and Gintis 2011). The literature further shows that such punishment may even be to the punisher's detriment, altruistically bearing a cost to enforce norms of reciprocity.

9. Progress— If I play by the rules and contribute to the game, my life, and the lives of those I care about will improve. Progress can be thought of as a form of distributional fairness over time. Progress, both in economic reality and its perception as either a moral good or right, is a phenomenon that appears to have developed in certain societies circa the 17th-19th centuries (Maddison 2007; Wootton 2018). It is not clear that similar notions exist in traditional societies, and it has historically been viewed differently in non-Western cultures. Nonetheless, today the "right to progress" has become a widely held idea across the globe (Wegener 1991; Alesina, Tella, and MacCulloch 2004; Rodon and Sanjaume-Calvet 2020; Day and Fiske 2017). Furthermore, expectations of progress in one's life, and emotions of hope for the future, are strongly associated with subjective well-being (Pleeging, Burger, and Exel 2021).

Another way to see the impact of these nine dimensions on perceptions of fairness is to think of a social contract with the opposite characteristics: Imagine you are offered a social contract to play a game where there is:

Relational unfairness

- 1. You do not have agency to make choices
- 2. You are excluded from critical aspects of the game
- 3. You will not be respected for your role and contributions

Procedural unfairness

- 4. You do not know the rules and/or they are unequally applied
- 5. You and others will not receive rewards and punishments based on merit
- 6. You will not be protected from misfortune

Distributional unfairness

- 7. You do not have the capabilities necessary to play nor opportunity to acquire them
- 8. You are not reciprocally rewarded for your contributions
- 9. And, finally, even if you play and contribute to the best of your abilities, your life and those you care about will not improve

Is this a fair game? Would you accept a social contract to play it? Probably not. Would anyone *voluntarily* agree to play such a game? It is highly unlikely. The next question then is, if even *one* of the above statements is true, would you regard the contract to play the game as fair? My hypothesis is that if even one of these statements is true, then that is sufficient to make the game unfair for most people. This is what I mean by saying that the nine dimensions are "distinct and non-substitutable". I do not mean to imply that in the real-world people cannot or do not make trade-offs across the attributes—they can and do. Instead, what I am proposing is that *all* nine are necessary to at least *some degree* for a contract to be viewed as fair. A zero value for any of the nine will trigger moral intuitions of unfairness. For example, even if a social contract is very high on meritocracy, that is no substitute for not having capabilities. Nor will investing in more capabilities make up for being excluded.

How universal are these nine dimensions? Moral psychology researchers have observed a high degree of universality in moral intuitions, social-emotional responses, and neural-cognitive

patterns, but also significant individual and cultural variability in how people weigh, trade-off, and interpret moral preferences (Greene et al. 2004; Singer 2005; Gintis et al. 2008; Shenhav and Greene 2010; Zak 2011b; Molly J. Crockett et al. 2014; Molnar-Szakacs 2011). Jonathan Haidt likens findings on the universality of dimensions of moral preferences to "taste buds" (Haidt 2012). Every human has the same five taste receptors (sweet, sour, salty, bitter, and umami), but individuals and cultures vary in their preferences as to how these universal tastes are combined in specific foods. Likewise, in the case of the above nine dimensions, I would propose that they are highly universal and cross-cultural, but individuals and cultures will vary in their preferences for how they are combined and traded-off in specific social contracts.

It is important to note that, despite the universality of dimensions, different interpretations and weightings can nonetheless result in highly contested views as to what specifically constitutes a fair contract. Examples include people with differing political views interpreting the provision of access to capabilities in different ways or debating how much security is "enough" in the welfare state. Or one branch of a religion that interprets sacred texts as justifying the exclusion of women from education or certain occupations (i.e., adherents view the texts as providing a merit-based justification for exclusion, the "merit" being "God says so") versus another branch of the same religion that interprets the texts as promoting moral equality of both women and men and therefore inclusion. Again, my claim is *not* that there is universality to the *specific* social contracts that people perceive as a fair or unfair, but rather that there is universality to the *evaluative framework* people use when making such judgements.

Social Contract Violation and Political Populism

This universality of evaluative framework can give us insights into how and why collaboration breaks down, and why the fairness of specific social contracts may be contested. As an example, in this section we will apply the nine dimensions to briefly analyze the breakdown of political collaboration in the U.S. I will argue that a major deterioration in the fairness of social contracts in the U.S. from the 1970s-2010s led to widespread perceptions of contract violation, which in turn laid the emotional foundations for a drop in political collaboration and rise in political populism.

Over the past decades the U.S. and various other countries have seen a breakdown in political collaboration, increased polarization, a loss of faith in democracy, a loss of trust in key institutions, and a rise of populist and authoritarian political figures (Pew 2016; Hawkins et al. 2019; Edelman 2020). Using a variety of metrics, (Putnam and Garrett 2020) identify the late 1960s to early 1970s as a peak in U.S. social, cultural, and political cohesion. By 2015, this cohesion had deteriorated to levels not seen since the Civil War. A variety of explanations have been put forward to explain this broad trend, including increases in economic inequality, economically "left behind" regions, cultural and demographic factors, and changes in the media landscape. Surveys and studies of recent election results find, however, that instead of material explanations (e.g., economic, education, demographics), the most explanatory variables are attitudinal and emotional (Inglehart and Norris 2017; Cox, Lienesch, and Jones 2017; Hawkins et al. 2019; Green and McElwee 2018; Mutz 2018; Ward et al. 2020). Notably, that voters who support populist candidates report feelings of a loss of agency over their lives and communities (e.g., the Brexit slogan "Take back control"), alienation and exclusion from the broader culture (e.g., perceptions that their racial, ethnic, cultural, or religious group is becoming a "persecuted minority"), a loss of reciprocity (e.g., we pay our taxes while "others" get benefits), a view that powerful elites are playing by different rules (e.g., "the game is rigged"), significant fears of status loss, and a loss of feelings of security and hope for the future. Such attitudes align very closely with a negation of the attributes of fair social contracts and indicate significant feelings of contract violation (Table 2).

