Pre-existing fairness concerns restrict the cultural evolution and generalization of inequitable norms in children

Joël Berger[1,2], Sonja Vogt[3,4], Charles Efferson[3]

[1] Department of Social Sciences, University of Bern, Switzerland
[2] Department of Economics, Kalaidos University of Applied Sciences, Switzerland
[3] Faculty of Business and Economics, University of Lausanne, Switzerland
[4] Nuffield College, University of Oxford, UK

Correspondence: joel.berger@soz.unibe.ch, sonja.vogt@unil.ch, charles.efferson@unil.ch; Joël Berger, Fabrikstrasse 8, CH-3012 Bern

Keywords: coordination game, social norms, norm generalization, bargaining

Declarations of interest: none

Abstract

Many social exchanges produce benefits that would not exist otherwise, but anticipating conflicts about how to distribute these benefits can derail exchange and destroy the gains. Coordination norms can solve this problem by providing a shared understanding of how to distribute benefits, but such norms can also perpetuate group-level inequality. To examine how inequitable norms evolve culturally and whether they generalize from one setting to another, we conducted an incentivized lab-in-the-field experiment among kindergarten (5-6) and second-grader (8-9) children (4'228 decisions collected from 326 children living in Switzerland). In Part 1, we created two arbitrarily marked groups, triangles and circles. We randomly and repeatedly formed pairs with one triangle and one circle, and players in a pair played a simple bargaining game in which failure to agree destroyed the gains from social exchange. At the beginning of Part 1 we suggested a specific way to play the game. In symmetric treatments, this suggestion did not imply inequality between the groups, while in asymmetric treatments it did. Part 2 of the experiment addressed the generalization of norms. Retaining their group affiliations from Part 1, each child had to distribute resources between an in-group member and an out-group member. Children of both age groups in symmetric treatments used our suggestions about how to play the game to coordinate in Part 1. In asymmetric treatments, children followed our suggestions less consistently, which reduced coordination but moderated inequality. In Part 2, older children did not generalize privilege from Part 1. Rather, they compensated the underprivileged. Younger children neither generalized privilege nor compensated the underprivileged.

1. Introduction

Many social exchanges increase total welfare relative to the status quo. For example, if Joël has 10 apples and no bread, while Sonja has 3 loaves of bread and no apples, they can both enjoy the benefits of a more balanced diet if they trade. Exchanges of this sort yield novel gains, and in this sense the individuals involved have a shared interest in collaborating. At the same time, exchange usually involves a distributional conflict (Bowles, 2009) and dealing with this conflict can destroy the gains from trade or prevent exchange altogether (Bowles, 2009; De Dreu, Beersma, Steinel, & Kleef, 2007; Raiffa, 2002; Young & Burke, 2001; Young & Raiffa, 1983). In effect, the parties' interests are neither fully congruent nor fully opposed, but they are mixed. Particularly, the parties involved face a coordination problem (Bramoullé, 2007). They can divide up the gains from social exchange in different ways, but they have to agree. Most potential agreements, however, favor one party over the other, and anticipating this tension may prevent agreement altogether.

The cultural evolution of coordination norms can attenuate this problem by creating shared expectations about who gets what. A spectacular example comes from sharecropping. In some regions tenant farmers get one third of the harvest and land owners get two thirds, while in other regions farmers get two thirds and land owners one third (Young, 2013). In either case, the result is a stable disparity between farmers and land owners. Moreover, norms of this sort can be remarkably uniform in local areas. If so, the norm dominates the question of who gets what. Resource distribution is bizarrely insensitive, for example, to the fact that land holdings, and by extension the land owner's bargaining power, vary in terms of soil fertility. In effect, when considering how to distribute, farmers and land owners simply give up, defer to the norm, and get on with the business of growing corn (Young & Burke, 2001). If the local norm is inequitable, however, it can create and perpetuate status differentials that might spill over and affect other types of exchange in other social domains (Bowles, 2009; Henrich & Boyd, 2008; Holm, 2000; Siddique & Vlassopoulos, 2020). To the extent that this happens, the parties involved generalize

3

the norm beyond the original problem, namely how to distribute the gains from sharecropping, to other types of exchange, for example who defers to whom when a farmer and land owner meet in the street?

As a very different example of coordination and inequality, a prominent hypothesis posits that female genital cutting, especially infibulation, is a coordination norm that helps men manage paternity uncertainty (Efferson, Vogt, & Fehr, 2020; Howard & Gibson, 2019; Mackie, 1996; Vogt, Zaid, Ahmed, Fehr, & Efferson, 2016). Once a cutting norm is in place, parents face strong incentives to cut their daughters to improve the future marriage prospects of these daughters (Camilotti, 2016; Efferson, Vogt, Elhadi, Ahmed, & Fehr, 2015; Efferson et al., 2020; Howard & Gibson, 2019; Platteau & Auriol, 2018; Shell-Duncan, Wander, Hernlund, & Moreau, 2011; Vogt et al., 2016). To the extent that such a coordination norm evolves, it shifts the distribution of reproductive benefits away from women and toward men, once again with the potential to entrench status differentials that spill over to other types of social exchange (Henrich & Boyd, 2008).

These examples share the feature that the norm helps the parties involved pick an equilibrium. "Norm" is a rich word with a contentious history (Cialdini, Reno, & Kallgren, 1990; Coleman, 1990; Fehr & Schurtenberger, 2018; Ullmann-Margalit, 1977), but we will use it in this precise limited sense. Specifically, a norm is a common pattern of behavior together with a shared belief that others will tend to adhere to this common pattern behavior. These commonalities combine to help people choose a specific equilibrium and routinely coordinate in a setting with multiple equilibria. This definition corresponds to the notion of a descriptive norm in the sense of Bicchieri (2016; 2005).

Coordinating on a mutual solution increases the welfare of all relative to miscoordinating, and this is true even for the underprivileged party. The farmer and the landowner who agree on a sharecropping arrangement both do better than if the land remained dormant. Still, the solution can be inequitable, as one party is underprivileged relative to another, and inequality can accumulate when underprivilege persists over time and generalizes across contexts.

In spite of these diverse and far-reaching social effects hypothesized to attend cultural evolution under coordination incentives, we do not have an empirical grip on if and how inequitable coordination norms might generalize from one decision-making domain to another after they have evolved culturally in their original domain. Addressing these problems is our main task in the present paper.

In terms of the norm generalization question, the central challenge is methodological. Namely, how do we identify when a norm that evolved in one domain generalizes to shape behavior in another domain? Researchers typically cannot do so without some control over the norm in the primary domain. To illustrate the methodological challenge, imagine the cultural evolution of a sharecropping norm in a local population. To keep the example simple, imagine that only two norms are possible. One norm stipulates that the landowner gets 1/3 of the crop and the farmer 2/3, while the other stipulates that the landowner gets 2/3 and the farmer 1/3. Finally, assume the norm that actually evolves culturally is the latter norm. Our question of interest is, do farmers and landowners generalize landowner privilege in sharecropping to how they interact, say, when they run into each other in town? Imagine that we, as researchers, observe the initial cultural evolution of the sharecropping norm. A year later, we further observe that farmers start to give way and make room for landowners when the two meet in town.

This could occur because both farmers and landowners generalize landowner privilege from sharecropping to social encounters in town, but we cannot conclude this from our observations. Another very different possibility is equally plausible. To illustrate, assume that due to chance self-regarding individuals are over-represented among landowners in this local population, but this is not the case among farmers. The over-representation among landowners is simply ordinary sampling error. Because landowners happen to be more self-regarding on average than farmers, they push their own interests more aggressively than farmers do when the sharecropping norm evolves culturally. Perhaps landowners negotiate more forcefully in the early days of sharecropping before a specific distributional norm has been established. This biases the cultural evolution of the norm in favor of landowners. Then, landowners do exactly the same thing a year later when the norm is evolving for social encounters in town. Landowner privilege does

not generalize. Rather, the extra degree of self-regard among landowners, which is just an accident, separately biases norm evolution in both domains in the same way. Landowners end up privileged in both domains, but not because the privilege in sharecropping agreements generalizes to social encounters in town. Distinguishing this kind of mechanism from norm generalization will typically be difficult or impossible in field settings. In principle, however, a properly designed experiment can provide control over the norm in the first domain, which would then allow the researcher to identify the causal effect of this norm generalizing to the second domain.

Accordingly, to examine the evolution and generalization of distributional norms, we conducted an incentivized lab-in-the-field experiment with children who were either in kindergarten (5-6 years) or second grade (8-9 years). In the first part of the experiment, we randomly divided children into two groups, triangles and circles. In each of 12 periods, we randomly formed pairs consisting of one triangle and one circle, and the players in a pair played a simple bargaining game. This game involved coordination incentives in the sense that the game had two pure-strategy equilibria. To take some control over the coordination norms that evolved under repeated play of the game, we began each session by randomly choosing a descriptive norm (Bicchieri, 2005; Cialdini et al., 1990) suggesting how to play the game, in the following called "suggested equilibrium". In some treatments, this exogenously suggested equilibrium implied inequality, with one group privileged and the other underprivileged. In other treatments, the suggested equilibrium implied equality. In either case, participants were free to choose in all periods, and so the suggested equilibrium was not binding in any way. Rather, it was simply an exogenous prod for the cultural evolutionary process. If the players in an experimental session converge on a specific equilibrium, this is the norm that endogenously evolves within the context of the session. In the second part of the experiment, children retained their group affiliations but played a new game with material incentives explicitly decoupled from the first game. Specifically, each child split a sum of money between two other children, an in-group member and an out-group member. The key question here is whether behavior in this second part of the experiment depended on the first part of the experiment.

This design allowed us to examine the evolution, persistence, and generalization of norms. Specifically, we expected norms to evolve readily and persist when we suggested equilibria that implied equality. However, when the suggested equilibria implied inequality, we expected the kindergarteners to be more likely to follow the suggestion than second-graders. In particular, existing evidence indicates that pre-existing fairness norms should be weaker among kindergarteners than among second-graders (Blake et al., 2015; Peter R. Blake & McAuliffe, 2011; Fehr, Bernhard, & Rockenbach, 2008; Fehr, Glätzle-Rützler, & Sutter, 2013; House et al., 2020, 2013; House & Tomasello, 2018; Kogut, 2012; Smith, Blake, & Harris, 2013).

For this reason, we expected kindergarteners to use an inequitable equilibrium to help coordinate more readily than second-graders. In effect, we expected pre-existing social preferences to be less of a hindrance for kindergarteners than second-graders. In the same way, when faced with inequality from the first part of the experiment, we expected kindergarteners to generalize privilege to the second part of the experiment more readily than second-graders. In contrast, the second-graders should already have internalized societal fairness norms that would prohibit the generalization of inequitable norms from one context to another. The same holds true for adults (House, Kanngiesser, Barrett, et al., 2020; House et al., 2013; House & Tomasello, 2018). It is for this reason that we conducted our experiments with two developmental samples, kindergarteners and second-graders.

The remainder of the article has the following structure. Section 2 elaborates the theoretical background and hypotheses. Section 3 outlines the methods and design. Section 4 presents the results, and section 5 concludes with a discussion.

2. Theory and Hypotheses

2.1 The generalization of norms

Humans routinely encounter novel situations in which they have to make decisions. A straightforward approach in such cases is to apply the culturally evolved norms that shape choices in situations that seem similar to the novel task at hand. Indeed, this kind of norm generalization follows naturally if norms and the values they represent are internalized (Bowles, 1998; Chudek & Henrich, 2011; Kelly & Hoburg, 2017), which is simply a way of saying that social norms become constituents of our culturally evolved preferences.

