Forthcoming in *Science Advances*

The emergence of cooperative behaviors, norms, and strategies across five diverse societies

Dorsa Amir[1,2*], Richard E. Ahl[2], Matthew R. Jordan[2], Hannah Bolotin[2], Michael Bogese[2], Gorana T. González[2,6], Tara Callaghan[3], Lawrence S. Sugiyama[4], Emily Otali[5], Patrick Tusiime[5], Samantha Bangayan[2], J. Josh Snodgrass[4,7], Katherine McAuliffe[2]

1. Duke University, Department of Psychology and Neuroscience, Durham, NC 27708 USA
2. Boston College, Department of Psychology and Neuroscience, Chestnut Hill, MA 02467, USA
3. St. Francis Xavier University, Department of Psychology, Antigonish, Nova Scotia B2G 2W5, Canada
4. University of Oregon, Department of Anthropology, Eugene, OR 97403, USA
5. Kibale Forest Schools Program, Kabarole, Uganda
6. University of Massachusetts, Amherst, Amherst, MA 01003, USA
7. Hokkaido University, Global Station for Indigenous Studies and Cultural Diversity, Hokkaido, Japan

*Corresponding author: Dorsa Amir, dorsa.amir@duke.edu

Teaser: Children's cooperative behaviors and norms develop along distinct cultural pathways shaped by local norms.

Short title: Cooperation across cultures

Abstract:

Human cooperation involves a set of interconnected behaviors that develop in conjunction with the cultural environment. Despite recent advances in Western, industrialized contexts, we know far less about how cooperative behaviors emerge across cultures, how normative environments shape their development, and how these behaviors relate to one another. Here, we examined the development of four cooperative behaviors — fairness, trustworthiness, forgiveness, and honesty — in children (N=413) aged 5–13 from five societies: urban United States, rural Uganda, Canada, Peru, and the hunter-horticulturalist Shuar of Ecuador. We also collected normative judgments from peers (N=163) and adults (N=86) of each community. We find substantial variation in cooperative behaviors and norms across populations, but, more generally, that children's behaviors and norms tend to converge toward community-specific norms in middle childhood. We also identify three cooperative strategies — Maximization, Generic Cooperation, and Partner-Contingent Cooperation — whose prevalence shifts with age and differs across societies. Together, these findings illuminate how cooperative behavior develops within and across cultures.

Keywords: *cooperation, norms, development, cross-cultural, cognition*

Introduction

Cooperation is multi-faceted & shaped by norms

Human cooperation is multi-faceted (*1*), encompassing a wide range of interrelated preferences and behaviors. Here, we focus on four behaviors that play a key role in cooperation: those related to fairness, trustworthiness, forgiveness, and honesty. Many of our cooperative interactions reflect our preferences toward *fairness*, defined here as a preference for equitable outcomes, even when such outcomes come at a cost to oneself (*2*). These preferences guide behavior (*3*), influencing individuals' decisions to cooperate or defect in social exchange situations (*2*). Many human interactions are repeated in nature, which means, among other things, that one's ability to predict whether another party will reciprocate (i.e., whether they are trustworthy) is important in their decision to cooperate. Thus, interpersonal trust — the acceptance of risk and vulnerability from one party in anticipation of a positive result from another party's actions (*4*, *5*) — is also determinative of cooperation. An important feature of interpersonal trust relates to an individual's *trustworthiness* — or the degree to which an individual produces the positive outcome one has been entrusted to produce, even if it comes at a cost (*4*). When transgressions occur, computational models (*6*) and empirical research (*7*) both suggest that *forgiveness* — a suite of prosocial motivational changes in the face of transgressions (*8*) — can help restore cooperation. Lastly, *honesty* also influences cooperation, as individuals frequently assess the validity of others' intentions and prefer honest individuals in cooperative interactions (*9*).

Social norms — the informal rules that govern behavior in groups or societies(*10*) — shape cooperative behaviors across diverse societies by enabling the coordination and alignment of individuals' interests. Many cooperative interactions can be thought of as *mixed-motive games*, or social dilemmas in which there is an individual incentive to be selfish and free-ride(*11*, *12*). As a result, groups can find themselves in inefficient equilibria over the long term. Social norms, however, can reshape the payoffs of these interactions, by, for instance, attaching social costs to defection (e.g. in the form of punishment of non-cooperators) or by attaching social benefits to cooperation (e.g. in the form of reputational gains). In this way, norms can help stabilize long-term cooperative interactions by shifting individuals' behaviors toward actions that may be individually costly in the short term, but can produce benefits in the long-term.

Though the content of norms regarding cooperation can vary extensively across societies(*13*), there is some evidence for cross-cultural regularities in *norm psychology*. Even early in childhood, individuals appear receptive to local norms that govern cooperation and use them to guide downstream behaviors(*14*). This work points to early development as a key window for norm acquisition and adoption. In other words, in early life, children must learn the content of *their* group's social norms and practice the process of integrating those norms — which can often be costly in the short term — into their behavioral repertoires(*15*). Indeed, some have argued that the extended period of dependence in early life may have emerged in part to support this complex form of cultural learning(*16*, *17*).

Despite the theoretical importance of cultural variation and development in shaping cooperative behaviors, relatively little empirical work to date has examined how this process unfolds among diverse societies which vary across dimensions like normative content and social structure. A small but emerging body of work has begun to demonstrate both cultural variability and convergence in children's cooperative behaviors. For instance, work using economic games has shown that preferences for fairness, reciprocity, and sharing emerge early but vary in expression across societies(*14*, *15*, *18–23*). Further, children's developing behaviors are informed by local social and moral norms(*3*, *12*, *15*). However, much of this past work (including our own), has relied primarily on isolated behaviors in individual decision-making tasks(*19*, *24*), despite the important influence of norms on cooperative behaviors(*25*) and variation in those norms across cultures(*26*). To advance our understanding of human cooperation, we build on these lines of inquiry to produce a unified, holistic assessment of the development of cooperative behaviors, capturing not just children's behaviors, but also peers' and adults' normative judgments in those same communities. Further, we explore the concurrent development of four key cooperative behaviors within the same individual. Our battery of cooperative tasks allows us to assess whether cooperative behaviors develop in tandem, as part of a larger interrelated suite of behaviors, or whether they develop in dissociable ways across childhood.

To assess the ways in which cooperative behaviors can vary across age and culture, we recruited N = 413 children between the ages of 5 and 13 years old (M = 8.78, SD = 2.40, 214 girls) across five diverse societies: the forager-horticulturalist Shuar of Ecuador, urban-dwelling children in the US, and rural children in Peru, Uganda, and Canada. To ensure meaningful cross-cultural comparisons, we selected societies varying in subsistence systems, market exposure, and economic development, dimensions previously linked to differences in cooperative behavior and decision-making under uncertainty(*18*, *27*, *28*). A summary of our participant sample can be found in **Table 1.** A more detailed demographic breakdown can be found in the **Supplement.** Our motivation, sampling strategy, and research questions were pre-registered on AsPredicted and are accessible at the following link: https://aspredicted.org/ZR9_Z1Y.

Table 1: **Participant demographics across societies.** Demographic features of participants across locations. Sample sizes for each task varied slightly based on task-specific exclusion criteria. The sample size reported here is the largest N in each site.