Unfair game/broken contract	Example attitudes
1. Loss of agency	Others are controlling our lives; we need to take back control
2. Exclusion	My group is being discriminated against, excluded from opportunities
3. Loss of dignity, status	People like me used to be valued members of society, now we are not
4. Rules violations	The game is rigged; powerful people and favored groups play by different rules
5. Less meritocratic	I work hard but I can't seem to get ahead, while less deserving people do
6. Decreasing security	I worry about my finances, health, retirement, crime, and our nation's security
7. Insufficient capabilities	I've worked hard all my life, but my skills are no longer valued
8. Loss of reciprocity	I work hard and deserve what I get, but others don't and get a free ride
9. Loss of progress, hope	Things are getting worse not better, I fear for my children's future

Table 2. Attitudes of supporters of populist political candidates and causes align closely with feelings of social contract violation

Such feelings of social contract violation have become widespread but have been most heavily concentrated over the past decades in two broad groupings. First, are white, working-class, largely Christian, largely male, ex-urban voters. For these voters, the "others" who have violated the contract are people of different political beliefs, racial groups, religions, immigrants, and gender identities, as well as foreign countries (Pew 2021). The overall feeling of these voters is that their own group has worked hard, contributed to society, and played by the rules, but has lost opportunities and status because of unfair play by the "others". Furthermore, the rule-setters and referees who are supposed to ensure a fair game-the "cultural elite" of political, business, media, and academic leaders-have not only allowed the contract to be broken, but they have been complicit in it, helping break the contract to serve their own interests. Thus, perceptions (and misperceptions) of contract violation have contributed to increases in racism, anti-immigrant sentiment, misogyny, and an anti-elite backlash. Right-wing political figures, parties, and media were the first to pick up on these growing feelings in the 2000s, and began exploiting them, creating a major political re-alignment that shifted many white, working-class voters from left-leaning to right-leaning political parties in the U.S. and Europe, and leading to a dramatic rise in right-wing populism exemplified by Brexit and the election of Donald Trump.

But while white, working-class voters drove the rise in right-wing populism, they were not the only ones feeling social contract violation. The second broad group reporting such feelings includes struggling lower income families, citizens in deprived urban communities, people from historically excluded racial, religious, and gender groups, and young people who fear for their future, have also expressed feelings like those in Table 2. However, for them, the "others" violating the social contract include billionaires who don't pay their taxes, large corporations who exploit workers and profit at the expense of others, "privileged" groups who benefit from historical injustices, and a political class that rigs the game. Such voters were attracted to left-wing populists such as Bernie Sanders in the U.S. and Jeremy Corbyn in the UK.

It is notable that, preceding the 2016 populist wave, the 2008 financial crisis resulted in angry, grass-roots populist movements on both the political right (e.g., the Tea Party movement) *and* the political left (e.g., the Occupy Wall Street movement), both of which articulated broken contract narratives but differed as to who was doing the violating. While the politics and policies of the right-wing and left-wing populists differ starkly—and the racism and sexism of right-wing populist figures and segments of voters has no justification and must be condemned—the *emotional structure* of popular support has been similar on both sides, founded on feelings of moral outrage over a broken social contract.

These feelings of a broken contract among large segments of voters are not unfounded. While material explanations may not be directly causal in explaining the populist rise, underneath each of these attitudinal dimensions are changes and trends in the structure of the economy and society that have arguably provoked these feelings. (Putnam and Garrett 2020) identify the late-1960s to early 1970s as the turning point in U.S. social cohesion. Beginning in the mid-1970s, the structure of the U.S. economy underwent a profound shift: productivity growth and worker income growth de-coupled, incomes for about 90 percent of households stagnated in real terms while almost all the gains of growth flowed to the top 1 percent of earners, the middle-class shrank as a percentage of the population, social mobility declined, and various measures of economic insecurity increased.[†] While technological change and globalization contributed to these trends, particularly from the 1990s onwards, cross-country studies suggest much of this change resulted from shifts in economic ideology and policy that started during this period (Nolan 2018). A shift towards more "neoliberal" economic policies, both in right-wing (e.g., Reagan, Thatcher) and leftwing (e.g., Clinton, Blair) governments resulted in a shifting of the tax burden away from the wealthiest and corporations to middle and lower-income workers, relative reductions in public investment (e.g., education, infrastructure), weakening of the social safety net, changes to labor market regulations that reduced union and worker power, central bank policies that prioritized low inflation over employment and wage growth, and trade policies that favored corporate over worker interests. At the same time, changes in corporate practices (e.g., moving from balanced stakeholder to shareholder value-maximizing governance, outsourcing, offshoring, reductions in pension and health benefits, less secure employment) shifted gains in productivity away from workers and towards shareholders, while reducing worker power and security (Lazonick and O'Sullivan 2000). These economic changes coincided with a growing influence of money in U.S. politics and various failed attempts to regulate it in the 1980s-2000s (culminating in the Supreme Court's 2010 Citizen's United decision), as well as increasingly effective gerrymandering of Congressional and state legislative districts, and demographic shifts making the U.S. Senate less representative of the population (Smith 1995; Mansbridge 1999; Teachout 2016). Together, these changes made the U.S. democratic system less responsive to citizen concerns and more responsive to those of wellfunded interests (Lindsey and Teles 2017).

At the same time, there were also major social, cultural, and demographic shifts as well. The Civil Rights and feminist movements of the 1950s-60s empowered historically excluded and underrepresented groups, and demographic shifts saw the non-white population in the U.S. grow from 20% in 1980 to 40% by 2019 (Frey 2020), and the immigrant population of the U.S. tripled during this period from 4.8% in 1970 to 13.7% in 2020 (Budiman 2020). Furthermore, during this period, there were significant population movements from rural to urban and suburban, and from rustbelt regions to the sunbelt. While studies show that voters with racist, anti-immigrant, and sexist attitudes were a significant factor in Donald Trump's election (Hooghe and Dassonneville 2018; DeSante and Smith 2020; Bock, Byrd-Craven, and Burkley 2017; Cassese and Barnes 2019), there were also significant numbers of voters who were reacting to feelings of economic injustice and unresponsive political system. And again, such feelings were not limited to voters traditionally on the political right.