Experimental studies conducted in small-scale societies around the world provide suggestive evidence that is at least consistent with the generalization of fairness norms. These studies show that, when people find themselves playing an economic game they have never played before, they seem to adapt normative principles from their everyday lives (Henrich et al., 2005, 2010, 2006). For example, as the market integration of a society increases, fairness toward strangers in an experimental game also increases. The explanation offered is that people from societies relatively high on market integration are relatively comfortable trading with strangers, and they transfer associated fairness norms from everyday market settings to the novel economic game at hand. People with little or no market experience, in contrast, do not have such norms to draw on, and self-interest dominates as a result (Henrich et al., 2010).

While these results suggest evidence for the internalization and generalization of fairness norms, drawing inferences about the precise mechanisms at work is challenging for reasons similar to those outlined in the Introduction. One possibility, for example, is that markets create pressure for the cultural evolution of fairness norms, which are then generalized to bargaining

experiments, like ultimatum games, that share key features with market transactions. Another possibility, however, is that some societies, for whatever reason, value fairness, and fairness both facilitates market integration and manifests itself during ultimatum game. The former possibility is about norms that evolve in markets generalizing to affect play in experimental games. The latter possibility is about how variation in market integration and game play both respond to variation in fairness concerns that exist for other reasons.

Cleanly distinguishing between these two scenarios requires an approach in which the researcher manipulates normative information in one decision-making domain to see how this causes an effect in some other domain downstream. This is exactly what we did in our experiment, but this approach introduces another difficulty. Namely, participants do not come to an experiment as vacant recipients of norm manipulations; they come with some mix of norms already in place because of a process of enculturation that has already occurred. Accordingly, manipulating norms in a way that is detectable in the face of these pre-existing forces is potentially difficult, but at least two generic strategies are clear. First, the researcher can implement an exceedingly strong manipulation. This is fine, but in general strong manipulations risk inducing experimenter demand effects (Orne, 1962). Second, the researcher can work with participants that vary in terms of how much enculturation they have experienced, with at least some participants having experienced relatively little enculturation. This is the approach we took by choosing to work with kindergarteners and second-graders as participants.

2.2 Hypotheses

In order to explain our hypotheses, we sketch only the relevant details of our experimental design here. We explain the design in full in section 3. We did not pre-register our design or hypotheses because we collected all data before the recent replication crisis and the trend toward pre-registration it provoked (Aarts et al., 2015; Camerer et al., 2016).

Our experiment randomly divided participants into two arbitrary groups, triangles and circles. Each session involved two parts, where the first part varied by treatment, and the second part did not. In the first part, in each period a randomly paired triangle and circle had to play a simple bargaining game. This game was distinctive in that, at least for the treatments of interest here, a successful agreement required one player to play aggressively and the other play to defer (Henrich & Boyd, 2008). This highlights the problem. Who should play aggressively, the triangle or the circle, and who should defer?

We suggested one equilibrium of the anti-coordination game to the participants that they could potentially use to help them solve this problem by explaining how previous people had played the same game. This suggestion, however, implied a privileged status for one group, either triangles or circles, and an underprivileged status for the other group. Depending on the treatment, the suggested equilibrium in question was based on either the previous play of adults or children. Past research in diverse fields has identified the imitation of role models as a key mechanism for transmitting social norms to children (Ben-Ner, List, Putterman, & Samek, 2017; Blake, Corbit, Callaghan, & Warneken, 2016; Laland & Rendell, 2019; Lew-Levy, Lavi, Reckin, Cristóbal-Azkarate, & Ellis-Davies, 2018; Rushton, 1976). More specifically, while children imitate their peers mostly for social reasons, namely, for bonding, they prefer imitating adults over imitating peers when learning, and, specifically, when learning normative content (Rakoczy, Hamann, Warneken, & Tomasello, 2010; Zmyj & Seehagen, 2013).

At the same time, recent ethnographic and experimental evidence suggesting that the imitation of other children is also crucial and, in some cases, even more important than learning from adults (Gallois, Lubbers, Hewlett, & Reyes-García, 2018; Lew-Levy et al., 2020, 2018; Salali et al., 2019), specifically, for the case of prosocial behavior (Misch & Dunham, 2021). This contrast points toward a key question. Namely, for children in need of information about how to behave, are adults more or less compelling as normative role models than other children? We captured this distinction by manipulating whether we provided information about how adults had played the game previously versus children. Given how central adult role models are in the literature on

norm internalization, we predicted that adult role models would lead to more compliance with the suggested equilibrium than child role models. This is the *adult role model hypothesis*.

The second part of our experiment, which followed immediately after the first part, was a resource allocation game that was identical across all treatment. Specifically, each participant had to allocate resources between an in-group member and an out-group member. Participants retained their group affiliations from the first part of the experiment, but the two parts were otherwise not linked. In particular, the material incentives presented in the two parts of the experiment had no explicit connection to each other.

Our design allowed us to separate out any effects related to in-group favoritism. Our key task centered around identifying how norms from the first part of the experiment shaped allocations in the second part beyond any biases associated with in-group favoritism. In particular, we wanted to examine if players in the allocation game would favor the group privileged in the first part of the experiment, ignore what happened in the first part, or favor the group underprivileged in the first part. The first of these possibilities is our operational definition of norm generalization, while the last possibility represents a form of compensating for past inequalities. Strikingly, recent research has shown that, when adults from different ethnicities play experimental games, participants from the underprivileged ethnic group also play the game in a way that perpetuates their underprivileged status (Siddique & Vlassopoulos, 2020).

With respect to children, we know that fairness norms governing the distribution of resources are highly dependent on culture, and these norms develop starting from middle childhood (House, 2018; House, Kanngiesser, Clark Barrett, et al., 2020; House et al., 2013; House & Tomasello, 2018; Warneken, 2016, 2018). More precisely, it is a universal pattern that before middle childhood, children show no costly altruism (House, Kanngiesser, Barrett, et al., 2020; House et al., 2013). If young children (e.g. 5-6) exhibit social preferences at all, they exhibit aversion against underprivilege, called disadvantageous inequity aversion, but not yet aversion towards privilege, called advantageous inequity aversion (Blake et al., 2015; Blake & McAuliffe,

2011; Fehr et al., 2013). Starting from middle childhood (8-10), children's prosocial sharing increasingly approximates the normative expectations prevailing in a specific society, which also correspond to adult sharing behavior (House, Kanngiesser, Barrett, et al., 2020; House et al., 2013; House & Tomasello, 2018; Kogut, 2012). It is only starting from this age, that advantageous inequity aversion is emerging, and it is emerging only in some societies (Blake et al., 2015). Similarly, only starting from middle childhood, children follow experimental social norm treatments that suggest a generous distribution of resources. In contrast, at ages 4-5, children only react to norm treatments that suggest a selfish distribution (McAuliffe, Raihani, & Dunham, 2017).

At the same time, already 4-5-year old children are capable of coordinating their decisions to produce a mutual benefit (Duguid, Wyman, Grueneisen, & Tomasello, 2020) even when communication is not possible, for example, using majority information (Grueneisen, Wyman, & Tomasello, 2015b) or salience (Grueneisen, Wyman, & Tomasello, 2015a).

Given this evidence, we conjecture that both age groups would be able to coordinate on a suggested equilibrium. At the same time, we conjectured that pre-existing fairness norms in the older age group, the second-graders, would prevent the generalization of any inequitable norms that do arise and lead players to compensate the underprivileged group in the second part of the experiment. This is the *compensation hypothesis.* In contrast, we also conjectured that relatively weak pre-existing fairness norms in the younger age group, the kindergarteners, would lead players to generalize such norms to the second part of the experiment. This means all kindergarteners, whatever their group affiliation, would extend privilege from the first part of the experiment to the second part by giving more to members of the previously privileged group. This is the *generalization hypothesis*. It applies both to those privileged in the first part of the experiment and to those underprivileged. That said, because available evidence suggests that young children are especially unlikely to exhibit aversion to advantageous inequality, we thought that kindergarteners privileged by an inequitable norm would be especially prone to adhere to such a norm and generalize it to new settings.

3. Materials and Methods

3.1 Subjects

We conducted our experiment in 33 kindergarten and primary school classes in the German-speaking part of Switzerland. In total, 326 children participated (169 female, 157 male, see Table 1). These classes were recruited as follows. Headmasters of public schools located in the cities of Zurich, Luzern, and Aarau, or surrounding villages, were asked for permission of inviting kindergarten classes and second grade classes to participate in our study. Although we did not follow a random sampling scheme, our sample comprises classes from a variety of contexts, with respect to regional structure (urban, suburban, rural) and socio-economic status (relatively high, intermediate, low). Despite of pronounced variation, low-status classes are moderately underrepresented, as schools located in low status neighborhoods more frequently self-selected out of the study. The data are available within the Electronic Supplementary Material 2.

The kindergarteners had a mean age of 5.58 and the second-graders a mean age of 8.54. The study was approved by all relevant local educational authorities and by the Human Subjects Committee of the Faculty of Economics, Business Administration, and Information Technology at the University of Zurich. In addition, we received informed consent from the parents of all participating children. No deception was used in the experiment.

3.2 Decision-making tasks and experimental design

The first part of the experiment consisted of an anti-coordination game (e.g. Bramoullé, 2007) that participants played for 12 periods (Fig. 1). By "anti-coordination" game, we simply mean a game in which multiple pure-strategy equilibria exist, as with a coordination game, but these equilibria require players to choose behaviors with different labels. Importantly, an anti-

coordination game is strategically equivalent to a coordination game with the choices for one player relabeled. Because the difference between a coordination and anti-coordination game is simply a question of labeling, we will sometimes refer to an outcome in which paired subjects chose different options as successful "coordination." Similarly, we will also sometimes refer to an outcome in which they chose the same option as "miscoordination." As explained below, the payoff matrix used for the anti-coordination game was one source of treatment variation (Fig. 1), and thus the specific anti-coordination game played varied from one session to another.

In the second part of the experiment, which followed immediately after repeated play of the anti-coordination game, each child divided resources between an in-group member and an out-group member. This allocation game was the same in all treatment conditions, and so it did not vary from one session to another. Altogether, a session consisted 12 periods of an anti-coordination game that varied by treatment, followed by an allocation game that never varied. This paradigm allowed us to isolate the two primary questions of interest. First, how did our treatments shape the coordination norms that evolved during repeated play of a bargaining game with or without the potential to create systematic inequality? Second, to what extent did any coordination norms that evolved spill over to affect choices in a subsequent game with material incentives unrelated to the first game?

Except for two sessions with only eight children available, each session involved 10 participants. At the beginning of a session, the 10 participants were randomly divided into two groups, triangles and circles. In each period of the anti-coordination game, all participants were randomly paired to play the game. Each pair consisted of a triangle player and a circle player, and thus all play was across the group boundary. With random matching by period, each child would necessarily play with at least one partner multiple times. Participants, however, had no identifying information other than their group affiliations, and thus they could not know the specific individual with whom they were playing at any given point in time.

To choose, each child had to select one of two options, "up" or "down." In practice, the options were presented as two baskets, each with some number of gold coins inside. If the players chose different baskets, each would get the gold coins in the basket she chose. If the players chose the same basket, the gold coins disappeared, and neither child received any coins for the period in question. As one treatment dimension, we varied the allocation of gold coins in the baskets (Fig. 1). For the **symmetric** treatment, both baskets contained one gold coin. For the **asymmetric** treatment, the up basket had two gold coins, and the down basket had one.