			Task Battery		**Child Norm Battery**		**Adult Norm Battery**	
Country	**Region (Society)**	**Economy**	**N (males)**	**Mean age (range)**	**N (males)**	**Mean age (range)**	**N (males)**	**Mean age (range)**
Ecuador	Amazonian Ecuador (Shuar)	Horticulture, fishing, hunting, gathering, limited agro-pastoralism and sporadic wage labor	67 (35)	8.97 (5-13)	24 (15)	9.62 (6-13)	13 (5)	44.5 (21-74)
United	Boston	Professional,	81 (42)	8.19 (5-13)	37 (20)	8.19 (5-13)	20 (7)	33.6 (21-73)

States		trade/service, labor						
Peru	San Pedro de Saño	Agricultural labor, trade/service	86 (44)	9.20 (5-13)	46 (24)	9.15 (6-13)	18 (8)	32.1 (18-50)
Canada	Antigonish	Professional, trade/service, labor	71 (33)	8.03 (5-12)	25 (12)	7.20 (5-10)	13 (5)	21.5 (18-33)
Uganda	Fort Portal (Tooro)	Agricultural labor, trade/service	108 (45)	9.27 (5-13)	31 (16)	9.06 (6-13)	22 (10)	31.9 (23-56)

We designed a battery of four tasks to explore the development of fairness, trustworthiness, forgiveness, and honesty (**Figure 1**). Participants were presented with all four tasks in either forward or reversed order, counterbalanced (See **Supplement** for detailed scripts and protocol). Participants were told that their choices affected both their own resources and those of three anonymous partners, with a unique anonymous partner in three of the tasks. Our Fairness Task utilized the Inequity Game, a resource allocation task that measures children's acceptances and costly rejections of various distributions(*29*) and has successfully been used to study fairness behavior across societies(*3*, *27*). Children made decisions in seven test trials: two equal trials (EQ) where the actor and recipient both received one resource, and four advantageous inequity (AI) trials, in which the actor received four resources, and the recipient received one. These first six trials were randomized. The last trial was always an Extreme AI trial (XAI) where the actor received four resources, and the recipient received zero. Rejections in the AI and XAI trials indicated a costly aversion to unequal outcomes that favored the actor over the recipient, also known as *advantageous inequity aversion* (*3*, *27*). We chose to focus on rejections of AI as they show considerable variability across age and societies and represent a costly signal of fairness(*30*, *31*) .

Our Trustworthiness Task utilized an intuitive and apparatus(*4*) that mimicked the dynamics of the second player's behavior in the Trust Game(*32*), measuring how many resources participants would share with another player who gave up resources for their benefit. Past work using this method has validated that this task does capture trustworthiness rather than simple generosity(*4*). Our Forgiveness Task measured how many of five resources participants gave back to a transgressor, as opposed to discarding them(*33*). Past work with this method similarly has validated that children are appropriately integrating information about intent and remorse, and responding differentially to such features (*33*). And finally, our Honesty Task measured how many resources children would give back to an experimenter after they privately opened a box with four resources in it when they were only supposed to receive one. To our knowledge, this is the first developmental assessment of a suite of cooperative tasks across cultures.

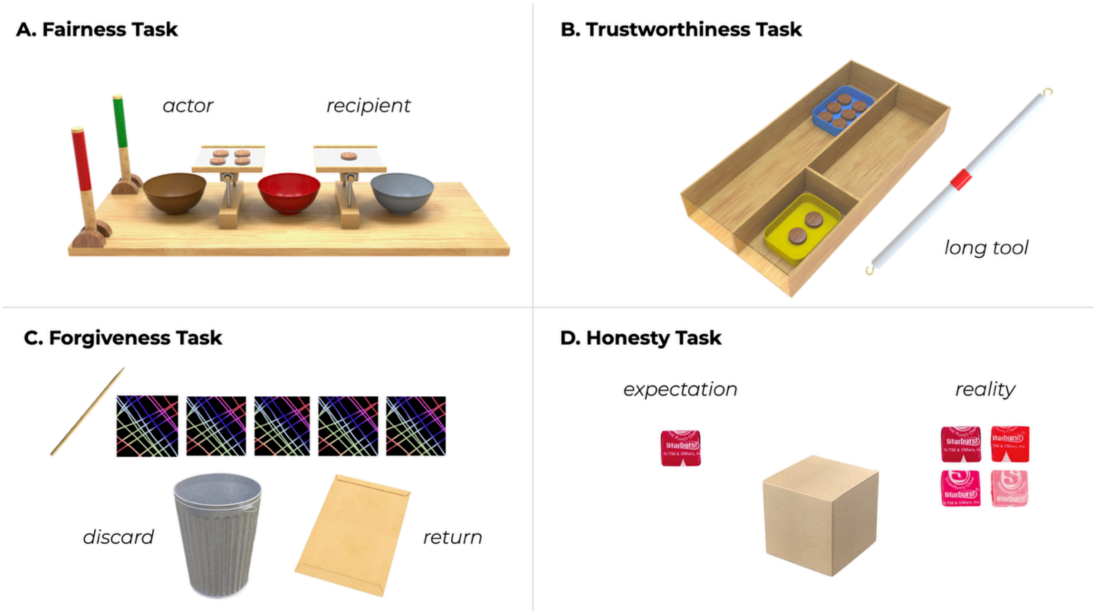

Figure 1: Methods and stimuli for the four behavioral tasks. (A) depicts the Inequity Game, an apparatus with two trays: one for the actor, one for the recipient. The image shows a 4-1 Advantageous Inequity distribution. The actor has access to a green handle which enacts the distribution, and a red handle which rejects the distribution. The resources in this task were candies. Fairness is proxied through the number of rejected unequal trials. **(B)** depicts the Trustworthiness Task, a child-friendly version of the Trust Game. The apparatus has two trays: a yellow tray that is closer to the opening that has two resources, and a blue tray that is further away from the opening and has six resources (candies). The blue tray can only be accessed by the other player forfeiting their small tool, which can be joined with the actor's small tool to form a long tool. Trustworthiness is proxied through the number of resources sent back to the other player. **(C)** depicts the Forgiveness Task, in which children must decide how many of five scratch sheets to return to a transgressor who accidentally drew on them and how many to throw away. Forgiveness is proxied through the number of sheets returned to the transgressor. **(D)** depicts the Honesty Task, in which children are given a box they are told should have one resource, but in actuality has four resources inside (candies). Honesty is proxied through the number of resources returned to the experimenter.

In addition to these behavioral measures, we also collected normative judgments through norm surveys of participants' peers (N = 163 children between the ages of 5 and 13 years old (M = 8.69, SD = 2.34, 76 girls)) — none of whom, with the exception of children in Ecuador (N = 24), had participated in the behavioral tasks, to prevent spillover from their own behavioral choices into normative judgements. We focused on children between the ages of 5 and 13 as this period of childhood is key for moral development and the integration of social norms (e.g. *18*, *27*), and because the structural demands of the tasks relied on cognitive and physical abilities (e.g. numeracy, verbal instructions) that emerge at age five and later. We also recruited local adults from each of five surveyed communities (N = 86 adults between the ages of 18 and 74 years old, M = 32.6, SD = 12.6, 35 men). In these normative surveys,

participants were introduced to five vignettes that mirrored the choices that children faced in the behavioral task battery, in either forward or reversed order, counterbalanced (see **Supplement** for full protocol and more details). In this way, we were able to capture not just children's actual behaviors in each of these decision-making contexts, but also others' judgments of what their behavior *should* be in each context, allowing us to construct a more complete picture of the normative environment in each community.

Results

Does cooperative behavior vary across age and culture?

Our primary, pre-registered analysis was centered on assessing developmental differences in the four cooperative behaviors — fairness, trustworthiness, forgiveness, and honesty – across cultures. We built a number of linear models and linear probability models to predict the respective dependent variable of interest for each task (*34*). In the Fairness Task, the dependent variable was whether a participant rejected on a given trial (binary, accept or reject). In the Trustworthiness Task, the dependent measure was the number of candies shared with the partner that trusted them (numeric, 0-6). In the Forgiveness Task, the dependent variable was the number of sheets given back to the transgressor (numeric, 0-5). In the Honesty Task, the dependent variable was the number of candies returned to the experimenter (numeric, 0-4). In all analyses, Ecuador was modeled as the baseline society, as the socioecological environment of the children in Ecuador (small-scale hunter-horticulturalists) is the most different from the environments of the other children. We expected their behavior to vary the most from children in the industrialized regions, as evident in previous work among these same communities (*28*). The primary, pre-registered predictors were age (numeric, years), and society (factor, five levels).

To address whether children's behaviors and normative judgments converged toward adult norms with age, we computed difference scores between children's behaviors and norms as compared to the adult norms. These difference scores were our dependent variables in a number of linear (probability) models, where age (continuous) and society (factor, five levels) were the predictors. See **Supplement** for more details and model outputs.