Table 3 shows how these dimensions of economic, political, and social change map onto the attributes of fair social contracts. One can see how for very large numbers of citizens, there was a factual basis to perceptions of a deteriorating social contract.

[†] The research on these various trends is too voluminous to cite individually, however for various studies and data sources see for example, the University California Berkeley Center for Equitable Growth (<u>http://ceg.berkeley.edu/index.html</u>), Washington Center for Equitable Growth (<u>https://equitablegrowth.org/</u>), the Economic Policy Institute (<u>https://www.epi.org/</u>), and World Inequality Database (<u>https://wid.world/</u>).

Unfair game/broken contract	Example trends
1. Loss of agency	Loss of worker power, de-unionization; loss of local government autonomy, more centralized, less responsive political power
2. Exclusion	Racial and gender barriers; demographic change; cultural alienation; identity politics
3. Loss of dignity, status	Perceived relative status loss (esp. for white, male, working class)
4. Rules violations	Different rules for rich and powerful (e.g., corporate behavior, political capture)
5. Less meritocratic	Lower social mobility, "opportunity hoarding" by top 5%
6. Decreasing security	Weakening social safety net (public and private), greater downward mobility
7. Insufficient capabilities	Declining public investments, education quality, worker training
8. Loss of reciprocity	Decoupling of wages and productivity growth; declining tax fairness
9. Loss of progress, hope	Wage stagnation, declining optimism for future generations

Table 3. Each dimension of the U.S. social contract was broken or weakened for a large proportion of the population during the 1970s-2010s

This then leads to a key conclusion: the U.S. and other similarly affected countries cannot heal political divisions, renew faith in democracy, and reinvigorate collaboration at a national scale, unless the social contract is restored. A key aspect of the psychology of broken contracts is that feelings of contract violation must first be acknowledged and emphasized with before people are willing to listen and engage in contract reconstruction. Populist candidates have succeeded electorally by giving voice to resultant feelings of moral outrage (M. J. Crockett 2017; Brady et al. 2018), promising to fix the violation ("only I can fix it"), and contrasting themselves versus "out of touch elites" who "don't get it." It is essential that political and other leaders who genuinely do want to restore the social contract must first acknowledge it has been broken, show that they hear and empathize with the resulting emotions, and then provide specific solutions to restore the contract that map onto the nine attributes of contract fairness. But if the hypothesis is correct, that the attributes of fair social contracts have high universality, then those nine attributes can provide a template for reducing social divisions and increasing collaboration where there is broad agreement on goals but debate on specific policies. For example, restoring perceptions of reciprocity could be aided by both increased tax fairness (a traditional cause of the left) and welfare system reform (a traditional cause of the right). While increasing agency could be helped both by increasing worker power (a traditional cause of the left) and devolving central political power (a traditional cause of the right). Greater agreement on ends (a fairer social contract) and more constructive debates on means (specific policies) could help facilitate the return of a more functional politics.

Conclusions and Directions for Future Research

Social contract fairness is foundational to large-scale collaboration and social contract unfairness is a key factor in collaboration breakdown. Understanding what leads to perceptions of social contract fairness is amenable to empirical study. This paper has presented an empirically informed hypothesis that there are nine universal dimensions to social contract fairness. This hypothesis is testable. Tools and methods from various disciplines could be brought to bear to prove, disprove, or modify this hypothesis. For example, behavioral experiments could test the willingness of players to collaborate in games that varied in design along the nine dimensions. Other experiments could test the substitutability of the dimensions, as well as individual preference weightings. Or sociological surveys could be used to test perceptions of fairness or unfairness against the dimensions and their universality or variation across individuals and cultures. And organizational studies or anthropological observations could seek to document and analyze social contract designs "in the wild", assessing participant perceptions of fairness against the dimensions. Finally, it would even be possible to imagine field experiments, where social contract terms are varied for different groups to assess the impacts on perceptions, collaborative behaviors, and outcomes.

Findings from such work could yield prescriptive insights for identifying risks for social contract breakdown—the example of U.S. political polarization illustrates the stakes involved when social contract fairness in societal scale collaborations is allowed to degenerate. Practical insights and strategies are needed for social contract repair in many contexts—only through collaboration can we solve our greatest challenges.

Acknowledgements

The author acknowledges the Institute for New Economic Thinking and the Open Society Foundations for funding support and would like to thank participants at The New Institute, Foundations of Value and Values Symposium, Hamburg, Germany, April 2022, as well as the Ernst Strüngmann Forum, How Collabortion Arises and Why it Fails, Frankfurt, Germany, May 2022, for their valuable feedback. Finally, the author thanks Fiammetta Brugo for research assistance.

References

- Adams, J. Stacy. 1965. "Inequity in Social Exchange." Advances in Experimental Social Psychology 2: 267–99. https://doi.org/10.1016/s0065-2601(08)60108-2.
- Akbaş, Merve, Dan Ariely, and Sevgi Yuksel. 2019. "When Is Inequality Fair? An Experiment on the Effect of Procedural Justice and Agency." *Journal of Economic Behavior & Organization* 161: 114–27. https://doi.org/10.1016/j.jebo.2019.02.014.
- Alesina, Alberto, Rafael Di Tella, and Robert MacCulloch. 2004. "Inequality and Happiness: Are Europeans and Americans Different?" *Journal of Public Economics* 9–10 (88): 2009–42.
- Anderson, Elizabeth S. 1999. "What Is the Point of Equality?" *Ethics* 109 (2): 287–337. https://doi.org/10.1086/233897.
- Axelrod, Robert, and William D. Hamilton. 1981. "The Evolution of Cooperation." *Science* 211(4489): 1390–96.
- Bandura, Albert. 1997. Self-Efficacy: The Exercise of Control. Edited by Freeman. New York: Freeman.
- . 2006. "Toward a Psychology of Human Agency." Perspectives on Psychological Science 2 (1): 164–80.
- Batson, C. Daniel, Christopher L. Kennedy, Lesley-Anne Nord, E. L. Stocks, D'Yani A. Fleming, Christian M. Marzette, David A. Lishner, Robin E. Hayes, Leah M. Kolchinsky, and Tricia Zerger. 2007. "Anger at Unfairness: Is It Moral Outrage?" *European Journal of Social Psychology* 37 (6): 1272–85. https://doi.org/10.1002/ejsp.434.