As a second treatment dimension, we provided players with a suggestion about how to play the anti-coordination game. Specifically, before beginning the main experimental study, we ran multiple sessions in order to observe the evolution of different descriptive norms related to game play, without making use of deception. The individuals taking part in these sessions were different from the individuals taking part in the main study. In some of these sessions, play tended to converge on triangles playing up, while in other sessions play tended to converge on circles playing up. Altogether, these pre-experimental sessions provided us with a pool of sessions from which to draw when implementing different treatments in the main experiment. Specifically, for a given session in the main experiment, we selected a pre-experimental session that tended to converge on either triangles playing up or circles playing up. In the **triangles-up** treatment, we explained that others had already played the same game, and these other players had converged on triangles playing up and circles down. In the **circles-up** treatment, we explained that others had already played the same game, and these other players had converged on circles playing up and triangles down. In effect, we seeded each session with a suggestion about how to play the game, called suggested equilibrium, and in this way we exogenously manipulated which of the two pure-strategy equilibria was focal. Suggested equilibria had the potential to facilitate coordination under both the asymmetric and the symmetric treatments, but under the asymmetric treatment the suggestion had the ancillary effect of privileging one group at the expense of the other group.

a

		Member of Circle group	
		up	down
Member of Triangle group	up	0, 0	2, 1
	down	1, 2	0, 0

b

		Member of Circle group	
		up	down
Member of Triangle group	up	0, 0	1, 1
	down	1, 1	0, 0

Fig. 1. Anti-coordination game. Both versions of the game have two pure-strategy equilibria, with each equilibrium requiring the players to choose different behaviors. In this sense, we have two versions of an "anti-coordination" game. **a** In the asymmetric treatment, successful coordination yields unequal payoffs, with one pure-strategy equilibrium favoring triangles, and the other favoring circles. **b** In the symmetric treatment, in contrast, successful coordination always yields equal payoffs.

For the third and final treatment dimension, we varied the role models we referred to when conveying the suggested equilibrium. In particular, we ran our pre-experimental sessions with either children or adults. This means we had four types of pre-experimental session from which to draw, namely children converging on triangles up, children converging on circles up, adults converging on triangles up, and adults converging on circles up. The distinction between children and adults provided the basis for our final treatment dimension. In the **child-demonstrators** treatment, we communicated the descriptive norm by referring to the fact that children had already played the game. In the **adult-demonstrators** treatment, we referred instead to previous play among adults.

In sum, we assigned treatments to sessions based on a 2 x 2 x 2 experimental design (refer to Table 1 including caption for more information). Altogether, we varied (i) whether the anti-coordination game was **symmetric** or **asymmetric**, (ii) whether the suggested equilibrium was **triangles-up** or **circles-up**, and (iii) whether this norm followed from the previous play of **child demonstrators** or **adult demonstrators**. This design allows detecting effects at conventional levels of power (see Supplementary Fig. E3.).

Table 1. Design and participant numbers by age and treatment

Age Group	Treatment condition		N sessions	N subjects	N obser- vations	% of total sample
Kindergarteners (5-6 years)	Symmetric, adult demonstrators	circle up	2	20	260	6.1
		triangle up	2	20	260	6.1
	Symmetric, child demonstrators	circle up	2	18	234	5.5
		triangle up	3	28	354	8.4
	Asymmetric adult demonstrators	circle up	1	10	130	3.1
		triangle up	2	20	260	6.1
	Asymmetric, child demonstrators	circle up	3	30	390	9.2
		triangle up	2	20	260	6.1
Second-graders (8-9 years)	Symmetric, adult demonstrators	circle up	2	20	260	6.1
		triangle up	2	20	260	6.1
	Symmetric, child demonstrators	circle up	2	20	260	6.1
		triangle up	2	20	260	6.1
	Asymmetric, adult demonstrators	circle up	2	20	260	6.1
		triangle up	2	20	260	6.1
	Asymmetric, child demonstrators	circle up	2	20	260	6.1
		triangle up	2	20	260	6.1
Total			33	326	4228	100

Note: Conditional on implementing sessions with kindergarteners or second-graders we randomly assigned treatments to sessions subject to a stratification scheme that maintained an approximately equal number of sessions per treatment. Observations refer to decisions made in the first part (12 periods) and the second part (1 period) of the experiment, with the exception of Kindergarten/symmetric/child demonstrators/triangle up, where data of part 2 is missing.

3.3 Materials and procedures

To implement the experiment with young children, we developed a child-friendly, interactive video game using the programming language "Processing." For each session, we used multiple laptops (one server, 10 clients) and a wireless network to set up a mobile lab in each participating classroom (Fig. 2a).

Upon arrival, while the children were in a separate area with their teacher, the experimenters set up the computer network for the experiment and a toy store full of toys for incentivizing choices. Then the experimenters joined the children and explained the basics of what would happen. Specifically, the children learned that they would take part in games, and during these games the children could earn gold coins. The experimenter held up a single gold coin at this point and explained its value. At the end of the experiment, each child would have a chance to turn

in her accumulated gold coins for a toy from the toy store (Fig. 2b), with more coins providing a child with the option to choose a toy from a larger selection that included more valuable toys. The least expensive toys in the toy store had values of approximately 10 Swiss Francs (c. 10 USD), and the most expensive toys had values of approximately 30 Swiss Francs. At this point, the experimenter led the children to the toy store, where they observed different groups of toys separated both by value and whether they were targeted at girls or boys. Then, the experimenters covered the toy store with a blanket so the children would not be distracted during the experiment itself.

At this point, some children, specifically those who had been randomly selected to take part in another experiment (see Vogt, Efferson, Berger, & Fehr, 2015), left the room. For children selected to participate in the experiment described here, the experimenter read their names aloud and assigned each child to a specific client laptop. Later, the server for the experiment randomly divided these client locations into the two groups, triangles and circles.

Fig. 2. a A network of 11 laptops (one experimenter/server and 10 subjects/clients) was set up in each classroom. **b** A toy store was set up in the back of the school class for each session. The toy store included age-appropriate toys for girls and boys. We grouped toys into three price categories ranging from 10 to 30 CHF. Thus, the toys effectively had different prices, and the children could freely spend the gold coins they earned during the experiment subject to the constraint that a participant could not buy a toy she could not afford. Even though choices were incentivized in this way, we scaled payoffs and toy prices to ensure that every child was able to choose a toy from the lowest price category.

Once seated, the children learned the rules of the anti-coordination game. To explain the game, one experimenter read instructions aloud in Swiss German. At the same time, a different experimenter manually controlled an animated demonstration of the experiment so that the spoken instructions and the animated demonstration progressed in parallel. In effect, as the children listened to the spoken instructions, they could see a demonstration on the screens in front of them that matched exactly what they were hearing. The animated demonstration was simply a special version of video game used in the experiment. This version was special in that it allowed the experimenter to manually control the scenes depicted on the client screens, and this feature ensured that we could always progress through the spoken instructions and the

demonstration in synchrony. Moreover, the demonstration also acquainted participants with exactly the same screens they would encounter during the experiment proper.

In terms of content, the experimenter and demonstration detailed the following. The children had been divided into two groups, circles and triangles, and each child was a member of one group. Each child was represented on the screen with a single figure, either triangle or circle as appropriate. The figure representing the focal child was surrounded by red lines (Fig. 3a). To play the game once, players would be randomly paired. To represent random pairing, the five circle players spun around like a wheel, while the five triangle players did the same (Fig. 3a). After a few seconds of this, the focal player and her partner zoomed to the foreground, and the remaining players faded away (Fig. 3b). For the game itself, each player had to choose a basket (Fig. 3b). If the two chose different baskets, each would receive the gold coins in their respective baskets (Fig. 3c). If they chose the same basket, neither would receive anything (Fig. 3d). To represent the accumulation of coins over periods, piggy banks would fly in from the side of the screen when appropriate, and a player's figure would automatically throw her gold coins into her piggy bank. In the asymmetric treatment, the upper basket contained two coins, and the lower basket contained one. In the symmetric treatment, both baskets contained one coin.
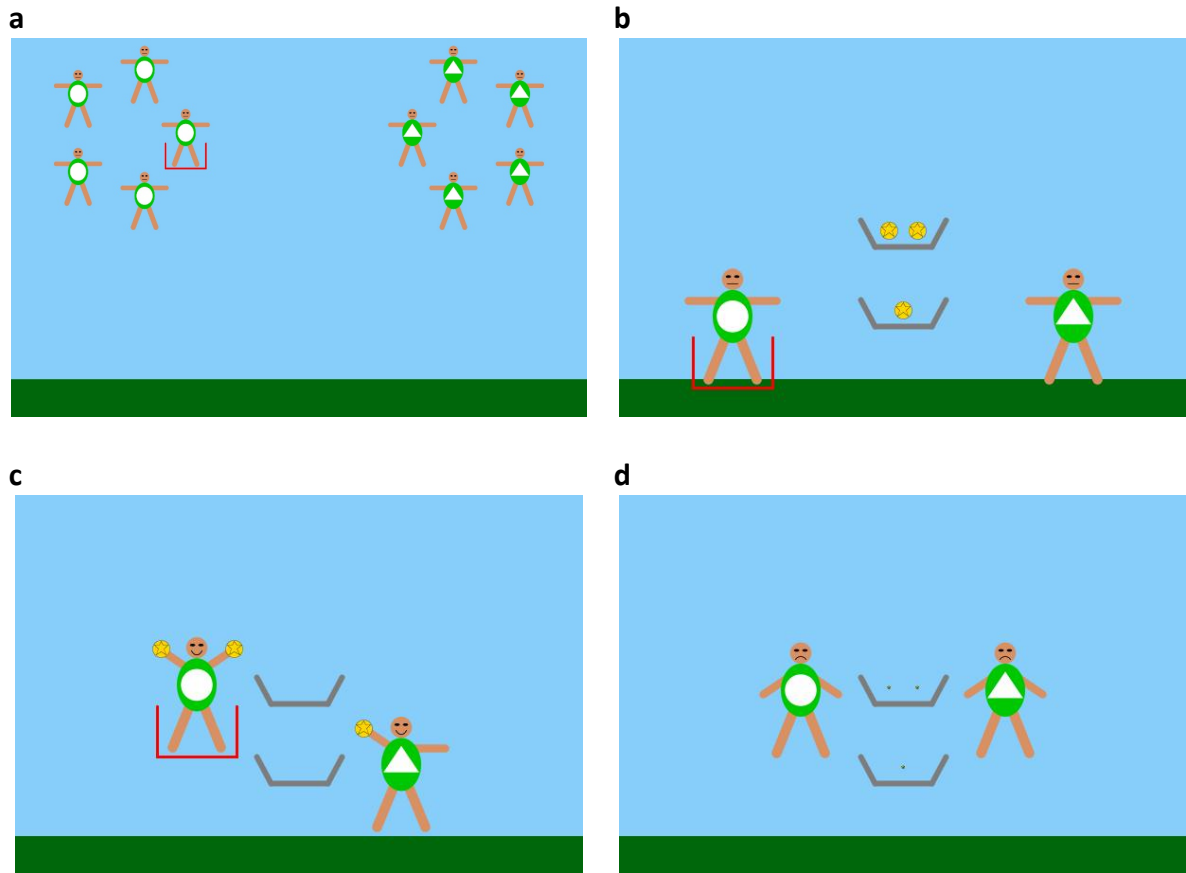
Fig. 3. Anti-coordination game. a At the beginning of each of the twelve periods of the anti-coordination game, the groups (circles and triangles) of players spin in two circles, an animated representation of random partner matching. **b** A focal player (with red lines) and a randomly chosen partner ready to play. **c** Successful coordination, where in this example the circle player chooses up and gets two coins, while the triangle player chooses down and gets one. **d** Miscoordination, where in this example both players choose up and neither player gets any payoff.

During the instructions, the children also learned that they would need to use four computer keys for the experiment. These keys were marked with stickers that looked exactly like graphical elements on the screen. One sticker was positioned high on the keyboard, and it depicted the up basket with either one or two stars, according to treatment. Another sticker was positioned low on the keyboard, and it depicted the down basket, which always had one star. In addition, a sticker on the left side of the keyboard had a circle that looked exactly like the circles on the bodies of the circle players. A sticker on the right of the keyboard had an analogous triangle. Before beginning the anti-coordination game, we checked understanding. In particular, we asked

participants to choose the upper basket, the lower basket, to press the key with the sticker representing the in-group, and to press the key with the sticker representing the out-group. When a participant chose correctly during these tests, a sun appeared on her screen. Otherwise, no sun appeared, and an experimenter would discuss the instructions further with the child. This happened one time per class, on average. Besides a test of whether the children understood how to choose a specific basket, and whether they knew their group-membership, no further comprehension checks were conducted. The coordination game, however, was played repeatedly for twelve periods. We expected that repeated play, combined with regular feedback about the majority behavior in groups would allow the children to develop a clear understanding of the game, and indeed coordination levels above chance confirm that this was the case.