Age-related changes in children's behavior

An analysis of developmental trends in children's behaviors, pooled across all societies, yielded two main take-aways. First, in many cases, cooperative behaviors tended to increase with age. For costly fairness — the rejection of AI trials — children were more likely to reject unequal allocations than equal allocations ($\beta = 0.20$, $SE = 0.02$, $p < 0.0001$). Rejections were approximately three times more likely in unequal trials than equal ones. Further, younger children were much more likely to accept any kind of distribution ($\beta = -0.02$, $SE = 0.01$, $p < 0.001$), but this age trend is driven entirely by younger children accepting unequal offers in their favor (see **Supplement**). Similarly, trustworthiness and forgiveness

behavior showed age-related increases. For trustworthiness, children shared an additional 0.12 Starbursts for each additional year of age, meaning the oldest children in the sample shared nearly one more Starburst than the youngest children ($\beta = 0.12$, SE = 0.24, $p < 0.001$). Similarly for forgiveness, we found that children returned 0.11 additional sheets for each year of age, which translates to the oldest children returning nearly one additional sheet to the transgressor as compared to the youngest children ($\beta = 0.11$, SE = 0.29, $p < 0.001$). An age-related increase was also observed for honesty, which was modeled as a binary variable (returned or not) because virtually all children either returned three candies or none. However, this age effect was non-significant but marginal ($\beta = 0.01$, SE = 0.01, $p = 0.08$).

Cross-societal differences in the development of children's behaviors

When considering how behaviors vary across communities, we found evidence for cross-societal differences across all cooperative behaviors (**Figure 2**). Consistent with our expectations, children in Ecuador were the most different from children in other sites, yet the specific differences varied depending on task. For fairness, we found participants in Uganda, Peru, USA, and Canada rejected significantly more than children in Ecuador (all ps < 0.001). Additionally, we found that the developmental trajectory of fairness showed cross-societal variation. Specifically, we found a two-way interaction between age and society ($\chi2(4) = 30.25$, $p < 0.0001$): compared to children in Ecuador, children in Peru ($\beta = 0.06$, SE = 0.02, $p < 0.0001$), Canada ($\beta = 0.08$, SE = 0.02, $p < 0.0001$), and the United States ($\beta = 0.06$, SE = 0.02, $p < 0.001$) were more likely to reject AI trials with increasing age. This is consistent with past findings suggesting that AI rejections increase with age in some, but not all, societies(*27*). It also reflects the union of two interesting trends: that children outside of Ecuador reject AI more with age, but also that children within Ecuador actually reject AI *less* with age ($\beta = -0.026$, SE = 0.007, $p < 0.001$) (see panel A in **Figure 2**).

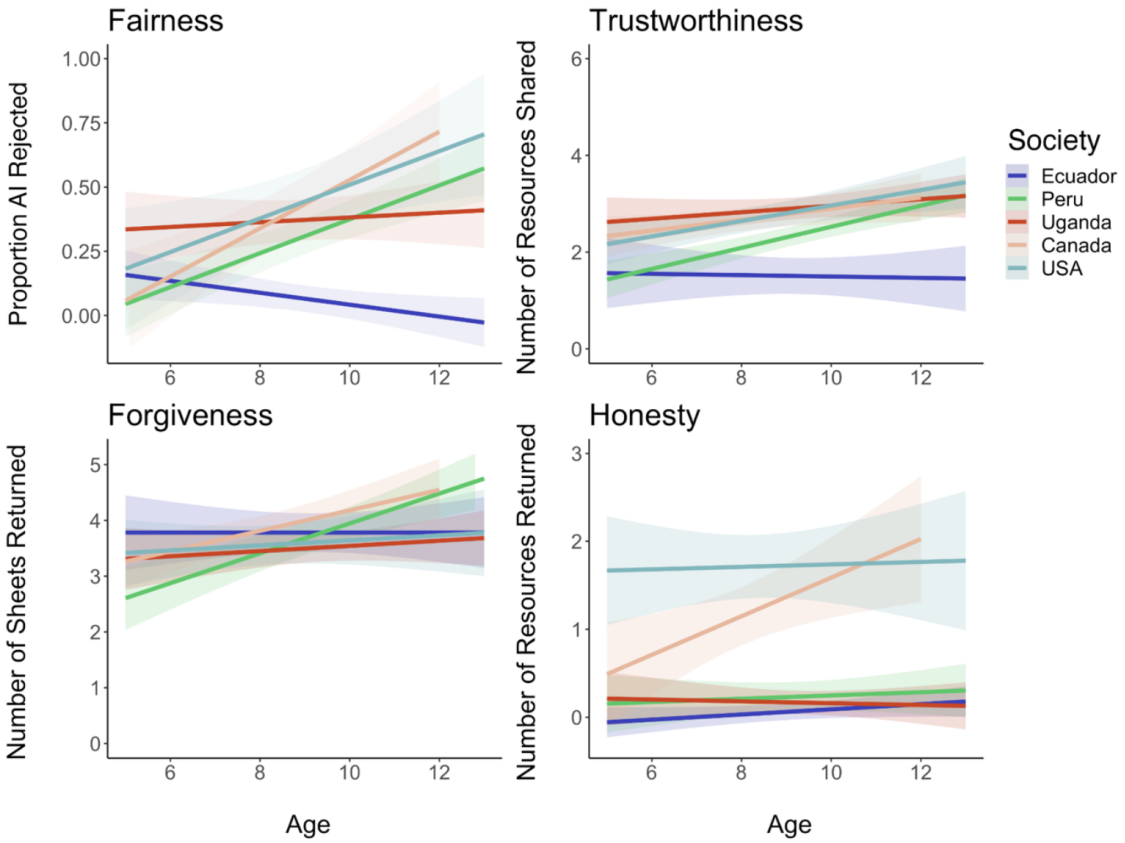

Figure 2: Children's behavior on the four cooperative preference tasks by age, society, and task. Shaded ribbons represent 95% confidence intervals.

These patterns were consistent, but in fact much stronger, in the XAI trials: we found a strong main effect of age, such that children were more likely to reject extreme AI with increasing age ($\beta = 0.07$, SE = 0.01, $p < 0.0001$), and substantial differences between societies, such that children in Peru, Uganda, Canada, and the United States were much more likely to reject the Extreme AI trial than children in Ecuador (all ps < 0.0001). These main effects were qualified by a two-way interaction between age and society (LRT, $\chi 2 = 3.29$, $p < 0.001$). Breaking this down further, we found that children in Peru ($\beta = 0.09$, SE = 0.03, $p < 0.001$), Uganda ($\beta = 0.07$, SE = 0.03, $p = 0.02$), Canada ($\beta = 0.01$, SE = 0.03, $p < 0.001$), and the United States ($\beta = 0.09$, SE = 0.02, $p < 0.01$) were all more likely than children in Ecuador to reject the Extreme AI trial as they aged. Indeed, only one participant in the entire Ecuadorean sample rejected this trial type.

For trustworthiness, children in Peru, Uganda, Canada, and the United States shared more resources with the trustor than children in Ecuador (all ps < 0.001). And, again, we found a two-way interaction between society and age ($\chi^2(4) = 14.32$, $p = 0.02$): compared to children in Ecuador, children in Peru ($\beta = 0.23$, SE = 0.07, $p < 0.01$), and the United States ($\beta =$

0.17, SE = 0.08, p = 0.03) shared more resources with the trustor as they aged (see panel B of **Figure 2**). For forgiveness, we did not see a main effect of society, but found a significant two-way interaction between society and age ($\chi^2(4) = 23.60$, p =0.01): compared to children in Ecuador, children in Peru ($\beta = 0.27$, SE = 0.09, $p < 0.01$), returned more sheets to the transgressor with increasing age (see panel C of **Figure 2**). However, behavior was much more similar in the forgiveness task across societies and age than in any of the other tasks. Finally, for honesty, we found children in Canada ($\beta = 0.38$, SE = 0.06, $p < 0.0001$) and the United States ($\beta = 0.55$, SE = 0.06, $p < 0.0001$) were significantly more likely to return resources than children in Ecuador (see panel D of **Figure 2**). We also found a marginal but insignificant two-way interaction between age and society ($\chi^2(4)= 1.04$, p =0.08) for honesty.

What inferences can be drawn about children's behaviors?

Considering the behavioral data altogether, we find both differences and similarities across cultures. First, we see substantial societal differences in children's fairness and trustworthiness behaviors. Most notably, Shuar children in Ecuador differ from children in other societies in a number of ways. While other children tend to reject standard AI with age, we see the opposite pattern emerge in Ecuador. And when it came to XAI, there was very little variance: virtually all children in Ecuador accepted this trial type. Relatedly, we see very little behavioral change in trustworthiness, forgiveness, and honesty behavior across age in Ecuador. For forgiveness, in contrast, we find overlap across these diverse sites, with forgiving the transgressor (returning most or all of the sheets to the transgressor) being the clear modal behavior in every community surveyed.