- Baumard, Nicolas, Olivier Mascaro, and Coralie Chevallier. 2012. "Preschoolers Are Able to Take Merit into Account When Distributing Goods." *Developmental Psychology* 48 (2): 492–98. https://doi.org/10.1037/a0026598.
- Bellani, Luna, and Conchita D'Ambrosio. 2011. "Deprivation, Social Exclusion and Subjective Well-Being." Social Indicators Research 104 (1): 67–86. https://doi.org/10.1007/s11205-010-9718-0.
- Binmore, Ken. 1994. *Game Theory and the Social Contract: Just Playing*. Vol. 2. 2 vols. Cambridge, MA: MIT Press.
- Bird, Douglas W., Rebecca Bliege Bird, Brian F. Codding, and David W. Zeanah. 2019. "Variability in the Organization and Size of Hunter-Gatherer Groups: Foragers Do Not Live in Small-Scale Societies." *Journal of Human Evolution* 131 (131): 96–108. https://doi.org/10.1016/j.jhevol.2019.03.005.
- Blake, Peter, K. McAuliffe, J. Corbit, T. C. Callaghan, O. Barry, A. Bowie, L. Kleutsch, et al. 2015. "The Ontogeny of Fairness in Seven Societies." *Nature* 528 (7581): 258--261. https://doi.org/10.1038/nature15703.
- Blake, Peter, and Katherine McAuliffe. 2011. "I Had so Much It Didn't Seem Fair' Eight-Year-Olds Reject Two Forms of Inequity." *Cognition* 120 (2): 215--224. https://doi.org/10.1016/j.cognition.2011.04.006.
- Bock, Jarrod, Jennifer Byrd-Craven, and Melissa Burkley. 2017. "The Role of Sexism in Voting in the 2016 Presidential Election." *Personality and Individual Differences* 119: 189–93. https://doi.org/10.1016/j.paid.2017.07.026.
- Bos, Wouter van den, Eric van Dijk, Michiel Westenberg, Serge A.R.B. Rombouts, and Eveline A. Crone. 2010. "Changing Brains, Changing Perspectives." *Psychological Science* 22 (1): 60–70. https://doi.org/10.1177/0956797610391102.
- Bowles, Samuel, and Herbert Gintis. 2011. A Cooperative Species: Human Reciprocity and Its Evolution. Princeton, NJ: Princeton University Press.
- Boyd, Robert, and Peter J. Richerson. 2005. *The Origin and Evolution of Cultures*. Edited by Oxford University Press.
- Brady, William J., Julian A. Wills, Dominic Burkart, John T. Jost, and Jay J. Van Bavel. 2018. "An Ideological Asymmetry in the Diffusion of Moralized Content on Social Media Among Political Leaders." *Journal of Experimental Psychology: General* 148 (10): 1802–13. https://doi.org/10.1037/xge0000532.
- Brosnan, Sarah F. 2011. "An Evolutionary Perspective on Morality." *Journal of Economic Behavior* & Organization 77 (1): 23–30. https://doi.org/10.1016/j.jebo.2010.04.008.
 - ———. 2013. "Justice- and Fairness-Related Behaviors in Nonhuman Primates." Proceedings of the National Academy of Sciences 110 (Supplement2): 10416--10423. https://doi.org/10.1073/pnas.1301194110.
- Budiman, Abby. 2020. "Key Findings About U.S. Immigrants." Pew Research Center. https://www.pewresearch.org/fact-tank/2020/08/20/key-findings-about-u-s-immigrants/.

- Cassese, Erin C., and Tiffany D. Barnes. 2019. "Reconciling Sexism and Women's Support for Republican Candidates: A Look at Gender, Class, and Whiteness in the 2012 and 2016 Presidential Races." *Political Behavior* 41 (3): 677–700. https://doi.org/10.1007/s11109-018-9468-2.
- Cesarini, David, Christopher T. Dawes, James H. Fowler, Magnus Johannesson, Paul Lichtenstein, and Björn Wallace. 2008. "Heritability of Cooperative Behavior in the Trust Game." *Proceedings of the National Academy of Sciences* 105 (10): 3721--3726. https://doi.org/10.1073/pnas.0710069105.
- Chang, Steve W. C., Nicholas A. Fagan, Koji Toda, Amanda V. Utevsky, John M. Pearson, and Michael L. Platt. 2015. "Neural Mechanisms of Social Decision-Making in the Primate Amygdala." *Proceedings of the National Academy of Sciences* 112 (52): 16012--16017. https://doi.org/10.1073/pnas.1514761112.
- Chen, Ya-Ru, Xiao-Ping Chen, and Rebecca Portnoy. 2009. "To Whom Do Positive Norm and Negative Norm of Reciprocity Apply? Effects of Inequitable Offer, Relationship, and Relational-Self Orientation." *Journal of Experimental Social Psychology* 45 (1): 24–34. https://doi.org/10.1016/j.jesp.2008.07.024.
- Cohn, Alain, E Fehr, B Herrmann, and Frédéric Schneider. 2011. "Social Comparison in the Workplace: Evidence from a Field Experiment," no. 5550: 1--32. http://papers.ssrn.com/sol3/papers.cfm?abstractid=1778894.
- Cook, Karen S., and Karen A. Hegtvedt. 1983. "Distributive Justice, Equity, and Equality." *Annual Review of Sociology* 9 (1): 217–41. https://doi.org/10.1146/annurev.so.09.080183.001245.
- Cox, Daniel, Rachel Lienesch, and Robert P. Jones. 2017. "Beyond Economics: Fears of Cultural Displacement Pushed the White Working Class to Trump." *PRRI/Atlantic Report*. https://www.prri.org/research/white-working-class-attitudes-economy-trade-immigrationelection-donald-trump/.
- Crockett, M. J. 2017. "Moral Outrage in the Digital Age." *Nature Human Behaviour* 1 (11): 769–71. https://doi.org/10.1038/s41562-017-0213-3.
- Crockett, Molly J. 2009. "The Neurochemistry of Fairness." Annals of the New York Academy of Sciences 1167 (1): 76–86. https://doi.org/10.1111/j.1749-6632.2009.04506.x.
- Crockett, Molly J., Zeb Kurth-Nelson, Jenifer Z. Siegel, Peter Dayan, and Raymond J. Dolan. 2014. "Harm to Others Outweighs Harm to Self in Moral Decision Making." *Proceedings of the National Academy of Sciences* 111 (48): 17320--17325. https://doi.org/10.1073/pnas.1408988111.
- D'Agostino, Fred, Gerald Gaus, and John Thrasher. 2021. "Contemporary Approaches to the Social Contract." In *The Stanford Encyclopedia of Philiosophy*, edited by Edward N. Zalta. https://plato.stanford.edu/archives/win2021/entries/contractarianism-contemporary/>.
- Day, Martin V., and Susan T. Fiske. 2017. "Movin' on Up? How Perceptions of Social Mobility Affect Our Willingness to Defend the System." *Social Psychological and Personality Science* 8 (3): 267–74. https://doi.org/10.1177/1948550616678454.