At the very end of the spoken instructions, the experimenter communicated the suggested equilibrium by explaining what had happened in the appropriate pre-experimental session (see Electronic Supplementary Material C for original instructions in Swiss German and D for an English translation). To translate, the experimenter said, "Recently, we played this game with grown-ups/children. The grown-ups/children played the game like this. The triangle (circle) group usually chose the upper basket with one/two gold coins, and the circle (triangle) group usually chose the lower basket with one gold coin." Simultaneously, to visually depict the suggested equilibrium, the screens would show two players choosing in a way consistent with what the experimenter was saying. It is worth noting that the experimenter red out the suggested equilibrium to all participants, therefore all children had common knowledge about the norm, that is they knew that also the others were familiar with the norm.

At this point, the children learned that they would wear headphones, through which they would receive additional information. The facilitator asked the children to put on the headphones and raise their hands if they heard the computer greeting them. After this final check, an audio file would repeat the appropriate descriptive norm with a parallel visualization. The children then learned that they would play 12 rounds of the game, and for each round they would have a randomly selected partner. At the end of each round, each child would also hear recordings stating

whether she had successfully coordinated or not, the coins she and her partner each received, and the most common choice (up versus down) in her own group. This feedback of aggregate group behavior would allow the children to apply frequency-dependent social learning strategies, as predicted by models of cultural evolution (McElreath et al., 2008; Richerson & Boyd, 2008). The children were not allowed to communicate nor could they see the computer screen of any other child. The children were not allowed to communicate nor could they see the computer screen of any other child. We did not collect any verbal communication data from the children. Fig. E1 in the Electronic Supplementary Material depicts a representative client station.

At the end of the 12 periods for the first part of the experiment, the game reported to each subject how many coins she had earned. The game also provided information about the relative performance of the two groups. In symmetric treatments, the two groups necessarily earned the same number of tokens on average, and the computers informed subjects that the two groups had earned equivalent numbers of tokens. In asymmetric treatments, the groups earned different amounts on average, and the game informed subjects which of the two groups had earned more tokens. This latter information was, in effect, a summary of which of the two groups had emerged from the anti-coordination game as privileged and which group had emerged as underprivileged (cf. Henrich and Boyd, 2008). At this point, the children were asked to remove their headphones, and the experimenter explained the resource allocation game in the second part of the experiment.

For the resource allocation game, each subject was asked to distribute three coins between two other children, an in-group partner and an out-group partner (Fig. 4a). We explained that in this new game, each child, as a decision maker, was represented as the figure depicted above the two potential recipients. We then explained that, for each of the three coins, the child would have to choose a recipient, either the circle player or the triangle player (Fig. 4b). Importantly, decision makers only chose how to allocate coins to others. In this sense, we did not consider ego-centered inequity aversion in the resource allocation game (e.g. Fehr & Schmidt, 1999), but rather we examined preferences related to group-level inequality. This kind of resource allocation game is

an established paradigm to measure such preferences (Dunham, Baron, & Carey, 2011; McGuire, Elenbaas, Killen, & Rutland, 2019). Moreover, this game taps a social domain that is distinct from the coordination game. In particular, it centers on a number of isolated individual decisions and entails no strategic consideration as does the anti-coordination game. In the coordination game, ego's payoff is affected by his or her decision. In the resource allocation game, ego's decision does not impact his or her own payoff. As such, the combination of both games is well-suited to capture norm generalization from one domain to another domain, different from the first.
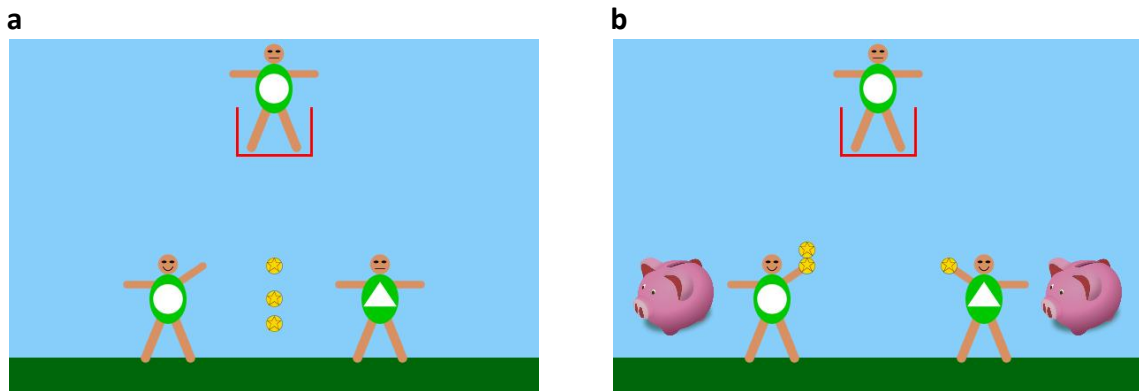

Fig. 4. Resource allocation game. In the second stage of the experiment, each subject split three coins between an in-group member and an out-group member.

4. Results

4.1 Norm compliance in the anti-coordination game.

Fig. 5 depicts compliance with the suggested equilibrium across treatment conditions (see Fig. E2 for coordination across Periods and age groups). In both age groups, compliance with the suggested equilibrium tended to be highest in symmetric treatments (kindergarteners: 0.638 with 95% CI of [0.544,0.722]; second-graders: 0.83 with 95% CI of [0.668,0.922]; relative frequencies averaged over the twelve periods, sample collapsed over child and adult demonstrators, standard errors adjusted for clustering at the session level). Compliance was lowest among those players

underprivileged by the suggested equilibrium in asymmetric treatments (kindergarteners: 0.419 with 95% CI of [0.345,0.496]; second-graders: 0.456 with 95% CI of [0.340,0.578]). Those privileged by the suggested equilibrium in the asymmetric treatment exhibited intermediate levels of compliance (kindergarteners: 0.542 with 95% CI of [0.473,0.608]; second-graders: 0.685 with 95% CI of [0.569,0.782]).
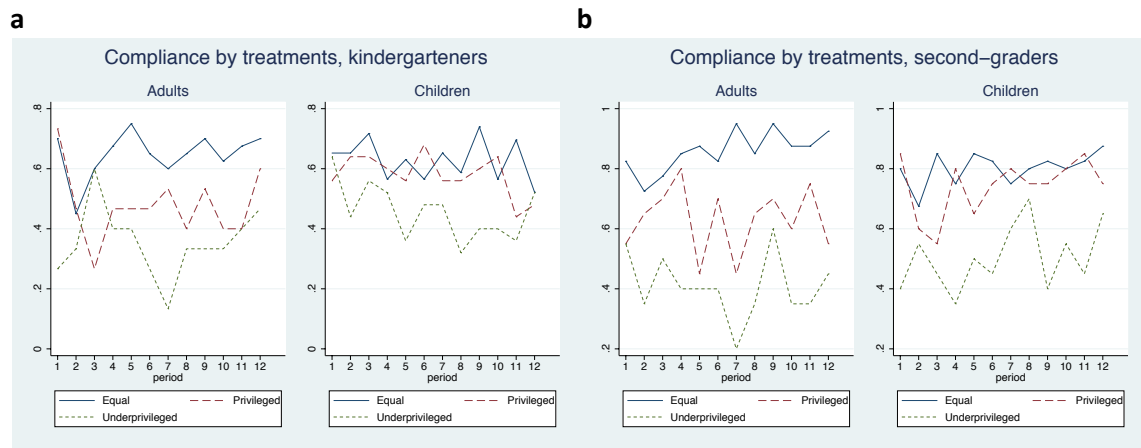

Fig. 5. Compliance in anti-coordination games across treatment conditions. Graphs show the dynamics of the proportion complying by treatment for kindergartners (**a**) and second-graders (**b**). We define compliance as an individual choice in accord with our suggestion about how to play the game at the beginning of Part 1. "Adult" refers to adult demonstrators and "child" to child demonstrators. Each subject is a member of a group, which is, according to this treatment, either privileged (asymmetric), underprivileged (asymmetric), or equal (symmetric). See Table 1 for the numbers of observations by treatment.

Table 2 gives an overview of the predicted compliance rates across treatment conditions for both age groups (predictive margins derived from a full factorial logit model of compliance over all twelve periods). In symmetric treatments, compliance rates were consistently above the 50% we would expect if choices were random. In asymmetric treatments, the 95% confidence intervals of compliance rates consistently overlap with 50%. The compliance in excess of chance in symmetric treatments shows that the children knew how to use the suggested equilibrium as a coordination device. One explanation for lower compliance among the underprivileged, and, to a smaller extent, also among the privileged, is some degree of inequity aversion in both kindergartners and second-graders, with disadvantageous inequity aversion mattering more than

advantageous inequity aversion (Fehr & Schmidt, 1999). An alternative explanation for lower norm adherence in asymmetric treatments is the higher complexity of this treatment – although, the latter explanation fails to account for the fact that the underprivileged adhere less to the suggested equilibrium than the privileged.

Table 2. Compliance over all twelve rounds of the anti-coordination game

	Predictive margins	95%-CI	
Second-grade, child, equal groups	0.825	0.713	0.937
Second-grade, child, unequal groups	0.665	0.492	0.838
Second-grade, adult, equal groups	0.895	0.714	1.076
Second-grade, adult, unequal groups	0.535	0.468	0.602
Kindergarten, child, equal groups	0.622	0.574	0.670
Kindergarten, child, unequal groups	0.472	0.405	0.539
Kindergarten, adult, equal groups	0.670	0.518	0.822
Kindergarten, adult, unequal groups	0.420	0.364	0.476

Note: Predictive margins of compliance derived from a full factorial logit model over the 12 periods of the anti-coordination game. 95% confidence intervals are based on standard errors clustered at the session level. Random play would result in compliance rates of 50%.

To analyze compliance more completely, with a particular focus on the hypothesis that adult demonstrators should yield greater compliance than child demonstrators, we used logistic regressions to examine compliance both in the first period and over all 12 periods. Importantly, in the first period, the suggested equilibrium was the only information participants had to help them coordinate, and so limiting attention to this period provides the purest test of whether adult demonstrators had a relatively strong effect on choices. We also analyze the data over all twelve periods. In this case, the data do not simply represent the effects of the suggested equilibrium; they additionally reflect any effects associated with endogenous social dynamics. It should be noted that with respect to the anti-coordination game, we only have one hypothesis: the adult role model hypothesis. We therefore treat the remainder of these analyses as exploratory.

To recap, our treatment dimensions include adult demonstrators versus child demonstrators, the symmetric treatment versus the asymmetric treatment, and a descriptive norm suggesting triangles up versus circles up. Importantly, the game and the suggested equilibrium jointly defined the link between groups and suggested status. Under the asymmetric treatment, the suggested equilibrium, aside from providing a coordination device, additionally suggested a privileged status for one group and an underprivileged status for the other group. Under the symmetric treatment, the suggested equilibrium implied equal status regardless of which equilibrium we suggested, which meant the suggestion only served as a coordination device. Our task is to examine how compliance rates depended jointly on the demonstrators, suggested relative group status, and the age (kindergarteners versus second-graders) of the participant.

To avoid a profusion of three-way interactions, we work with dummy variables that specify combinations given by the demonstrators and suggested group status. For example, "Adult/privileged" refers to a participant privileged by a suggested equilibrium based on the prior play of adult demonstrators. We additionally include a dummy (Kindergarten) for participant age, and we interact this dummy with all combinations defined over demonstrators and group status. Altogether, this leaves the omitted category as second-graders facing a suggested equilibrium that implies equal status and follows from the prior play of child demonstrators. As controls, we use participant gender and period. We only include period, of course, when analyzing data from all 12 periods, and here we also rely on robust standard errors clustered at the session level. Regarding the participants' behavior in part one, we only formulated one hypothesis. The hypothesis, that adult demonstrators will more effectively create norm adherence – a pattern, that could be more pronounced in the older age group.