There are several inferences that can be drawn from this set of behavioral data. First, and most clearly, the documented variation across societies suggests that the behavior of children in some industrialized societies such as the United States cannot be assumed to generalize to those living in other societies without empirical assessment (*16*). This is in line with a number of other investigations documenting much greater variability in children's behaviors when cultural diversity is more broadly sampled (e.g. *28*, *35*). Second, across these distinct tasks, we nevertheless observed interesting regularities in trajectories across societies, most notably that children in all sites tended to demonstrate more self-interested behavior in early life, with more similar starting points at younger ages across our tasks, providing additional support for this recurrent finding in the literature (e.g. *18*, *27*). Further, in line with past work (*12*, *14*, *15*, *27*), we find that middle childhood appears to be a significant period for potential change, as children's behaviors begin to more clearly dissociate and diverge — as in the case of fairness, trustworthiness, and honesty.

How do we interpret the differing behavior of Shuar children in these tasks? We believe their behavior, especially with respect to accepting AI and XAI offers, likely reflects differing motivations, in particular as they pertain to the equality-efficiency tradeoff (e.g. *36*, *37*). Indeed, looking holistically across the tasks that involve resources, Shuar children quite effectively maximize their resources, ultimately gaining the most candy. To our eye, Shuar children appear to be prioritizing efficiency over perceived equality. This behavior converges

with past work among this population which has suggested that socioeconomic factors such as low market integration may lead to resource-maximization behaviors across risk and time preference tasks (*28*). Rather than challenging the idea that norms shape the emergence of these behaviors, it's much more likely that prioritizing efficiency *itself* constitutes a locally adaptive norm — one that favors minimizing waste and maximizing returns, rather than enforcing strict equality. Further, as our tasks are capturing what could be referred to as *generic* cooperation — cooperation with an anonymous other — it may also be the case that children who accept AI are maximizing resources in order to engage in cooperation with known others *outside* the task; an important question for future research to address. As we'll discuss in Section 4, these behaviors are fully aligned with children's normative judgments about this decision, as well.

How dissociable are cooperative behaviors across age and society?

Our battery of tasks affords us the opportunity to address questions about whether and to what extent cooperative behaviors are related within individuals throughout development (as in (*38*)), but now contextualized within the normative environment of each community. By examining individuals' cooperative behaviors across tasks, we aim to determine which strategies—i.e., combinations of choices across the tasks—are used by the children in our sample. If we can identify distinct strategies, we can then 1) describe the nature of each strategy, 2) examine how the prevalence of those strategies changes across development, and 3) use the prevalence of strategies across societies to recover our different populations.

As pre-registered, we conducted a number of exploratory analyses to consider how behaviors in the four tasks may relate to each other within subjects, and how these patterns may vary across societies. To identify strategies, we used cluster analysis (analytical details can be found in the Supplement). We clustered individuals using only children for which we had complete data across all tasks (resulting in a slightly reduced dataset of N = 366 children), aggregating behavioral data from the trustworthiness task, the honesty task, the forgiveness task, and the AI and XAI trials of the fairness task. The model identified three clusters, each of which corresponds to a distinct cooperative strategy. We have labeled them as such: 1) **Maximization**: this strategy prioritizes maximizing individual resources when faced with unspecified targets, 2) **Generic Cooperation**: this strategy defaults to cooperation, even in cases when the target is unspecified and when there is no partner involved, and 3) **Partner-Contingent Cooperation**: this strategy cooperates in the presence of unspecified others, but maximizes individual resources when no other agent is involved. These labels are intended as descriptive shorthand to capture the main distinctions that emerged in our data and set our tasks apart, rather than as exhaustive or fixed theoretical categories. See the Supplement for the task-level behaviors that characterize these strategies.

After having identified our three strategies and labeled each child in our sample with a strategy, we next examined how the prevalence of these strategies changes throughout

development. As shown in **Figure 3, Panel A**, the majority of younger children use Maximization, and Maximization becomes less prevalent across development. The opposite is true for Partner-Contingent Cooperation, which is the rarest strategy among the youngest children, but becomes the most common strategy used among the oldest children. Generic Cooperation remains equally prevalent (at around 20% of our sample) across development.

Finally, we examined the frequencies of these three strategies across development and societies. Specifically, we aimed to determine if the strategic make-ups of our diverse societies reflect local social incentives. As shown in **Figure 3, Panel B**, we can see the five societies tend to coalesce into three distinct patterns: Ecuador, in which nearly the entire sample uses Maximization throughout development; Peru and Uganda, in which the vast majority of children use either Maximization or Partner-Contingent Cooperation, and children switch from Maximization to Partner-Contingent Cooperation as they develop; and Canada and USA, in which all three strategies are present, and both Generic Cooperation and Partner-Contingent Cooperation replace Maximization across development. Interestingly, these patterns generally align with subsistence practice and degree of market integration of each community: the Shuar in Ecuador are forager-horticulturalists with the lowest degree of industrialization, the communities in Peru and Uganda are largely agricultural and industrializing, and the communities in Canada and the US are fully industrialized. To clarify, we do not claim these three strategies represent all societies or cooperative behaviors; through cluster analyses, our approach aims to identify underlying structure within our data to reveal patterns that might otherwise go unnoticed. Nonetheless, it is striking that these strategy profiles, derived solely from children's behavioral patterns, so closely recover the underlying subsistence structure of each society.