- DeSante, Christopher D., and Candis Watts Smith. 2020. "Fear, Institutionalized Racism, and Empathy: The Underlying Dimensions of Whites' Racial Attitudes." *PS: Political Science & Politics* 53 (4): 639–45. https://doi.org/10.1017/s1049096520000414.
- Diamond, Jared M. 1997. *Guns, Germs, and Steel: The Fates of Human Societies*. Edited by Norton and New York. New York: Norton.
- Dovidio, John F. 1984. "Helping Behavior and Altruism: An Empirical and Conceptual Overview." *Advances in Experimental Social Psychology* 17: 361–427. https://doi.org/10.1016/s0065-2601(08)60123-9.
- Dunbar, Robin. 1992. "Neocortex Size as a Constraint on Group Size in Primates." Journal of Human Evolution 22 (6): 469–93. https://doi.org/10.1016/0047-2484(92)90081-j.
 - ——. 1993. "Coevolution of Neocortical Size, Group Size and Language in Humans." *Behavioral and Brain Sciences* 16 (4): 681–94. https://doi.org/10.1017/s0140525x00032325.
- Dworkin, Ronald. 1981. "What Is Equality? Part 2: Equality of Resources." *Philosophy & Public Affairs*, no. 10: 283–345.
- Edelman. 2020. "Edelman Trust Barometer 2020." New York: Edelman.
- Engel, Robin Shepard. 2005. "Citizens' Perceptions of Distributive and Procedural Injustice During Traffic Stops with Police." *Journal of Research in Crime and Delinquency* 42 (4): 445–81. https://doi.org/10.1177/0022427804272725.
- Fehr, Ernst, Helen Bernhard, and Bettina Rockenbach. 2008. "Egalitarianism in Young Children." *Nature* 454 (7208): 1079–83. https://doi.org/10.1038/nature07155.
- Fehr, Ernst, Urs Fischbacher, and Simon G\u00e4chter. 2002. "Strong Reciprocity, Human Cooperation, and the Enforcement of Social Norms." *Human Nature* 13 (1): 1–25. https://doi.org/10.1007/s12110-002-1012-7.
- Fehr, Ernst, and Simon Gachter. 2017. "Fairness and Retaliation : The Economics of Reciprocity Published by : American Economic Association Stable URL : Http://Www.Jstor.Org/Stable/2646924 Fairness and Retaliation : The Economics of Reciprocity" 14 (3): 159--181.
- Folger, Robert. 1986. "Rethinking Equity Theory: A Referent Cognitions Model." In *Justice in Social Relations*, by In H. W. Beirhoff, R. L. Cohen, and & J. Greenberg (Eds.), 145–162. New York: Plenum Press.
- Folger, Robert, and Russell Cropanzano. 2001. "Fairness Theory: Justice as Accountability." Edited by Stanford University Press. In J. Greenberg & R. Cropanzano (Eds.), Advances in Organizational Justice, 1–55.
- Fong, Christina M., Samuel Bowles, and Herbert Gintis. 2006. "Chapter 23 Strong Reciprocity and the Welfare State." *Handbook of the Economics of Giving, Altruism and Reciprocity* 2: 1439–64. https://doi.org/10.1016/s1574-0714(06)02023-9.
- Fox, John, and Melvin Guyer. 1978. "Public Choice and Cooperation in N-Person Prisoner's Dilemm.Pdf." *Journal of Conflict Resolution* 3 (22): 469–81.

Frey, William H. 2020. "The Nation Is Diversifying Even Faster than Predicted, According to New Census Data." Washington, DC: The Brookings Institution. https://www.brookings.edu/research/new-census-data-shows-the-nation-is-diversifyingeven-faster-than-predicted/.

Gauthier, David P. 1986. Morals by Agreement. Oxford UK: Oxford University Press.

Gintis, Herbert. 2003. "Solving the Puzzle of Prosociality." Rationality and Society 15 (2): 155--187. https://doi.org/10.1177/1043463103015002001.

------. 2004. "The Genetic Side of Gene-Culture Coevolution : Internalization of Norms and Prosocial Emotions" 53: 57--67. https://doi.org/10.1016/s0167-2681(03)00104-5.

——. 2011. "Gene-Culture Coevolution and the Nature of Human Sociality." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 366 (1566): 878--88. https://doi.org/10.1098/rstb.2010.0310.