Table 3 presents the results. When limiting attention to the first period (Table 3, first and second columns), participants complied with the suggestion significantly less when the game was asymmetric and the suggested equilibrium followed from the prior play of adults (Table 3, Adult/underprivileged, b = -1.186, z = -1.98, P = 0.048, 95% confidence interval (CI) [-2.359, -0.012] and Adult/privileged, b = -1.186, z = -1.98, P = 0.048, CI [-2.359, -0.012]). With a symmetric

treatment and child demonstrators, second-graders exhibited a significantly reduced tendency to comply when underprivileged by the suggested equilibrium (Table 3, Child/underprivileged, b = -1.792, z = -2.97, P = 0.003, CI [-2.975, -0.608]). Importantly, however, this reduction in compliance was significantly attenuated among kindergarteners (Table 3, Kindergarten x Child/underprivileged, b = 1.739, z = 2.18, P = 0.029, CI [0.178, 3.299]). These results reveal a common, but not universal tendency for participants to push back against a suggested equilibrium that implies unequal status.

To test the hypothesis that adult demonstrators should generate more compliance than child demonstrators, we also test the relevant linear combinations from the results in Table 3. In direct contrast to the hypothesis, adult demonstrators significantly reduced compliance in asymmetric treatments in the older age group (Table 3, linear combinations Child/underprivileged vs. Adult/underprivileged, χ^2 = 3.92, P = 0.048; Child/privileged vs. Adult/privileged, χ^2 = 3.92, P = 0.048). Also inconsistent with the hypothesis, when the game was symmetric, compliance did not differ between child and adult demonstrators (Child/equal vs. Adult/equal, χ^2 = 0.08, P = 0. 775). The pattern is similar, although less clear, for the younger age group, with zero effects in equal groups (Child/equal vs. Adult/equal, χ^2 = 0.010, P = 0.94) and in underprivileged groups (linear combinations Kindergarten x Child/underprivileged vs. Kindergarten x Adult/underprivileged, χ^2 = 0.26, P = 0.6102) and a statistically insignificant tendency for lower compliance of the privileged under adult compared to child demonstrators (linear combination Kindergarten x Child/privileged vs. Kindergarten x Adult/privileged, χ^2 = 3.09, P = 0.079). We also implemented a logit regression containing only the dummy 'adult role model', but no interactions between 'adult role model' and the other predictors. We find no main effect for adult role model (b = -0.110, z = -0.45, P = 0.652, CI [-0.587, 0.367]), Table A2 in the electronic supplement).

When analyzing choices over all twelve periods (Table 3, third and fourth columns), we again find a common but not universal tendency for significant reductions in compliance when the game was asymmetric, and by extension the suggested equilibrium implied group-level status

differences (Table 3, Adult/underprivileged, b = -1.771, z = -4.67, P = 0.000, CI [-2.514, -1.027]; Adult/privileged, b = -0.871, z = -2.36, P = 0.018, CI [-1.594, -0.148], and Child/underprivileged, b = -1.383, z = -2.81, P = 0.005, CI [-2.348, -0.418]). The only scenario in which an asymmetric treatment did not significantly reduce compliance was when participants were privileged by a suggested equilibrium based on the prior play of children (Table 3, Child/privileged, b = -0.345, z = -0.66, P = 0.508, CI [-1.367, 0.677]). Unlike the analyses restricted to the first period, there are no significant interactions with the age of participants – a result suggesting that it took longer for kindergarteners than for second-graders to fully comprehend the dynamics of the game. This means that kindergarteners and second-graders, over all twelve periods, responded in the same way to suggested equilibria, and this is true both in terms of a norm's origin (child demonstrators versus adult demonstrators) and in terms of the group status a norm implied (equal, privileged, underprivileged). In addition, as with the analyses based only on period one, the data from all twelve periods provide evidence against the hypothesis that adult demonstrators should hold more sway than child demonstrators. Specifically, testing the relevant linear combinations show that adult demonstrators had no effect on compliance in symmetric treatments (test of linear combinations Child/equal vs. Adult/equal, standard errors clustered at the session-level: χ^2 = 0.29, P = 0.593) and a negative effect in asymmetric treatments (test of linear combinations Child/underprivileged vs. Adult/underprivileged, standard errors clustered at the session-level: χ^2 = 21.77, P = 0.000; Child/privileged vs. Adult/privileged: χ^2 = 5.57, P = 0.018). Finally, the data show no relation between gender and compliance, neither in the first period nor over all twelve periods.

In sum, high compliance rates in symmetric treatments (Table 2) show that both kindergartners and second-graders understood how to use the suggested equilibrium as a coordination device. They often chose, however, to reduce their compliance with the suggested equilibrium in asymmetric treatments, treatments in which compliance produced inequality between the two groups. This reduced compliance was especially robust among subjects underprivileged by the suggested equilibrium (Table 3). Moreover, while the same tendency was present among subjects privileged by the norm, reduced compliance among this subset of

participants is not significant under all model specifications (Table 3). Finally, our data contradict the hypothesis that adult demonstrators should have more influence than child demonstrators. Indeed, adult demonstrators either reduced compliance relative to child demonstrators, or they had no effect (Table 3).

	Compliance in Period 1		Compliance in Periods 1-12	
Kindergarten	-0.758	-0.780	-0.873*	-0.872*
	(-1.51)	(-1.55)	(-2.36)	(-2.36)
Adult/equal	0.164	0.172	0.402	0.402
	(0.29)	(0.30)	(0.53)	(0.53)
Adult/underprivileged	-1.186*	-1.155	-1.771***	-1.773***
	(-1.98)	(-1.92)	(-4.67)	(-4.67)
Adult/privileged	-1.186*	-1.183*	-0.871*	-0.871*
	(-1.98)	(-1.97)	(-2.36)	(-2.37)
Child/underprivileged	-1.792**	-1.849**	-1.383**	-1.380**
	(-2.97)	(-3.04)	(-2.81)	(-2.82)
Child/privileged	0.348	0.370	-0.345	-0.346
	(0.47)	(0.50)	(-0.66)	(-0.66)
Kindergarten x Adult/equal	0.0544	0.0602	-0.319	-0.319
	(0.07)	(0.08)	(-0.39)	(-0.39)
Kindergarten x Adult/underprivileged	-0.455	-0.471	0.649	0.650
	(-0.51)	(-0.53)	(1.61)	(1.62)
Kindergarten x Adult/ privileged	1.569	1.571	0.255	0.256
	(1.76)	(1.76)	(0.63)	(0.64)
Kindergarten x Child/underprivileged	1.739*	1.847*	0.683	0.676
	(2.18)	(2.30)	(1.30)	(1.29)
Kindergarten x Child/privileged	-0.736	-0.764	0.141	0.143
	(-0.82)	(-0.85)	(0.26)	(0.26)
Period			0.00873	0.00873
			(0.83)	(0.83)
Female		0.282		-0.0197
		(1.12)		(-0.21)
Constant	1.386***	1.258**	1.343***	1.352***
	(3.51)	(3.06)	(3.76)	(3.74)
N	326	326	3912	3912

Table 3. Compliance with the suggested equilibrium. Binary logit model. Compliance (1) is the dependent variable. *N* refers to either (i) the number of participants for analyses restricted to the first period (columns 1 and 2) or (ii) to the total number of decisions for analyses using data from all 12 periods (columns 3 and 4). For models analyzing all periods, robust standard errors are based on

clustering at the session level. *z* statistics in parentheses. See Table A1 in the electronic supplement for similar results from an analysis based on clustered bootstrapping. * p < 0.05, ** p < 0.01, *** p < 0.001.

4.2 Generalization based on exogenous versus endogenous privilege

Norm generalization predicts that, among kindergarteners, privileged status in the asymmetric treatment becomes normatively acceptable and thus spills over to shape behavior in the subsequent resource allocation game. We examine this idea in two different ways based on two different operational definitions of privilege. To foreshadow the key distinction, one definition of privilege is based on the equilibrium we suggested at the beginning of Part 1. Because we exogenously manipulated the suggested equilibrium by session, the associated definition of privilege allows a causal analysis. However, as indicated by the above analyses of compliance, participants did not always adhere to the suggested equilibrium. In asymmetric treatments, this raises the possibility that the group favored by the suggested equilibrium actually ended up earning less on average than the disfavored group. In this way, realized privilege was endogenous, and the associated definition of privilege only allows a correlational analysis. It should be noted that in Part 2, our hypotheses (compensation hypothesis and generalization hypothesis) originally referred to exogenous privilege. Although the hypotheses extend to endogenous privilege, we treat the corresponding analyses as exploratory.

With respect to exogenous privilege, we provided a suggestion about how to play the anti-coordination game at the beginning of each session. Under asymmetric payoffs, this suggested equilibrium implied privileged status for one group and underprivileged status for the other. This form of privilege is exogenous in the sense that we assigned privileged status by session following a stratification scheme (Fig. 6a). As explained above, compliance rates under the symmetric treatment were significantly above chance for both kindergarteners and second-graders. Under asymmetric payoffs, however, which is our primary interest when analyzing spillovers, compliance did not typically differ from chance because both the privileged and underprivileged rebelled against our suggestion for how to play. Nonetheless, we can still analyze if exogenous

privilege in the anti-coordination game predicts allocations in the resource allocation game. If so, the estimated effect would be causal. Altogether, exogenous privilege, or exogenous (group) status, refers to our experimental manipulations, which implied either equal status under the symmetric treatment or a privileged group and an underprivileged group under the asymmetric treatment.

With respect to endogenous privilege, or endogenous (group) status, under asymmetric payoffs one group always ended up earning more on average than the other group. We told participants which of the two groups was privileged in this sense at the end of the final period of the anti-coordination game. We did not tell participants how much more the privileged group earned on average than the underprivileged group. Rather, we only reported which group came out on top. We refer to this as endogenous privilege (Fig. 6b) because privileged status followed from actual choices rather than something we manipulated as experimenters. By extension, when we analyze the effect of endogenous privilege in the anti-coordination game on play in the resource allocation game, the estimated effect is correlational. Altogether, endogenous privilege captures three different types of group status. Either the groups played the symmetric game, in which case they were necessarily equal, or they played the asymmetric game, in which case endogenous choices yielded a privileged group and an underprivileged group. Refer to Supplementary Tables B4 and B5 for more information about the bivariate distribution of exogenous and endogenous group status.