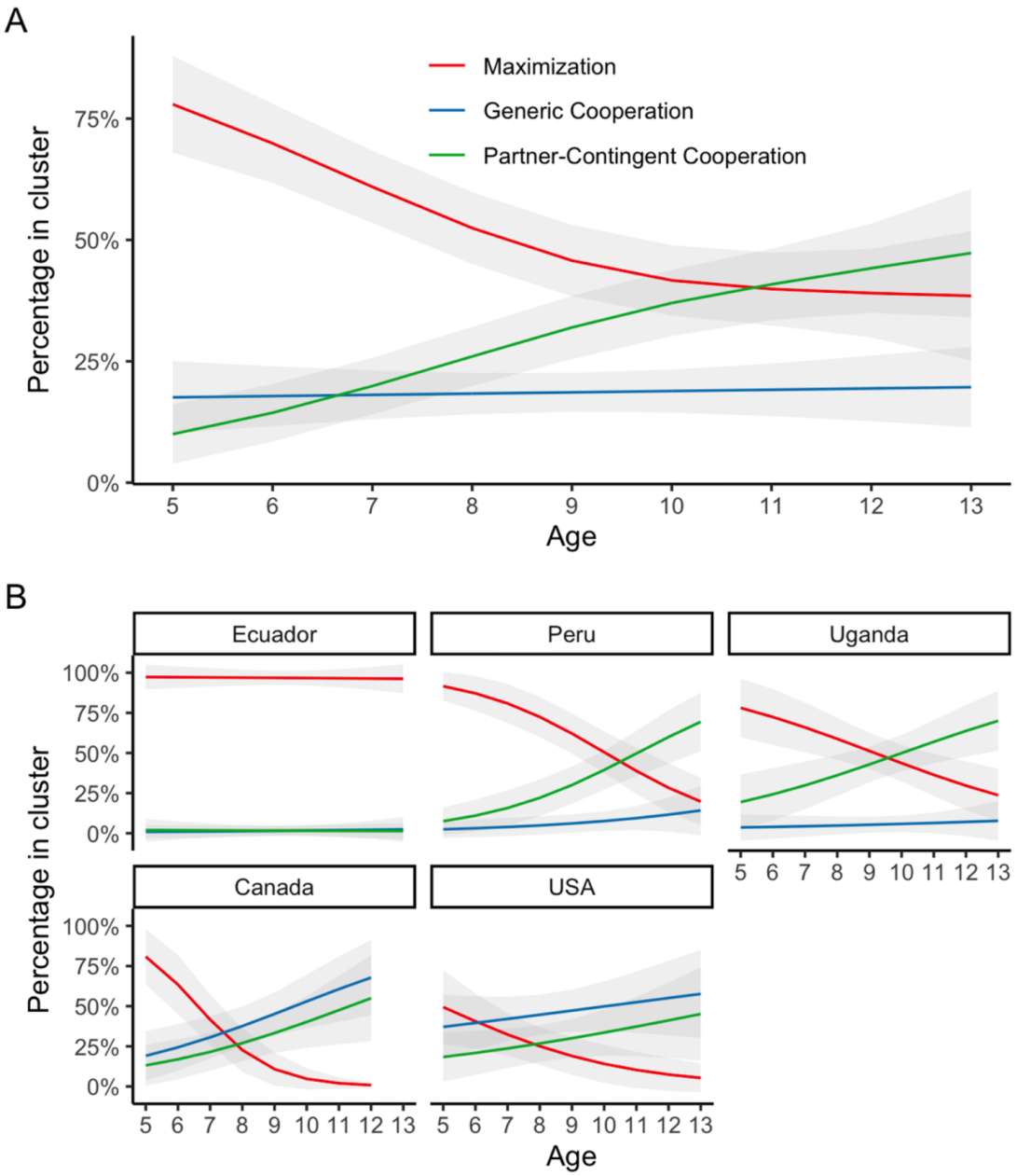

Figure 3. Cooperative strategies across age and society. (Panel A) A plot showing the prevalence of the Maximization, Generic Cooperation, and Partner-Contingent Cooperation strategies across development. Shaded areas are 95% confidence intervals. Lines represent GAM-predicted cluster prevalence by age (see Supplement for full analysis details). **(Panel B)** A plot showing the prevalence of the three strategies across development within each society. Shaded areas are 95% confidence intervals. Lines represent GAM-predicted cluster prevalence by age (see Supplement for full analysis details)

This set of analyses, using data from all tasks, provides new insights into the strategies that children in our sample employ across cooperative settings. First, our bottom-up approach for identifying patterns in behavior reveals the emergence of three distinct

strategies in our dataset. Second, there are clear developmental trends in the prevalence of these three strategies, as Maximization becomes less prevalent, and Generic and Partner-Contingent Cooperation more prevalent, across development. And third, Generic Cooperation is only present in the most industrialized societies (Canada and USA). We view this as an important proof of concept: even within a dynamic, cross-cultural battery, children show patterned strategic behavior across tasks — evidence that this approach can yield meaningful structure beyond a single-task design.

A norms-and-behavior approach to cooperation

While we see evidence of age- and society-related differences in children's behaviors in the behavioral tasks — with more homogenous behavior in early ages, followed by divergence across societies with age — it is challenging to make deeper inferences without cultural contextualization. The normative data from adults, provided by the additional surveys we conducted, allows us to formally ask: do children's behaviors and judgments converge toward adult norms with age? And, do we see differences between societies? Here, we are well positioned to examine norms-and-behavior in tandem across development, across diverse societies, and across distinct domains of cooperative behavior.

To examine how local norms shape children's behaviors and judgments, we employed a two-pronged, exploratory approach that capitalized on our survey data from adults and children. First, we explored whether children's behaviors begin to align with adult-endorsed views of what *should* be done (e.g. the "right" thing to do, or the *injunctive norm* (*10*)) (using our *adult norms* dataset). Second, we examined the development of children's own notions of the "right" behavior (our *child norms* dataset) and compared these judgments to adult norms. To measure adult norms, we used the mean response from the adult norm survey in each society as a stand-in for the average adult judgment (see **Supplement** for more). To assess whether children's behaviors converged toward these adult averages with age, we created difference scores for each child in each task, which was calculated as the absolute value of the difference between the average adult-endorsed behavior and the child's actual behavior. Taking an absolute value thus captures the size of the deviation from the adult norm, but in both directions. Take trustworthiness as an example. If the average adult in a given society said to share three resources in the trustworthiness task, and the child shared only one, the difference score for that child in the trustworthiness task would be two. We computed difference scores in this way for all of the participants' actions in all our tasks. For the fairness tasks, children's rejections were averaged across the four AI trials and subtracted from the adult norms to create a difference score. Our models were based on the logic that if children are converging toward the adult average with age, the difference scores should significantly decrease with increasing age.

In **Figure 4**, the data from the behavioral tasks (solid lines), the normative judgments of children (dashed lines) and the normative judgments of adults (solid circles) can be considered in unison. The point of this plot — which we refer to as a *comet plot* — is to demonstrate that children's behaviors and norms tend to converge toward a target: local,

adult norms — the adult norms form the head of the comet and children's trajectories trail toward them like a comet's tail. This visual metaphor captures the developmental movement or "drift" of children's responses over time, showing how their behaviors and normative judgments tend to gradually converge toward the adult norm that anchors each cultural context.

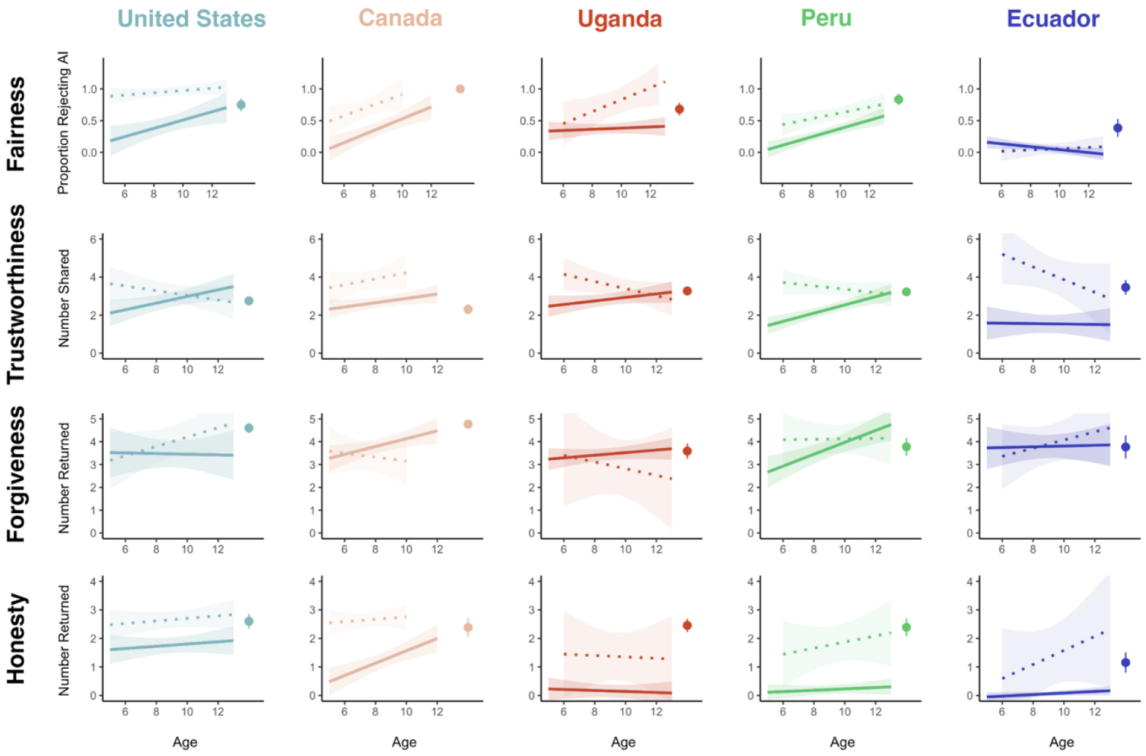

Figure 4: **Convergence of children's behavior and norms toward adult norms.** A *comet plot* showing children's behaviors (solid lines with 95% shaded confidence intervals), children's normative judgments (dashed lines with 95% shaded confidence intervals), and mean adult normative judgments (solid circles with error bars representing standard errors), by age and society. The fairness data is focused on rejections of AI trials.