- Gintis, Herbert, Joseph Henrich, Samuel Bowles, Robert Boyd, and Ernst Fehr. 2008. "Strong Reciprocity and the Roots of Human Morality." *Social Justice Research* 21 (2): 241--253. https://doi.org/10.1007/s11211-008-0067-y.
- Gredebäck, Gustaf, Katharina Kaduk, Marta Bakker, Janna Gottwald, Therese Ekberg, Claudia Elsner, Vincent Reid, and Ben Kenward. 2015. "The Neuropsychology of Infants' pro-Social Preferences." *Developmental Cognitive Neuroscience* 12: 106–13. https://doi.org/10.1016/j.dcn.2015.01.006.
- Green, Jon, and Sean McElwee. 2018. "The Differential Effects of Economic Conditions and Racial Attitudes in the Election of Donald Trump." *Perspectives on Politics* 17 (2): 358–79. https://doi.org/https://doi.org/10.1017/S1537592718003365[Opens in a new window].
- Greenberg, Jerald. 1987. "Reactions to Procedural Injustice in Payment Distributions: Do the Means Justify the Ends?" *Journal of Applied Psychology* 72 (1): 55–61. https://doi.org/10.1037/0021-9010.72.1.55.
 - . 1988. "Equity and Workplace Status: A Field Experiment." *Journal of Applied Psychology* 73 (4): 606–13. https://doi.org/10.1037/0021-9010.73.4.606.

. 1990. "Organizational Justice: Yesterday, Today, and Tomorrow." *Journal of Management*, no. 16.2: 399–432.

- Greene, Joshua D., Leigh E. Nystrom, Andrew D. Engell, John M. Darley, and Jonathan D. Cohen. 2004. "The Neural Bases of Cognitive Conflict and Control in Moral Judgment." *Neuron* 44 (2): 389–400. https://doi.org/10.1016/j.neuron.2004.09.027.
- Gross-Manos, Daphna. 2017. "Material Well-Being and Social Exclusion Association with Children's Subjective Well-Being: Cross-National Analysis of 14 Countries." *Children and Youth Services Review* 80: 116–28. https://doi.org/10.1016/j.childyouth.2017.06.048.
- Guth, W, and R Tietz. 1990. "Ultimatum Bargaining Behavior a Survey and Comparison of Experimental Results." *Journal of Economic Psychology* 11 (3): 417--449.
- Haidt, Jonathan. 2012. The Righteous Mind: Why Good People Are Divided by Politics and Religion. New York: Pantheon.

- Hamilton, Marcus J, Bruce T Milne, Robert S Walker, Oskar Burger, and James H Brown. 2007.
 "The Complex Structure of Huntergatherer Social Networks." *Proceedings of the Royal Society B: Biological Sciences* 274 (1622): 2195–2203. https://doi.org/10.1098/rspb.2007.0564.
- Haruno, Masahiko, and Christopher D. Frith. 2010. "Activity in the Amygdala Elicited by Unfair Divisions Predicts Social Value Orientation." *Nature Neuroscience* 13 (2): 160--161. https://doi.org/10.1038/nn.2468.
- Hawkins, Stephen, Daniel Yudkin, Miriam Juan-Torres, and Tim Dixon. 2019. "Hidden Tribes: A Study of America's Polarized Landscape." More In Common. https://hiddentribes.us/.
- Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, and Herbert Gintis. 2004. Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies. Oxford, UK: Oxford University Press.
- Henrich, Joseph, Jean Ensminger, Richard McElreath, Abigail Barr, Clark Barrett, Alexander Bolyanatz, Juan Camilo Cardenas, et al. 2010. "Markets, Religion, Community Size, and the Evolution of Fairness and Punishment." *Science (New York, N.Y.)* 327 (5972): 1480--4. https://doi.org/10.1126/science.1182238.
- Hooghe, Marc, and Ruth Dassonneville. 2018. "Explaining the Trump Vote: The Effect of Racist Resentment and Anti-Immigrant Sentiments." *PS: Political Science* #x0026; *Politics* 51 (3): 528– 34. https://doi.org/10.1017/s1049096518000367.
- House, Bailey, Joseph Henrich, Barbara Sarnecka, and Joan B. Silk. 2013. "The Development of Contingent Reciprocity in Children." *Evolution and Human Behavior* 34 (2): 86–93. https://doi.org/10.1016/j.evolhumbehav.2012.10.001.
- Inglehart, Ronald, and Pippa Norris. 2017. "Trump, Brexit, and the Rise of Populism: Economic Have-Nots and Cultural Backlash." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2818659.
- Jayawickreme, Eranda, and James O. Pawelski. 2013. "Positivity and the Capabilities Approach." *Philosophical Psychology* 26 (3): 383–400. https://doi.org/10.1080/09515089.2012.660687.
- Killen, Melanie, Kerry Pisacane, Jennie Lee-Kim, and Alicia Ardila-Rey. 2001. "Fairness or Stereotypes? Young Children's Priorities When Evaluating Group Exclusion and Inclusion." *Developmental Psychology* 37 (5): 587–96. https://doi.org/10.1037/0012-1649.37.5.587.
- Kim, Tae-Yeol, and Kwok Leung. 2007. "Forming and Reacting to Overall Fairness: A Cross-Cultural Comparison." Organizational Behavior and Human Decision Processes 104 (1): 83–95. https://doi.org/10.1016/j.obhdp.2007.01.004.
- Konow, J. 2000. "Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions." *American Economic Review* 4 (90): 1072–91.
- Konstantareas, M. Mary, and Nancy Desbois. 2001. "Preschoolers Perceptions of the Unfairness of Maternal Disciplinary Practices." *Child Abuse & Neglect* 25 (4): 473–88. https://doi.org/10.1016/s0145-2134(01)00221-6.
- Kulik, Carol T, and Maureen L Ambrose. 1992. "Personal and Situational Determinants of Referent Choice." *Academy of Management Review* 17 (2): 212–37. https://doi.org/10.5465/amr.1992.4279534.