4.3 Decisions in the resource allocation game

Across treatment conditions and age groups, the average transfer to the in-group partner was 2.09, which suggests the existence of a kind of baseline in-group-favoritism (see Fig. 6 for treatment-specific average transfers). More interestingly, do spillovers from the anti-coordination game also affect choice in the resource allocation game? To answer this question, we ran OLS regressions with the number of coins allocated to the in-group member as the dependent variable and the same set of predictors as in our analyses of compliance. Of main interest are the interaction terms kindergarten x adult/underprivileged, kindergarten x adult/privileged,

kindergarten x child/underprivileged, kindergarten x child privileged, as these terms will tell us, whether the younger age group generalized inequity from the anti-coordination game to the resource-allocation game (generalization hypothesis), while the older age group did not (compensation hypothesis). As a power analysis shows our models are sufficiently powered to detect effects at conventional levels of power (Electronic Supplementary Tables B6 and B7).
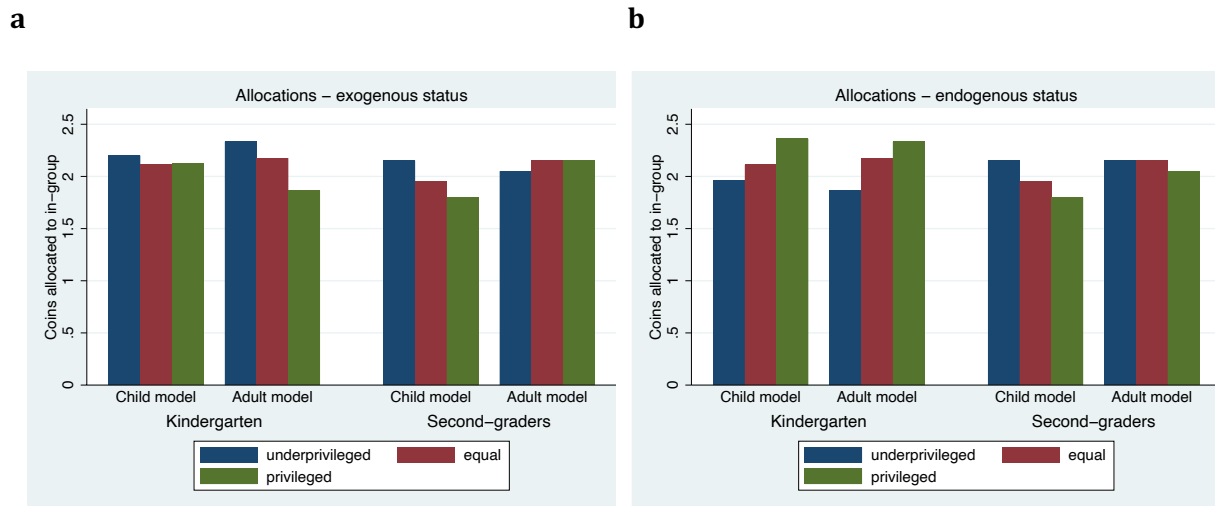
a

b

Fig. 6. Average number of coins allocated to the in-group member by treatment condition and age group. a Exogenous group status. **b** Endogenous group status.

To be more specific, we ran two sets of regressions, one with exogenous status as a predictor (Table 4, first and second columns) and the other with endogenous status (Table 4, third and fourth columns). We clustered the standard errors at the session level. Both models confirm that participants exhibited a baseline degree of in-group favoritism. This follows from the fact that in all of the four regressions the intercepts are significantly different from 1.5, which is the expected in-group allocation in the absence of any preference for favoring one group or the other (F-test of constant a = 1.5, standard errors clustered at the session-level, Table 4, first column: a = 1.95, F = 27.54, P = 0.000; second column: a = 2.01, F=27.83, P = 0.000; third column: = 1.95, F = 27.54, P = 0.000; fourth column: a = 2.00, F = 27.84, P = 0.000). The presence of in-group favoritism shows that the children understood the game and did not allocate coins randomly. When modeling in-group allocations as a function of exogenous status, we find only one significant effect.

Namely, relative to the omitted category, participants from the older age group allocated more coins to the in-group member when assigned an underprivileged status based on the prior play of children (Table 4, Child/underprivileged, b = 0.200, t = 2.07, P = 0.047, CI [0.003, 0.397]). In this case, participants were members of the group underprivileged by the suggested equilibrium. In turn, when given the opportunity in the allocation game, these participants compensated their group affiliates, who shared their underprivileged status. This result, of course, is exactly the opposite of norm generalization.

Concerning endogenous status, as before we find a positive effect associated with Child/underprivileged (b = 2.00, t = 2.07, P= 0.047, CI [0.003, 0.397]). In addition, we find a positive effect associated with Adult/underprivileged (b = 2.00, t = 2.07, P= 0.047, CI [0.003, 0.397]). These effects show that in the older age group, the underprivileged counteract the underprivileged status of their in-group members – a behavioral pattern that can be explained with a combination of disadvantageous inequity aversion and in-group favoritism and is in line with our compensation hypothesis. In contrast, among the kindergarteners privileged groups and underprivileged groups do not significantly differ from equal groups when allocating coins. Interestingly, the qualitative pattern is the opposite of that among second-graders, but the evidence for generalizing endogenous privilege among kindergarteners was not significant (test of linear combinations, standard errors clustered at the session-level, Kindergarten x Child/privileged vs. Kindergarten x Child/equal: F = 2.57, P=0.119; Kindergarten x Adult/privileged vs. Kindergarten x Adult/equal: F = 1.23, p = 0.2769; see Tables B1 and B2, Electronic Supplementary Material, for more information). Altogether, endogenously underprivileged second-graders compensated their underprivileged group affiliates, while kindergarteners neither compensated for nor generalized endogenous status. These results are consistent with the notion that development of internalized fairness norms is still underway between the ages of 5 and 8. They also support our compensation hypothesis, but not our generalization hypothesis. As discussed, our analyses of endogenous status are exploratory and the corresponding results should therefore be interpreted with suitable caution.

As a robustness check, given the relatively limited number of clusters, we replicated the regression model of allocations with bootstrapped standard errors clustered at the session level. Again, we find a tendency for the underprivileged to allocate more to the in-group. With bootstrapping, however, the results are somewhat less stable (see Electronic Supplementary Table B3).

In sum, when modeling in-group allocations as a function of exogenous status, we find that second-graders allocated more coins to the in-group member when assigned an underprivileged status based on the prior play of children. When modeling in-group allocations as a function of endogenous status, we find that second-graders allocated more coins to the in-group members when assigned an underprivileged status based on the prior play of children or adults, although with respect to adults this effect does not reach statistical significance under all model specifications. In the younger age group, we find neither the generalization of privilege nor compensation for being underprivileged for either specifications of group status.

	Exogenous group status		Endogenous group status	
	(1)	(2)	(3)	(4)
Kindergarten	0.161	0.164	0.161	0.164
	(1.40)	(1.43)	(1.40)	(1.43)
Adult/equal	0.200	0.197	0.200	0.197
	(1.73)	(1.69)	(1.73)	(1.70)
Adult/underprivileged	0.1000	0.0837	0.200*	0.186
	(0.63)	(0.53)	(2.07)	(1.92)
Adult/privileged	0.200	0.197	0.1000	0.0972
	(1.65)	(1.65)	(0.57)	(0.55)
Child/underprivileged	0.200*	0.223*	0.200*	0.203*
	(2.07)	(2.41)	(2.07)	(2.05)
Child/privileged	-0.150	-0.160	-0.150	-0.142
	(-0.81)	(-0.89)	(-0.81)	(-0.77)
Kindergarten x Adult/equal	-0.136	-0.133	-0.136	-0.133
	(-0.67)	(-0.64)	(-0.67)	(-0.65)
Kindergarten x Adult/underprivileged	0.122	0.134	-0.444	-0.427
	(0.66)	(0.71)	(-1.32)	(-1.31)
Kindergarten x Adult/privileged	-0.444	-0.437	0.122	0.121
	(-1.29)	(-1.33)	(0.61)	(0.60)
Kindergarten x Child/underprivileged	-0.111	-0.152	-0.351	-0.356
	(-0.45)	(-0.62)	(-1.42)	(-1.48)
Kindergarten x Child privileged	0.159	0.176	0.399	0.384
	(0.56)	(0.62)	(1.60)	(1.55)
Female		-0.130		-0.113
		(-1.59)		(-1.54)
Constant	1.950***	2.012***	1.950***	2.003***
	(22.74)	(20.73)	(22.74)	(21.00)
N	316	316	316	316

Table 4. Allocation to in-group. OLS regression. Sum of coins allocated to the in-group depending on group status (Models 1 and 2: exogeneous group status. Models 3 and 4: endogenous group status). Cluster-robust standard errors. N refers to the number of participants. z statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

5. Discussion

We used a computerized lab-in-the-field experiment to examine the cultural evolutionary processes shaping the emergence and generalization of coordination norms in children. To do so, we developed an accessible video game that allowed young children to repeatedly play an incentivized anti-coordination game. Using this novel approach, we find that both kindergarteners and second-graders used information about how others had played the game, information we manipulated exogenously (called "suggested equilibrium"), to improve coordination. Specifically, the subjects in the control version with symmetric payoffs consistently complied with the suggested equilibria. Importantly, because payoffs were symmetric, triangles and circles earned equal payoffs when choosing different baskets, and in this sense following the suggested equilibria was cost-free for everyone.

In contrast, in the asymmetric treatment, the equilibria we suggested implied a way of playing the game that would have created one privileged group and one underprivileged group. Interestingly, in this case both groups disregarded the suggestion we provided, which in turn limited the extent to which robust coordination norms formed and persisted.

As an alternative explanation to consciously disregarding the inequitable suggestion, the children might have had more difficulty in understanding the asymmetric version than the symmetric version of the anti-coordination game. Although we tested whether the children were able to select on their computers each of the two strategies available, we did not test whether they comprehended the strategic nature of the anti-coordination game. It is, therefore, possible that some of the subjects, for example, those in asymmetric treatments, or the kindergarteners, had more difficulty than others to understand the dynamics of the game. In that case, differences in comprehension might have led to differences in play. Although we cannot completely rule out this possibility, the following pattern, found in asymmetric treatments, suggests that comprehension was generally good. Namely, those privileged by the norm showed a stronger tendency to adhere to this norm than the underprivileged. Even though less pronounced in kindergarteners, this finding holds for both age groups.

Even though participants to some extent spurned suggestions about how to play to the asymmetric game, at the end of the first part of the experiment, one group (i.e. triangle or circle) had nonetheless always accumulated a higher average payoff than the other group. In this way, one group was endogenously privileged, while the other was endogenously underprivileged. In the resource allocation game, which followed the first part of the experiment, second-graders in the endogenously underprivileged group compensated for this disparity. They did so by giving more to their group affiliates, and in particular this in-group favoritism exceeded that present in control treatments with symmetric payoffs and hence no group-level inequality. Interestingly, this behavior was the exclusive domain of underprivileged second-graders. Neither second-graders from the endogenously privileged group nor kindergartners compensated for underprivileged status in the same way.

In addition, no group of participants generalized privilege, whether exogenous or endogenous, from the anti-coordination game with asymmetric payoffs to the subsequent resource allocation game. In particular, this result follows from two different but related phenomena. First, under asymmetric payoffs, we provided an exogenous suggestion about how to play the game, and doing so implied an advantage for one group but not the other. This exogenous advantage, however, did not lead to an advantage for the same group in the resource allocation game. Moreover, regardless of whether participants followed our suggestion, repeated play of the asymmetric game meant that one group had higher average payoffs than the other group after 12 periods. This endogenous advantage also did not lead to an advantage for the same group in the resource allocation game. All in all, though some participants compensated for endogenous underprivilege, no group of participants perpetuated advantage, whether exogenous or endogenous, by carrying it over from the asymmetric anti-coordination game to the resource allocation game.

While previous studies have found that preschoolers favor in-group members over out-group members (Benozio & Diesendruck, 2015; Fehr et al., 2013) and that kindergarteners mitigate inequality when they are underprivileged (Blake & McAuliffe, 2011; House et al., 2013), our study is the first to show that children around the ages of 5-6 and 8-9 forego received privilege

by limiting adherence to an exogenous but inequitable norm that could serve as a coordination device. In addition, at least for second-graders, they compensated endogenously underprivileged group affiliates at a level above and beyond simple in-group favoritism. A recent study with a coordination game involving asymmetric payoffs came to a very different conclusion and found that children aged eight frequently used inequitable norms to coordinate and generate inequitable payoffs (Grueneisen & Tomasello, 2019). Crucially, the children in this study were allowed to talk in face-to-face interactions, while our study did not allow communication. Indeed, all interactions were computerized and therefore anonymous in our case. This methodological difference suggests that impression management, already a driver of cooperative behavior among five-year-olds (Rapp, Engelmann, Herrmann, & Tomasello, 2019), could be a plausible explanation for the divergent findings.