Children's behaviors and adult norms

Age-related convergence of children's behavior toward adult norms

Looking first at the relationship between children's behaviors and adult norms, we found that children's costly inequity aversion and trustworthiness begin to align with adult norms with age. Specifically, we found that difference scores between children's behaviors and adult norms significantly decreased with age (fairness AI: $\beta = -0.03$, $SE = 0.002$, $p < 0.001$; fairness XAI: $\beta = -0.05$, $SE = 0.01$, $p < 0.001$; trustworthiness: $\beta = -0.08$, $SE = 0.02$, $p < 0.001$). By

contrast, we did not see this age-related pattern when examining children's forgiveness or honesty behavior relative to adult norms (ps > 0.05). Thus, in the context of fairness and trustworthiness, the gap between children's behavior and adult norms generally closes over time, whereas we do not see this same narrowing in the context of forgiveness and honesty. For forgiveness, this likely reflects the high degree of consensus between adults and children from an early age — there is little room for further convergence because both already view forgiveness as the appropriate response.

Society-level differences in the convergence of children's behavior toward adult norms

Next, we explored society-level differences in the gap between children's behavior and adult norms. We found that for fairness (both AI & XAI trials), trustworthiness and honesty, the difference between children's behavior and adult norms varied across society. Specifically, for both fairness and trustworthiness, difference scores in Peru, Uganda, Canada, and the United States were larger than those in Ecuador (fairness (both AI & XAI): *p*s < 0.05; trustworthiness: *p*s < 0.0001), meaning that the gap between what children were doing and what adults viewed as normative was generally more pronounced outside of Ecuador. For honesty, we found that children in Peru, Uganda, and Canada had larger difference scores than children in Ecuador (all *p*s < 0.0001). In the context of forgiveness, on the other hand, we did not see society-level variation in the difference between adult norms and child behavior (all ps > 0.05), revealing quite a high degree of agreement between forgiveness behavior and adult forgiveness norms across sites.

In addition to examining the main effects of age and society, we were interested in whether the age-related patterns we observed in the alignment of children's behavior and adult norms showed variation across societies. Thus, we also checked for the existence of two-way interactions between age and society across our tasks. We found non-significant two-way interactions between age and society in forgiveness ($\chi^2(4) = -3.54$, $p = 0.30$), and a non-significant but marginal two-way interaction for honesty ($\chi^2(4) = -3.93$, $p = 0.06$). For fairness (AI trials) and trustworthiness, however, we found that the gap between children's behavior and adult norms decreased with age in some societies relative to others. Specifically, in the standard AI trials, we observed a two-way interaction between society and age ($\chi^2(4) = 10.54$, $p < 0.001$), difference scores among children in Peru ($\beta = -0.06$, SE = 0.01, $p < 0.001$), Canada ($\beta = -0.1$, SE = 0.01, $p < 0.001$), and the United States ($\beta = -0.06$, SE = 0.01, $p < 0.001$) become significantly smaller across age, as compared to those of children in Ecuador. This mirrored the pattern found in the XAI trial (see **Supplement**). For trustworthiness, where we also observed a two-way interaction between society and age ($\chi^2(4) = 15.63$, $p < 0.001$), we found that difference scores among children in Peru ($\beta = -0.23$, SE = 0.06, $p < 0.001$) and Uganda ($\beta = -0.13$, SE = 0.06, $p = 0.04$) become significantly smaller across age than among children in Ecuador.

What inferences can we draw about the relationship between adult norms and children's behaviors?

Taken together, our comparisons of children's behavior to adult norms revealed differences between what children actually do and the local, adult-endorsed norms that govern what "should" be done. Moreover, this gap between actual behavior and normative behavior showed pronounced variation across age, society and cooperative domain. In general, however, children's fairness and trustworthiness behaviors became gradually more aligned with adult norms with age, yet this pattern was not observed in Ecuador, where fairness and trustworthiness continued to show a gap between behavior and adult norms even in the oldest children (see **Figure 4**). Forgiveness showed a different pattern: we saw high levels of agreement between children's actual behavior and adult norms. One reason for this could be that in our forgiveness task, unlike the other tasks, children were not required to go against self-interest to adhere to the norm. Finally, honesty showed an altogether different pattern. Here, we saw that the gap between children's behavior and adult norms remained pronounced over development, with adults generally endorsing more honesty (i.e., more resources returned to the experimenter) than was reflected in children's actual behavior. This gap was especially large in Peru, Uganda and Canada relative to Ecuador.

There are several inferences that can be drawn from these data. First, norm alignment is domain-specific and culturally moderated: children increasingly align with adult norms over development, but this convergence occurs selectively — most clearly in fairness and trust, less so in honesty, but not in forgiveness, where consensus is already high. Second, local context shapes the extent of convergence; in some cases, such as in Ecuador, the gap persists into later childhood. It's possible that convergence may occur primarily in adolescence in some societies; a question for future work. Together, these patterns suggest that the development of cooperation reflects calibration to community-specific norms rather than uniform progression toward a universal standard.

Children's norms and adult norms

Why are there gaps between children's behaviors and adult norms? One interpretation is that children know what should be done but cannot or will not behave in accordance with local norms — a phenomenon sometimes referred to as the *knowledge-behavior gap* (*24*). In this view, children's normative understanding aligns with adults', but their behavior lags behind. Another possibility is that children's own norms, like their behavior, are *themselves* in flux across development. We were well positioned to examine this question because we elicited normative judgments not only from adults but also from children across all five sites. Although our per-age sample sizes were modest, we were nevertheless able to ask whether children's normative judgments correlated with adult normative judgments.

To do so, we followed a similar method as above by computing difference scores between adult norms and child norms. For fairness (both AI & XAI), trustworthiness, and

forgiveness, we found that the difference between children's normative judgments and adult norms decreased with age (significant effect of age; fairness AI: $\beta = -0.03$, SE = 0.11 $p < 0.0001$; fairness XAI: $\beta = -0.04$, SE = 0.01 $p < 0.001$, trustworthiness: $\beta = -0.12$, SE = 0.04, $p < 0.001$; forgiveness: $\beta = -0.07$, SE = 0.03, $p = 0.04$). Echoing our norms-and-behavior findings from above, this pattern suggests that children's own normative judgments evolve over time, gradually coming into closer alignment with adult judgments.

We also observed significant society-level differences for fairness, trustworthiness and honesty. In the context of fairness, the difference between child and adult norms for standard AI was larger in the United States ($\beta = -0.16$, SE = 0.07, $p = 0.02$) than in Ecuador. This pattern was also found for XAI ($\beta = -0.35$, SE = 0.08, $p < 0.0001$), plus a significant difference between Canada and Ecuador ($\beta = -0.21$, SE = 0.09, $p < 0.05$). For trustworthiness, difference scores were, in general, lower in the United States ($\beta = -0.67$, SE = 0.26, $p = 0.01$) and Peru ($\beta = -0.49$, SE = 0.25, $p = 0.05$), as compared to Ecuador. For honesty, difference scores of children in Canada ($\beta = -0.89$, SE = 0.19, $p < 0.0001$) and the United States ($\beta = -0.77$, SE = 0.19, $p < 0.0001$) were significantly lower than those of kids in Ecuador. We found no evidence for a two-way interaction between society and age across any task (all $ps > 0.05$).

Taken together, these findings suggest that young children's conception of what is "right" in these cooperative contexts does not map cleanly onto adults' normative judgments. Rather, like their behavior, children's normative judgments — at least for fairness, trustworthiness, and forgiveness — appear to undergo developmental refinement, likely reflecting increasing exposure to and internalization of local cultural norms.

Discussion

Human behavior is remarkable in a number of ways. We cooperate at unparalleled levels with other humans, guided by complex cultural rules and norms that help govern these cooperative interactions, and we have an extended period of dependence that allows us to fine-tune and integrate these complex cultural norms into our behaviors. This investigation brings these features together to better understand and contextualize how cooperation can develop across diverse societies. In that vein, this investigation also integrated participants from both WEIRD (Western, Educated, Industrialized, Rich, and Democratic) and non-WEIRD communities, helping correct a persistent bias in the scientific literature (*16*, *39*, *40*).