- Kuwabara, Ko, Robb Willer, Michael W. Macy, Rie Mashima, Shigeru Terai, and Toshio Yamagishi. 2007. "Culture, Identity, and Structure in Social Exchange: A Web-Based Trust Experiment in the United States and Japan." Social Psychology Quarterly 4 (70): 461–479.
- Lazonick, William, and Mary O'Sullivan. 2000. "Maximizing Shareholder Value: A New Ideology for Corporate Governance." *Economy and Society* 29 (1): 13–35. https://doi.org/10.1080/030851400360541.
- Lindsey, Brink, and Steven Teles. 2017. The Captured Economy: How the Powerful Become Richer, Slow Down Growth, and Increase Inequality. Oxford UK: Oxford University Press.
- LoBue, Vanessa, Tracy Nishida, Cynthia Chiong, Judy S. DeLoache, and Jonathan Haidt. 2011. "When Getting Something Good Is Bad: Even Three-year-olds React to Inequality." *Social Development* 20 (1): 154–70. https://doi.org/10.1111/j.1467-9507.2009.00560.x.
- MacDonald, Geoff, and Mark R. Leary. 2005. "Why Does Social Exclusion Hurt? The Relationship Between Social and Physical Pain." *Psychological Bulletin* 131 (2): 202–23. https://doi.org/10.1037/0033-2909.131.2.202.
- Maddison, Angus. 2007. Contours of the World Economy 1-2030 AD: Essays in Macroeconomic History. Oxford, UK: Oxford University Press.
- Mansbridge, Jane. 1999. "Should Blacks Represent Blacks and Women Represent Women?A Contingent 'Yes." The Journal of Politics 3 (61): 628–57.
- Marwell, Gerald, and Ruth E. Ames. 1981. "Economists Free Ride, Does Anyone Else? Experiments on the Provision of Public Goods, IV." *Journal of Public Economics* 15 (3): 295–310. https://doi.org/10.1016/0047-2727(81)90013-x.
- Masten, Carrie L., Naomi I. Eisenberger, Jennifer H. Pfeifer, and Mirella Dapretto. 2010. "Witnessing Peer Rejection during Early Adolescence: Neural Correlates of Empathy for Experiences of Social Exclusion." *Social Neuroscience* 5 (5–6): 496–507. https://doi.org/10.1080/17470919.2010.490673.
- Mccrink, Koleen, Paul Bloom, and Laurie R Santos. 2010. "Children's and Adults' Judgments of Equitable Resource Distributions" 1: 37--45. https://doi.org/10.1111/j.1467-7687.2009.00859.x.
- Milinski, Manfred, Dirk Semmann, and Hans-Jürgen Krambeck. 2002. "Reputation Helps Solve the 'Tragedy of the Commons." *Nature* 415 (6870): 424–26. https://doi.org/10.1038/415424a.
- Moll, Henrike, and Michael Tomasello. 2007. "Cooperation and Human Cognition: The Vygotskianintelligence Hypothesis." *Philosophical Transactions of the Royal Society B: Biological Sciences, 362(1480), 639-648.* 362 (1480): 639–48. https://doi.org/10.1098/rstb.2006.2000.
- Molnar-Szakacs, Istvan. 2011. "From Actions to Empathy and Morality A Neural Perspective." *Journal of Economic Behavior and Organization* 77 (1): 76--85. https://doi.org/10.1016/j.jebo.2010.02.019.
- Moor, Bregtje Gunther, Berna Güroğlu, Zdeňa A. Op de Macks, Serge A.R.B. Rombouts, Maurits W. Van der Molen, and Eveline A. Crone. 2012. "Social Exclusion and Punishment

of Excluders: Neural Correlates and Developmental Trajectories." *NeuroImage* 59 (1): 708–17. https://doi.org/10.1016/j.neuroimage.2011.07.028.

- Musgrave, Stephanie. 2023. "Collaboration in Great Apes." In *How Collaboration Arises and Why It Fails*, edited by Andreas Roepstrorff and Paul Verschure. Ernst Strüngmann Forum Reports. Cambridge, MA: MIT Press.
- Mutz, Diana C. 2018. "Status Threat, Not Economic Hardship, Explains the 2016 Presidential Vote." *Proceedings of the National Academy of Sciences* 115 (19): E4330–39. https://doi.org/10.1073/pnas.1718155115.
- Nagel, Thomas. 1979. *Mortal Questions*. Edited by Cambridge University Press. London: Cambridge University Press.
- Nolan, Brian. 2018. Inequality and Inclusive Growth in Rich Countries. Oxford UK: Oxford University Press.
- Nussbaum, Martha. 2002. "Capabilities and Social Justice." *International Studies Review* 2 (4): 123–35.

2003. "Capabilities as Fundamental Entitlements: Sen and Social Justice." *Feminist Economics* 9 (2–3): 33–59. https://doi.org/10.1080/1354570022000077926.

- Pavey, Louisa, Tobias Greitemeyer, and Paul Sparks. 2011. "I Help Because I Want to, Not Because You Tell Me To." *Personality and Social Psychology Bulletin* 38 (5): 681–89. https://doi.org/10.1177/0146167211435940.
- Pew. 2016. "Partisanship and Political Animosity in 2016." Washington DC: Pew Research Center. https://www.pewresearch.org/politics/2016/06/22/partisanship-and-politicalanimosity-in-2016/?utm_source=Pew+Research+Center&utm_campaign=7732039071-Weekly_June_23_20166_23_2016&utm_medium=email&utm_term=0_3e953b9b70-7732039071-399825545.
 - ——. 2021. "Diversity and Division in Advanced Economies." Pew Research Center. https://www.pewresearch.org/global/2021/10/13/diversity-and-division-in-advancedeconomies/.
- Pleeging, Emma, Martijn Burger, and Job van Exel. 2021. "The Relations between Hope and Subjective Well-Being: A Literature Overview and Empirical Analysis." *Applied Research in Quality of Life* 16 (3): 1019–41. https://doi.org/10.1007/s11482-019-09802-4.
- Putnam, Robert D., and Shaylyn Romney Garrett. 2020. The Upswing: How We Came Together a Century Ago and How We Can Do It Again. New York: Simon & Schuster.
- Rawls, John. 1971. A Theory of Justice. Cambridge, MA: Harvard University Press.
- Robeyns, Ingrid. 2006. "The Capability Approach in Practice." *The Journal of Political Philosophy* 14 (3): 351–376.
- Rodon, Toni, and Marc Sanjaume-Calvet. 2020. "How Fair Is It? An Experimental Study of Perceived Fairness of Distributive Policies." *The Journal of Politics* 82 (1): 384–91. https://doi.org/10.1086/706053.