The result that both kindergarteners and second-graders disregard the inequitable normative suggestion in the first part of the experiment is consistent with inequity aversion, with aversion to disadvantageous inequality looming larger than aversion to advantageous inequality (Fehr & Schmidt, 1999). We expected some inequity aversion of this sort, but we expected it to be relatively weak among kindergarteners. The results from our repeated asymmetric anti-coordination game do not support this idea. However, in the resource allocation game in the second part of the experiment, only the second-graders actively compensated their underprivileged group affiliates. This is consistent with our expectation that the older children would come to the experiment with more strongly developed fairness norms than the kindergarteners. Interestingly, the difference here could arise from the locus of inequality in the two parts of the experiment. Specifically, in the anti-coordination game, choices potentially created inequality with respect to a focal decision maker and her partner. In this sense, fairness pertained to ego-centered inequality. In the resource allocation game, the decision maker allocated coins to an in-group member and an out-group member. The decision maker's choices did not affect her own payoffs, and in this sense fairness pertained to group affiliation without being ego-centered. This suggests the possibility that ego-centered fairness concerns might develop before group-oriented fairness concerns. If so, it would potentially explain why

kindergarteners pushed back against inequitable norms in the first part of the experiment, but they did not compensate the underprivileged in the second part.

At this point, it is worth noting that differences in play between the two age groups could, in principle, also be an artifact of the design. Should, for example, the kindergarteners have had more difficulty understanding the rules of a game or the compensation exchange system, this could also explain the compensation of underprivilege among second-graders and the absence of compensation among kindergarteners. Qualitative evidence mitigates this concern. A majority of both age groups played both games with confidence, ad seldom did a participant seem confused. Concerning the compensation exchange system in particular, before the children could trade their coins for toys, participants of both age groups frequently asked the facilitators whether they had earned a sufficient number of coins to buy a certain toy. This shows that they clearly understood that it was the number of coins earned that determined which toy they could get.

Our findings also address the important distinction between adults versus children as role models. The literature on role models and social learning suggests that adult models are more important than child models (Rakoczy et al., 2010; Zmyj & Seehagen, 2013). More specifically, adults have accumulated knowledge from which children can typically benefit, and in this sense we might expect the evolution of social learning strategies that preferentially attend to adults (Kelly & Hoburg, 2017; Richerson & Boyd, 2020). Interestingly, recent observational data from small-scale societies (e.g., Lew-Levy et al., 2019), and recent experimental data (Misch & Dunham, 2021) suggest that with respect to cultural learning in general and norm internalization in particular, the imitation of other children plays an important role. In line with this recent research, our experimental data highlights the importance of children as role models. More precisely, we find that children as role models can be even more effective in transmitting normative information than adults. Future research could investigate the conditions under which adult role models are important in transmitting normative information and when children are important. For example, is it a question of normative content, circumstance, both?

We started with the notion that when bargaining coordination incentives can create pressure for individuals to coalesce around shared expectations. Once shared expectations are in

place, they provide people with a coordination device in the form of a norm, which limits miscoordination and improves welfare. At the same time, a norm can privilege some groups at the expense of other groups, and alternative norms can differ precisely in terms of which groups benefit. In this sense, when a population is structured into groups according to ethnicity, gender, location, or occupation, the groups in question can have very different preferences over the set of feasible norms. Tenant farmers would prefer a norm that favors them over a norm that favors landowners; landowners presumably have the opposite preference.

Crucially, this kind of heterogeneity can have dramatic consequences for the cultural evolution of norms (Berger et al., 2021; Efferson et al., 2020; Granovetter, 1978; Schelling, 1978), and the intuition is straightforward. Coordination incentives and conformity create pressure to coalesce around a share understanding of the right way to behave conditional on the group one belongs to. Heterogeneity does the opposite. The rest is a question of details related to the cost of miscoordinating, the structure of the heterogeneity in the population, pre-existing power relations, and both random and planned shocks that occur during the cultural evolutionary process. These details are decisive in terms of whether a norm forms and stabilizes (Berger et al., 2021; Efferson et al., 2020), and our results are consistent with this idea.

In particular, our symmetric treatment was effectively a way of instantiating preference homogeneity, and norms readily evolved and stabilized. This shows that even young children making decentralized choices can quickly recognize and embrace a new norm when heterogeneity is not getting in the way. Our asymmetric treatment, however, instantiated a form of heterogeneity that did get in the way, and this simple change destabilized norm adherence and persistence. In addition, second-graders compensated the underprivileged when given the chance in the second part of the experiment. This finding further indicates that pre-existing fairness norms were in conflict with the norm we suggested. Variation in the tendency to import norms in this way from outside the task at hand could be one important source of heterogeneity in cultural evolutionary processes.

Altogether, our findings support the pivotal role of heterogeneity in cultural evolution. Existing research shows that cultural evolutionary processes can structure variation to create a

kind of normative signature, with less variation within groups and more variation between groups than we might otherwise expect (Bell, Richerson, & McElreath, 2009; Eugster, Lalive, Steinhauer, & Zweimüller, 2011; Handley & Mathew, 2020; Lowes, Nunn, Robinson, & Weigel, 2017). That said, we also know that individuals vary, and in particular empirical research on cultural transmission clearly shows that individuals vary in terms of how they learn from others (Efferson, Lalive, Richerson, McElreath, & Lubell, 2008; Goeree & Yariv, 2015; Mesoudi, Chang, Dall, & Thornton, 2016; Muthukrishna & Henrich, 2016). Heterogeneity, especially individual heterogeneity within groups, works against the normative signature. Broadly speaking, we have evidence for two countervailing forces, and the net effect could easily vary from one setting to another. Advancing our understanding of the cultural evolution of norms will involve an extended examination of these issues. Specifically, under what conditions does heterogeneity prevent the evolution, persistence, and generalization of norms, and under what conditions do norms dominate in spite of the heterogeneity that is, somehow, inevitably present?

References

Aarts, A. A., Anderson, J. E., Anderson, C. J., Attridge, P. R., Attwood, A., Axt, J., … Zuni, K. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716

Bell, A. V., Richerson, P. J., & McElreath, R. (2009). Culture rather than genes provides greater scope for the evolution of large-scale human prosociality. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(42), 17671–17674. https://doi.org/10.1073/pnas.0903232106

Ben-Ner, A., List, J. A., Putterman, L., & Samek, A. (2017). Learned generosity? An artefactual field experiment with parents and their children. *Journal of Economic Behavior and Organization*, *143*, 28–44. https://doi.org/10.1016/j.jebo.2017.07.030

Benozio, A., & Diesendruck, G. (2015). Parochialism in preschool boys' resource allocation. *Evolution and Human Behavior*, *36*(4), 256–264. https://doi.org/10.1016/j.evolhumbehav.2014.12.002

Berger, J., Efferson, C., Vogt, S., Efferson, C., Vogt, S., & Efferson, C. (2021). Tipping pro-environmental norm diffusion at scale – Opportunities and limitations. In *Working Paper, University of Bern*. Bern.

Bicchieri, Christina. (2016). *Norms in the wild: How to diagnose, measure, and change social norms*. Oxford: Oxford University Press.

Bicchieri, Cristina. (2005). The grammar of society: The nature and dynamics of social norms. In *The Grammar of Society: The Nature and Dynamics of Social Norms*. https://doi.org/10.1017/CBO9780511616037

Blake, P. R., Corbit, J., Callaghan, T. C., & Warneken, F. (2016). Give as I give: Adult influence on children's giving in two cultures. *Journal of Experimental Child

Psychology, *152*, 149–160. https://doi.org/10.1016/j.jecp.2016.07.010

Blake, P. R., & McAuliffe, K. (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

Blake, P. R., McAuliffe, K., Corbit, J., Callaghan, T. C., Barry, O., Bowie, A., … Warneken, F. (2015). The ontogeny of fairness in seven societies. *Nature*, *528*(7581), 258–261. https://doi.org/10.1038/nature15703

Bowles, S. (1998). Endogenous Preferences: The Cultural Consequences of Markets and other Economic Institutions. *Journal of Economic Literature*, *36*(1), 75–111.

Bowles, S. (2009). Microeconomics: Behavior, institutions, and evolution. In *Microeconomics: Behavior, Institutions, and Evolution*. New York: Sage.

Bramoullé, Y. (2007). Anti-coordination and social interactions. *Games and Economic Behavior*, *58*(1). https://doi.org/10.1016/j.geb.2005.12.006

Camerer, C. F., Dreber, A., Forsell, E., Ho, T. H., Huber, J., Johannesson, M., … Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, *351*(6280), 1433–1436. https://doi.org/10.1126/science.aaf0918

Camilotti, G. (2016). *Fighting against harmful customs: The case of female genital cutting*. University of Namur.

Chudek, M., & Henrich, J. (2011, May). Culture-gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in Cognitive Sciences*, Vol. 15, pp. 218–226. https://doi.org/10.1016/j.tics.2011.03.003

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places. *Journal of Personality and Social Psychology*, *58*(6), 1015–1026. https://doi.org/10.1037/0022-3514.58.6.1015

Coleman, J. S. (1990). *Foundations of Social Theory*. Harvard, MA: Belknap Press.

De Dreu, C., Beersma, B., Steinel, W., & Kleef, G. Van. (2007). The psychology of

negotiation: Principles and basic processes. In *Social psychology: Handbook of basic*

principles (Vol. 2, pp. 608–629). Retrieved from

http://www.researchgate.net/publication/232513754_The_psychology_of_negotiat

ion_Principles_and_basic_processes/file/79e4150c5c5c62939a.pdf%5Cnhttp://psy

cnet.apa.org/psycinfo/2007-11239-026

Duguid, S., Wyman, E., Grueneisen, S., & Tomasello, M. (2020). The Strategies Used by

Chimpanzees (Pan troglodytes) and Children (Homo sapiens) to Solve a Simple

Coordination Problem. *Journal of Comparative Psychology*.

https://doi.org/10.1037/com0000220

Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of "Minimal" Group

Affiliations in Children. *Child Development*, *82*(3), 793–811.

https://doi.org/10.1111/j.1467-8624.2011.01577.x

Efferson, C., Lalive, R., Richerson, P. J., McElreath, R., & Lubell, M. (2008). Conformists

and mavericks: the empirics of frequency-dependent cultural transmission.

Evolution and Human Behavior, *29*(1), 56–64.

https://doi.org/10.1016/j.evolhumbehav.2007.08.003

Efferson, C., Vogt, S., Elhadi, A., Ahmed, H. E. F., & Fehr, E. (2015). Female genital cutting

is not a social coordination norm. *Science*, Vol. 349, pp. 1446–1447.

https://doi.org/10.1126/science.aaa7978

Efferson, C., Vogt, S., & Fehr, E. (2020). The promise and the peril of using social

influence to reverse harmful traditions. *Nature Human Behaviour*, *4*(1), 55–68.

https://doi.org/10.1038/s41562-019-0768-2

Eugster, B., Lalive, R., Steinhauer, A., & Zweimüller, J. (2011). The Demand for Social

Insurance: Does Culture Matter? *Economic Journal*, *121*(556), 413–448.

https://doi.org/10.1111/j.1468-0297.2011.02479.x

Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature*, *454*(7208), 1079–1083. https://doi.org/10.1038/nature07155

Fehr, E., Glätzle-Rützler, D., & Sutter, M. (2013). The development of egalitarianism, altruism, spite and parochialism in childhood and adolescence. *European Economic Review*, *64*, 369–383. https://doi.org/10.1016/j.euroecorev.2013.09.006

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, *114*(3), 817–868. https://doi.org/10.1162/003355399556151

Fehr, E., & Schurtenberger, I. (2018). Normative foundations of human cooperation review-article. *Nature Human Behaviour*, Vol. 2. https://doi.org/10.1038/s41562-018-0385-5

Gallois, S., Lubbers, M. J., Hewlett, B., & Reyes-García, V. (2018). Social Networks and Knowledge Transmission Strategies among Baka Children, Southeastern Cameroon. *Human Nature*, *29*(4), 442–463. https://doi.org/10.1007/s12110-018-9328-0

Goeree, J. K., & Yariv, L. (2015). Conformity in the lab. *Journal of the Economic Science Association*, *1*(1), 15–28. https://doi.org/10.1007/s40881-015-0001-7