Bringing together all of our data (as seen in **Figure 4**), along with the analyses detailed above, we wish to highlight a number of insights that this investigation offers into the development of cooperative behavior, more generally. First and most clearly, we found a number of regularities in the development of cooperation, even across diverse societies and subsistence structures. In many cases, we found that children's behaviors tended to converge toward local adults' norms throughout the period of middle childhood. Importantly, children's own normative judgments also appear to be updated in line with the local normative environment, suggesting that both judgments and decisions are being calibrated

to community-specific standards. This is particularly intriguing as the content of the adults' norms themselves can and *do* vary across these societies (see **Supplement**). What this suggests to us is a sensitive and flexible developmental system that functions to calibrate and align both evaluative preferences and decision-making toward local norms during early development. This conclusion is in line with previous work on "*universal norm psychology*" (*18*), which some argue represents a human-unique developmental calibration process of integrating local norms into judgments and behavior. To be clear, we do not provide evidence of universality here, nor do we make that claim, but our results do suggest cross-cultural regularities in the timing of norm integration and the sensitivity of this process to local norms, in line with the argument outlined above. In much the same way as children have both flexible and sensitive learning tools for acquiring a huge diversity of languages — the content of which is filled in through experience — children appear capable of learning other cultural products, such as norms, in a similarly sophisticated and nuanced fashion.

Our analyses also suggest that, rather than behavior across tasks being well-described by a unidimensional and domain-general cooperative phenotype, behavior may be more accurately characterized by distinct, context-sensitive strategies that vary across development and societies. Our exploratory cluster analyses, in this case, revealed three distinct strategies that described behavior across tasks. These strategies — which we termed Maximization, Generic Cooperation, and Partner-Contingent Cooperation — were sensitive to subtle contextual differences across tasks, emerged at different points in development, and were differentially represented across the societies sampled. This diverges from past work suggesting more domain-general cooperative phenotypes in adults (*41*). It may be the case that the documented temporal stability and domain-generality of those behaviors in the literature (*41*) are stronger in adulthood and/or among industrialized populations. That is, they may represent end points in a very specific population, whereas our data capture earlier, more differentiated stages of cooperative behavior in which behavior is still flexible and responsive to local contexts. This finding echoes work on the early emergence of prosocial behaviors (helping, comforting, sharing), which tend to be uncorrelated (*42*). Our investigation thus highlights the importance of considering suites of cooperative behaviors in tandem to allow for stronger inferences and a more nuanced, strategy-based lens on the development of cooperation. Future research can extend this work by testing whether the strategies identified here — and others that may emerge — replicate across additional sites and cultural contexts, helping to refine and strengthen the framework introduced here.

Another key finding from our investigation is the greater degree of self-interest in early life across tasks, which tended to diverge toward more other-regarding behaviors in culturally-specific ways throughout childhood. How do we interpret these results? One (likely incorrect) interpretation is that children in affluent, industrialized societies simply become more 'selfless' than those in other societies. We do not believe this is the case. Instead, it's important to recall that the tasks which involved costly sharing with another peer (the fairness and trustworthiness tasks) always involved anonymous and unnamed recipients of the same gender. It is plausible that what some might define as an *other-regarding* preference in this context is actually best understood as an *unknown-other* regarding

preference. That is, to Shuar children in Ecuador, for instance, the specific partner and their known identity may play a more important role in shaping costly decisions, as opposed to children in other societies. Given the importance of kin systems and social exchange networks to an individual's success (*43*, 44), it's likely that strong preferences for cooperating with unknown others in one-shot games may be more common in industrialized societies, not due to some moral trait of 'selflessness' but rather due to the returns of cooperation in that particular social network (*45*). This claim is further supported by studies among adults suggesting that recipient identities are key considerations for allocation decisions in societies that more heavily rely on informal institutions and kin-based networks (*46–48*).

An additional consideration is that concerns about efficiency or uncertainty about future rewards may contribute to both individual and cross-societal differences(*28*, *49*). For instance, children's tendency to accept advantageous allocations may reflect a pragmatic norm favoring efficient use of resources or minimizing waste, rather than indifference to norms. However, these efficiency-based motivations are likely themselves culturally informed and may coexist with, rather than contradict, broader normative expectations about cooperation. Similarly, greater uncertainty about future access to resources could make immediate maximization more attractive in some contexts, shaping children's responses in tasks involving tangible rewards. It is worth noting, however, that such trade-offs between efficiency, fairness, and self-interest are not unique to this study — they are a feature of virtually all decision-making tasks, across societies and developmental stages. We believe these factors are unlikely to account for the full pattern of results, but may help explain cross-societal variation in how cooperation is expressed, reflecting local calibrations of when, for example, efficiency or equity is valued most strongly.

This work has a number of limitations and opportunities for future research. First, while the sample sizes in this investigation mirror those of past work (e.g. *27*) and are fundamentally constrained by factors such as feasibility and community size, larger sample sizes would permit finer-grained analyses of within-society variability in addition to an assessment of interactions among demographic, ecological, and social factors. They would also allow for a better-powered assessment of some of the interactive effects reported here. Second, our design was cross-sectional, which limits our ability to draw strong comparisons about individual developmental change, especially in communities that are rapidly changing(*50*). We also collected, in the majority of sites, normative judgments and behaviors from *separate* groups of children, which reduces contamination between judgments and behaviors within subjects, but also introduces the limitation that we cannot directly link a given child's normative understanding to their own behavioral choices. Third, although we selected five societies that differ substantially in subsistence strategy and market integration, this set does not, of course, capture the full breadth of human cultural diversity. Future work incorporating additional populations, particularly those undergoing rapid social or economic transition, would help test the generalizability of these patterns and clarify the mechanisms linking local ecology, social norms, and cooperative behavior. Further, our developmental analyses were restricted to children aged 5–13 years. Thus, observed age-related decreases in self-interested behavior and convergence toward local norms reflect

patterns within this range and cannot necessarily be extrapolated to younger or older age groups. Future work including younger children and older adolescents would help clarify when these trajectories first emerge and how long they persist. Finally, while our approach of using average adult norms to compute difference scores provides a clear and interpretable index of developmental convergence, it does not incorporate uncertainty around those means. Future work with larger adult samples could model adult normative variation directly — e.g., via hierarchical or multilevel approaches — to more precisely estimate how children's behaviors track local normative distributions rather than single mean values. Such work could also explicitly design adult sampling strategies to capture within-community heterogeneity in normative beliefs, allowing for more fine-grained analyses of how children calibrate to the range of norms present in their social environment.

In sum, this investigation adopts a novel and integrative approach to understanding the development of cooperative behaviors, norms and strategies across diverse societies, and in this way, provides deeper insights into the nature of human cooperation. We find that children's cooperative behaviors and norms vary across societies, that they tend to diverge in middle childhood, in many cases toward adult norms, and that they appear to be distinct and dissociable behaviors with unique trajectories. Our results further demonstrate that behavior across the four cooperative domains is well-described by three distinct strategies — strategies that change in frequency across development and are differentially represented across societies. The presence and nature of these variations in strategy suggest that rather than developing along a more domain-general trajectory, children calibrate their behavioral strategies to the local social environment in culturally sensitive ways. This combination of a cross-cultural and developmental approach allows us to better assess the regularity with which behaviors occur, and dissociate what appear to be generalizable trends that appear with regularity — e.g. a trend toward more self-interest at younger ages — and the importance of middle childhood as a period of time for norm acquisition from more culturally-variable aspects of human cooperation — e.g. the content of cooperative norms. This work marks an important step toward a serious consideration of cultural values in shaping behavioral development and presents a reproducible model for future work.

Materials and Methods:

Our approach, methods, and analysis strategy were all pre-registered (AsPredicted #24522). Study methods were approved by the institutional review boards at the affiliated institution (Protocol #19.030.01), and the Uganda National Council for Science and Technology (Protocol SS308ES), and adhere to international standards for ethical human subjects research. All adult participants provided informed consent prior to participation, and all child participants provided verbal assent. In Uganda and Ecuador, additional community consent was obtained from local leaders in accordance with cultural and institutional ethical standards.

Children's Task Battery

Participants

Our participants were N = 413 children between the ages of 5 and 13 years old (M = 8.78, SD = 2.40, 214 girls) across five societies. An additional 64 children were tested but their data in one or more tasks were excluded for the following pre-registered reasons: participant failure to follow instructions (n=22), outside of age range (n=4), participant revoked assent (n=2), neurodevelopmental disorders (n=2), experimenter error (n=28), participant didn't like resource (n=3), significant disruption (n=2), and previous participation in the task (n=1). A breakdown of exclusions by task can be found in the **Supplement**.