- Ryan, Richard M., and Edward L. Deci. 2000. "Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being." *American Psychologist* 1 (55): 68.
- Sandel, Michael J. 2020. The Tyranny of Merit: What's Become of the Common Good? New York: Farrar, Straus and Giroux.
- Sasaki, Tatsuya, and Satoshi Uchida. 2013. "The Evolution of Cooperation by Social Exclusion." *Proceedings of the Royal Society B: Biological Sciences* 280 (1752): 20122498. https://doi.org/10.1098/rspb.2012.2498.
- Seabright, Paul. 2010. The Company of Strangers: A Natural History of Economic Life. Princeton, NJ: Princeton University Press.
- Sen, Amartya. 1985. Commodities and Capabilities. Amsterdam: North Holland.
 - -------. 2008. "The Idea of Justice." *Journal of Human Development* 9 (3): 331–42. https://doi.org/10.1080/14649880802236540.
- Shaw, Alex, Peter DeScioli, and Kristina R. Olson. 2012. "Fairness versus Favoritism in Children." *Evolution and Human Behavior* 33 (6): 736–45. https://doi.org/10.1016/j.evolhumbehav.2012.06.001.
- Shaw, Alex, and Kristina R. Olson. 2012. "Children Discard a Resource to Avoid Inequity." Journal of Experimental Psychology: General 141 (2): 382–95. https://doi.org/10.1037/a0025907.
- Shenhav, Amitai, and Joshua D. Greene. 2010. "Moral Judgments Recruit Domain-General Valuation Mechanisms to Integrate Representations of Probability and Magnitude." Neuron 67 (4): 667--677. https://doi.org/10.1016/j.neuron.2010.07.020.
- Shinn, Marybeth. 2015. "Community Psychology and the Capabilities Approach." *American Journal of Community Psychology* 55 (3–4): 243–52. https://doi.org/10.1007/s10464-015-9713-3.
- Singer, Peter. 2005. "Ethics and Intuitions." *Journal of Ethics* 9 (3–4): 331--352. https://doi.org/10.1007/s10892-005-3508-y.
- Smith, Richard. 1995. "Interest Group Influence in the US Congress." *Legislative Studies Quarterly* 20 (1): 89–139.
- Starmans, Christina, Mark Sheskin, and Paul Bloom. 2017. "Why People Prefer Unequal Societies." *Nature Human Behaviour* 1 (4): 1--7. https://doi.org/10.1038/s41562-017-0082.
- Stets, Jan E. 2004. "Theory and Research on Human Emotions." Advances in Group Processes, 51–76. https://doi.org/10.1016/s0882-6145(04)21003-7.
- Tanaka, Toshiko, Takao Yamamoto, and Masahiko Haruno. 2017. "Brain Response Patterns to Economic Inequity Predict Present and Future Depression Indices." Nature Human Behaviour. https://doi.org/10.1038/s41562-017-0207-1.
- Teachout, Zephyr. 2016. Corruption in America: From Benjamin Franklin's Snuff Box to Citizens United. Cambridge, MA: Harvard University Press.

- Tinghög, Gustav, David Andersson, and Daniel Västfjäll. 2017. "Are Individuals Luck Egalitarians? – An Experiment on the Influence of Brute and Option Luck on Social Preferences." *Frontiers in Psychology* 8: 460. https://doi.org/10.3389/fpsyg.2017.00460.
- Tomasello, Michael, and Malinda Carpenter. 2007. "Shared Intentionality." *Developmental Science* 10 (1): 121–25. https://doi.org/10.1111/j.1467-7687.2007.00573.x.
- Trivers, Robert L. 1971. "The Evolution of Reciprocal Altruism." *The Quarterly Review of Biology* 1 (46): 35–57.
- Tuscherer, Taylor, Donald F. Sacco, James H. Wirth, Heather M. Claypool, Kurt Hugenberg, and Eric D. Wesselmann. 2016. "Responses to Exclusion Are Moderated by Its Perceived Fairness." *European Journal of Social Psychology* 46 (3): 280–93. https://doi.org/10.1002/ejsp.2152.
- Ward, George, Jan-Emmanuel De Neve, Lyle H Ungar, and Johannes C Eichstaedt. 2020. "(Un)Happiness and Voting in U.S. Presidential Elections." *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspi0000249.
- Warneken, Felix, and Michael Tomasello. 2013. "The Emergence of Contingent Reciprocity in Young Children." *Journal of Experimental Child Psychology* 116 (2): 338–50. https://doi.org/10.1016/j.jecp.2013.06.002.
- Wegener, Bernd. 1991. "Relative Deprivation and Social Mobility: Structural Constraints on Distributive Justice Judgments." *European Sociological Review* 7 (1): 3–18. https://doi.org/10.1093/oxfordjournals.esr.a036575.
- Williams, Kipling D. 2007. "Ostracism." Annual Review of Psychology 58: 425-452.
- Wootton, David. 2018. Power, Pleasure and Profit: Insatiable Appetites from Machiavelli to Madison. Cambridge, MA: Belknap Press.
- Zak, Paul J. 2011. "The Physiology of Moral Sentiments." *Journal of Economic Behavior and Organization* 77 (1): 53--65. https://doi.org/10.1016/j.jebo.2009.11.009.