Granovetter, M. (1978). Threshold Models of Collective Behavior. *American Journal of Sociology*, *83*(6), 1420–1443. https://doi.org/10.1086/226707

Grueneisen, S., & Tomasello, M. (2019). Children use rules to coordinate in a social dilemma. *Journal of Experimental Child Psychology*, *179*(4), 362–374. https://doi.org/10.1016/j.jecp.2018.11.001

Grueneisen, S., Wyman, E., & Tomasello, M. (2015a). Children use salience to solve coordination problems. *Developmental Science*, *18*(3).

https://doi.org/10.1111/desc.12224

Grueneisen, S., Wyman, E., & Tomasello, M. (2015b). Conforming to coordinate: Children use majority information for peer coordination. *British Journal of Developmental Psychology*, *33*(1). https://doi.org/10.1111/bjdp.12078

Handley, C., & Mathew, S. (2020). Human large-scale cooperation as a product of competition between cultural groups. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-14416-8

Henrich, J., & Boyd, R. (2008). Division of labor, economic specialization, and the evolution of social stratification. *Current Anthropology*, *49*(4), 715–724. https://doi.org/10.1086/587889

Henrich, J., Boyd, R., Bowles, S., Camerer, C. F., Fehr, E., & Gintis, H. (2005). Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies. In *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. https://doi.org/10.1093/0199262055.001.0001

Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., … Ziker, J. (2010). Markets, religion, community size, and the evolution of fairness and punishment. *Science*, *327*(5972), 1480–1484. https://doi.org/10.1126/science.1182238

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., … Ziker, J. (2006). Costly punishment across human societies. *Science*, *312*(5781), 1767–1770. https://doi.org/10.1126/science.1127333

Holm, H. J. (2000). Gender-Based Focal Points. *Games and Economic Behavior*, *32*(2), 292–314. https://doi.org/10.1006/game.1998.0685

House, B. R. (2018). How do social norms influence prosocial development? *Current*

Opinion in Psychology, *20*, 87–91. https://doi.org/10.1016/j.copsyc.2017.08.011

House, B. R., Kanngiesser, P., Barrett, H. C., Broesch, T., Cebioglu, S., Crittenden, A. N., …

Silk, J. B. (2020). Universal norm psychology leads to societal diversity in prosocial

behaviour and development. *Nature Human Behaviour*, *4*(1).

https://doi.org/10.1038/s41562-019-0734-z

House, B. R., Kanngiesser, P., Clark Barrett, H., Yilmaz, S., Smith, A. M., Sebastian-Enesco,

C., … Silk, J. B. (2020). Social norms and cultural diversity in the development of

third-party punishment. *Proceedings of the Royal Society B: Biological Sciences*,

287(1925). https://doi.org/10.1098/rspb.2019.2794

House, B. R., Silk, J. B., Henrich, J., Barrett, H. C., Scelza, B. A., Boyette, A. H., … Laurence, S.

(2013). Ontogeny of prosocial behavior across diverse societies. *Proceedings of the*

National Academy of Sciences of the United States of America, *110*(36), 14586–

14591. https://doi.org/10.1073/pnas.1221217110

House, B. R., & Tomasello, M. (2018). Modeling social norms increasingly influences

costly sharing in middle childhood. *Journal of Experimental Child Psychology*, *171*,

84–98. https://doi.org/10.1016/j.jecp.2017.12.014

Howard, J. A., & Gibson, M. A. (2019). Is there a link between paternity concern and

female genital cutting in West Africa? *Evolution and Human Behavior*, *40*(1), 1–11.

https://doi.org/10.1016/j.evolhumbehav.2018.06.011

Kelly, D., & Hoburg, P. (2017). A tale of two processes: On Joseph Henrich's the secret of

our success: How culture is driving human evolution, domesticating our species,

and making us smarter. In *Philosophical Psychology* (Vol. 30).

https://doi.org/10.1080/09515089.2017.1299857

Kogut, T. (2012). Knowing what I should, doing what I want: From selfishness to

inequity aversion in young children's sharing behavior. *Journal of Economic*

Psychology, *33*(1), 226–236. https://doi.org/10.1016/j.joep.2011.10.003

Laland, K. N., & Rendell, L. (2019). Social learning: Theory. In *Encyclopedia of Animal Behavior*. https://doi.org/10.1016/B978-0-12-813251-7.00057-2

Lew-Levy, S., Kissler, S. M., Boyette, A. H., Crittenden, A. N., Mabulla, I. A., & Hewlett, B. S. (2020). Who teaches children to forage? Exploring the primacy of child-to-child teaching among Hadza and BaYaka Hunter-Gatherers of Tanzania and Congo. *Evolution and Human Behavior*, *41*(1), 12–22. https://doi.org/10.1016/j.evolhumbehav.2019.07.003

Lew-Levy, S., Lavi, N., Reckin, R., Cristóbal-Azkarate, J., & Ellis-Davies, K. (2018). How Do Hunter-Gatherer Children Learn Social and Gender Norms? A Meta-Ethnographic Review. *Cross-Cultural Research*, *52*(2), 213–255. https://doi.org/10.1177/1069397117723552

Lowes, S., Nunn, N., Robinson, J. A., & Weigel, J. L. (2017). The Evolution of Culture and Institutions: Evidence From the Kuba Kingdom. *Econometrica*, *85*(4), 1065–1091. https://doi.org/10.3982/ecta14139

Mackie, G. (1996). Ending footbinding and infibulation: A convention account. *American Sociological Review*, *61*(6), 999–1017. https://doi.org/10.2307/2096305

McAuliffe, K., Raihani, N. J., & Dunham, Y. (2017). Children are sensitive to norms of giving. *Cognition*, *167*, 151–159. https://doi.org/10.1016/j.cognition.2017.01.006

McElreath, R., Bell, A. V., Efferson, C., Lubell, M., Richerson, P. J., & Waring, T. (2008). Beyond existence and aiming outside the laboratory: Estimating frequency-dependent and pay-off-biased social learning strategies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1509), 3515–3528. https://doi.org/10.1098/rstb.2008.0131

McGuire, L., Elenbaas, L., Killen, M., & Rutland, A. (2019). The role of in-group norms and

group status in children's and adolescents' decisions to rectify resource inequalities. *British Journal of Developmental Psychology*, *37*(3). https://doi.org/10.1111/bjdp.12274

Mesoudi, A., Chang, L., Dall, S. R. X., & Thornton, A. (2016). The Evolution of Individual and Cultural Variation in Social Learning. *Trends in Ecology and Evolution*, *31*(3), 215–225. https://doi.org/10.1016/j.tree.2015.12.012

Misch, A., & Dunham, Y. (2021). (Peer) Group influence on children's prosocial and antisocial behavior. *Journal of Experimental Child Psychology*, *201*. https://doi.org/10.1016/j.jecp.2020.104994

Muthukrishna, M., & Henrich, J. (2016). Innovation in the collective brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, Vol. 371. https://doi.org/10.1098/rstb.2015.0192

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, *17*(11), 776–783. https://doi.org/10.1037/h0043424

Platteau, J. P., & Auriol, E. (2018). Eradicating women-hurting customs. In S. Anderson, L. Beaman, & J. Plateau (Eds.), *Towards Gender Equity in Development* (pp. 319–356). Oxford University Press.

Raiffa, H. (2002). *Negotiation analysis : the science and art of collaborative decision making / Howard Raiffa with John Richardson, David Metcalfe.* Retrieved from https://encore.lib.warwick.ac.uk/iii/encore/record/C__Rb1759099__SNegotiation analysis: the science and art of collaborative decision-making__Orightresult_U__X4?lang=eng&suite=cobalt

Rakoczy, H., Hamann, K., Warneken, F., & Tomasello, M. (2010). Bigger knows better: Young children selectively learn rule games from adults rather than from peers.

British Journal of Developmental Psychology, *28*(4).

https://doi.org/10.1348/026151009X479178

Rapp, D. J., Engelmann, J. M., Herrmann, E., & Tomasello, M. (2019). Young children's

reputational strategies in a peer group context. *Developmental Psychology*, *55*(2),

329–336. https://doi.org/10.1037/dev0000639

Richerson, P. J., & Boyd, R. (2008). *Not by Genes Alone: How Culture Transformed Human*

Evolution. Chicago: University of Chicago Press.

Richerson, P. J., & Boyd, R. (2020). The human life history is adapted to exploit the

adaptive advantages of culture. *Philosophical Transactions of the Royal Society B:*

Biological Sciences, *375*(1803), 20190498. https://doi.org/10.1098/rstb.2019.0498

Rushton, J. P. (1976). Socialization and the altruistic behavior of children. *Psychological*

Bulletin, *83*(5), 898–913. https://doi.org/10.1037/0033-2909.83.5.898

Salali, G. D., Chaudhary, N., Bouer, J., Thompson, J., Vinicius, L., & Migliano, A. B. (2019).

Development of social learning and play in BaYaka hunter-gatherers of Congo.

Scientific Reports, *9*(1), 11080. https://doi.org/10.1038/s41598-019-47515-8

Schelling, T. C. (1978). *Micromotives and macrobehavior*.

https://doi.org/10.1177/027614678200200210

Shell-Duncan, B., Wander, K., Hernlund, Y., & Moreau, A. (2011). Dynamics of change in

the practice of female genital cutting in Senegambia: Testing predictions of social

convention theory. *Social Science and Medicine*, *73*(8), 1275–1283.

https://doi.org/10.1016/j.socscimed.2011.07.022

Siddique, A., & Vlassopoulos, M. (2020). Competitive Preferences and Ethnicity:

Experimental Evidence from Bangladesh. *Economic Journal*, *130*(627), 793–821.

https://doi.org/10.1093/ej/uez063

Smith, C. E., Blake, P. R., & Harris, P. L. (2013). I Should but I Won't: Why Young Children

Endorse Norms of Fair Sharing but Do Not Follow Them. *PLoS ONE*, *8*(3). https://doi.org/10.1371/journal.pone.0059510

Spielman, D. A. (2000). Young children, minimal groups, and dichotomous categorization. *Personality and Social Psychology Bulletin*, *26*(11), 1433–1441. https://doi.org/10.1177/0146167200263010

Ullmann-Margalit, E. (1977). *The Emergence of Norms*. Oxford: Oxford University Press.

Vogt, S., Efferson, C., Berger, J., & Fehr, E. (2015). Eye spots do not increase altruism in children. *Evolution and Human Behavior*, *36*(3), 224–231. https://doi.org/10.1016/j.evolhumbehav.2014.11.007

Vogt, S., Zaid, M. N. A., Ahmed, H. E. F., Fehr, E., & Efferson, C. (2016). Changing cultural attitudes towards female genital cutting. *Nature*, *538*(7626), 506–509. https://doi.org/10.1038/nature20100

Warneken, F. (2016). Insights into the biological foundation of human altruistic sentiments. *Current Opinion in Psychology*, *7*, 51–56. https://doi.org/10.1016/j.copsyc.2015.07.013

Warneken, F. (2018). How Children Solve the Two Challenges of Cooperation. *Annual Review of Psychology*, *69*(1), 205–229. https://doi.org/10.1146/annurev-psych-122216-011813

Young, H. P. (2013). The economics of convention. *Law and Economics: A Reader*, *10*(2), 142–158.

Young, H. P., & Burke, M. A. (2001). Competition and custom in economic contracts: A case study of Illinois agriculture. *American Economic Review*, *91*(3), 559–573. https://doi.org/10.1257/aer.91.3.559

Young, H. P., & Raiffa, H. (1983). The Art and Science of Negotiation. In *Journal of Policy Analysis and Management* (Vol. 3). https://doi.org/10.2307/3324050

Zmyj, N., & Seehagen, S. (2013). The Role of a Model's Age for Young Children's

    Imitation: A Research Review. *Infant and Child Development, 22*(6).

    https://doi.org/10.1002/icd.1811