Our sampling strategy was pre-registered with a goal of approximately 60-90 children between the ages of 5 and 13 in each society. In recruiting, we aimed for approximately 20-30 children in each of the three age groups — 5-7 year olds, 8-10 year olds, and 11-13 year olds — with the decision to test closer to 30 or to 20 in a given society made solely based on participant availability during the times of our data collection trips. In Uganda, our targets were slightly exceeded due to high levels of interest in the community. Our pre-registered recruitment targets were informed by experience from previous field research at these sites(*27*, *28*).

Procedure

The protocol began with an introductory section, followed by four tasks — fairness, trustworthiness, forgiveness, and honesty — presented in two orders (forward and reverse) that were counterbalanced across participants, and a concluding section. We find no order effects on behavior — see **Supplement**. The script's language matched the primary language across sites. Scripts were collaboratively translated and reverse-translated with the input of native speakers and professional translators.

The fairness, forgiveness, and trust tasks involved a number of comprehension questions to ensure children understood the mechanics of the task and how to make their choices accurately. Comprehension question success rates were extremely high, with first-answer success rates an average of 98.6% in the fairness task, 99.5% in the trust task, and 100% in the forgiveness task. There were no comprehension questions in the short honesty task. If children failed to answer correctly, the experimenter would give them the correct information and repeat the comprehension check again. If they failed a second time (which was exceedingly rare), the experimenter took note, repeated the correct information once more, and continued with the protocol.

In the introductory section, the experimenter introduced themselves, and explained that the participant would get a chance to play a number of activities in which they could take home some candy. The experimenter then introduced the children to small, fruit-flavored candies called Skittles and Starbursts. Each society was randomly assigned one

color of both Skittles and Starbursts. The participants were given a chance to try both candies and report if they liked them. The experimenter only continued if the participant shared that they either liked the candy, or liked the candy enough to play.

The experimenter then explained that all of their candies from the session would be put into a paper bag with a brown sticker on it. The brown sticker indicated that the bag belonged to the participant. The experimenter then brought out three other paper bags, with three different colored stickers on them — gray, black, and white — and explained that these bags belonged to three other children, gender-matched to the participant, that could not be here that day, but they would get their bags later, and that they all really liked candy, just like the participant. The participant was then given two general comprehension questions to ensure that they understood two key points: that the other children also liked candy, and that the other children would get their bags later.

The fairness task was focused on the development of inequity aversion, assayed through rejections of unfair offers in the Inequity Game. The apparatus had two trays — one for the participant, and one for the other child — with two corresponding bowls for each player, and one central "rejection bowl". The participant sat in front of two handles, one green and one red. The other child was represented by the paper bag with a gray sticker on it. The experimenter explained that the Skittles put on the tray closer to the participant were for them, and the Skittles put on the farther tray were for the child with the gray-sticker paper bag. The experimenter explained through demonstration with one skittle per tray that if the participant pulled the green handle, the trays would tilt outward to the respective bowls of the participant and the other child; the red lever would make the trays tilt inward, and no one would get those candies. There were then two comprehension questions to ensure that the participant understood which tray belonged to which child.

Following these instructions, the experimenter then led the participant through three narrated, lower-cost trials. We will refer to these as "practice trials" because they familiarized children to the mechanics of the apparatus, but they were not called such to the participant, and came with real costs and rewards. The protocol closely mirrored that of previous research utilizing this apparatus. The first practice trial was always an Equal trial (1-1). After loading the trays, the experimenter prompted the participant to pull a lever, and narrated the outcome. The second and third practice trials were one Disadvantageous Inequity (DI) trial (0 for the participant, 1 for the other child) and one Advantageous Inequity (AI) trial (1 for the participant, 0 for the other child), counterbalanced. If the participant chose one color handle for all three practice trials, they were asked to indicate which handle would lead to the alternative outcome (e.g. asked "which handle will make the candy drop into the middle?" if they picked the green handle for all three practice trials).

After the three practice trials, the participant then moved on to seven test trials. Four of these test trials were AI trials where the participant received four candies, and the other child received one. Two of these trials were Equal trials, in which both parties received one candy. The order of these six trials was randomized. The last trial was always an Extreme AI (XAI) trial, where the participant received four candies, and the other child received none. The dependent variable in the Fairness Task was whether the participant accepted or

rejected each type of offer. In AI trials, rejections indicated costly adherence to fairness. For Equal trials, outcomes were identical for both parties (0-0 if rejected, 1-1 if accepted), and thus rejections do not indicate fairness for this trial type.

In the Trustworthiness Task, the experimenter introduced participants to a child-friendly version of the Trust Game. This game was played with a tray-like apparatus with high sides and a plexiglass covering. The inside of the tray was split into two channels, within which were two small, metal dishes filled with Starbursts. The first dish had two Starbursts in it and was much closer to the opening on the side of the tray. The other dish was further in, and had six Starbursts in it. The experimenter explained that in this game, the participant could not reach in through the top, and the only way to reach the candies inside was to use a special tool — a short, plastic tube with a hook on the end. The experimenter demonstrated that the tool could be used to hook the closer dish with two Starbursts in it, but it was too short to reach the dish with six Starbursts. The only way to reach the dish with six Starbursts was to join two short tools together using a connector piece to make one long tool. This long tool could reach the dish with six Starbursts. Participants were then asked a comprehension question to ensure they understood that only the long tool could reach the further dish.

The experimenter then gave the participant one short tool, and explained that another child who really likes Starbursts played this game before them. The paper bag with a black sticker was their bag, and stood in for them. The experimenter explained that the previous child was given a choice between either using their short tool to reach the dish with two Starbursts, or giving their short tool to the participant. The other child chose to give up their tool, forfeiting their chance at two Starbursts in order to give the participant access to the dish with six Starbursts. They knew that the participant could share some of these six Starbursts with them today. The experimenter then asked two comprehension questions to ensure the child understood that the other player made the choice to give the participant their tool, and that they did not receive any Starburst as a result.

The experimenter then gave the second tool to the participant, instructing them to assemble the longer tool and pull out the dish with six Starbursts. Upon retrieving the dish with six Starbursts, the experimenter asked the participant to decide how many of the Starbursts to keep for themselves versus give to the other child, by allocating the Starbursts between their respective paper bags. To give the participant privacy, the experimenter pulled out a bell, and told the participant they would be turning around to finish working, but that the participant should ring the bell when they are done allocating the candies. If the child did not ring the bell within 15 seconds, the experimenter, without turning around, asked whether they were done making their decision. After the bell was rung, the experimenter turned around, and ended the activity. The dependent variable in the Trustworthiness Task was the number of Starbursts the participant shared with the other player.

In the Forgiveness Task, experimenters introduced participants to an activity which involved the use of a novel resource: a magic color scratch sheet. These sheets were small pieces of rainbow-colored paper with a black, carbon film on top. Using a stylus, the experimenter demonstrated that by scratching off the carbon film, the colors underneath could be revealed. The experimenter instructed the child to practice drawing on a test sheet

for 15 seconds. They then told the participant that there were five more of these sheets for them to draw on. However, when they pulled out the small box that was supposed to hold the sheets, they discovered that the sheets had already been drawn on. Experimenters were instructed to express surprise and disappointment, by saying, “Oh no, it looks like someone before you already drew on them and you won’t get to complete this activity. We don’t have any more of these.” The experimenter then revealed a note in the box from another experimenter which stated, “A kid that was here before you did not know these were for you, and they drew on them, accidentally. They are really sorry that they used the scratch sheets.” This is analogous to the Accidental-Remorse condition from Amir (*33*). Following this, the experimenter then asked the participant to make a decision about the scratch sheets that had already been drawn on. They could throw as many of the sheets as they wanted away into a small trash can, or they could return as many of the sheets as they wanted to the child that drew on them by placing them into the corresponding paper bag with a white sticker on it. The experimenter then asked two comprehension questions to ensure the participant knew how to enact the two options, and then turned around and instructed the participant to ring the bell when they were done with their decision. The dependent variable in the Forgiveness task was the number of sheets the participant returned to the other player.

In the Honesty Task, the experimenter told the participant that as a prize for participating, they would receive a gift of one Starburst. The experimenter handed the participant a small, cardboard box and reiterated that there should be one Starburst inside. They then asked the participant, “Can you make sure you can open the box? I’m going to turn around and finish writing something down”. Upon opening the box, the participant learned that instead of one Starburst, there were actually four Starbursts inside. The experimenter then waited to see what the child would do. The experimenter remained turned around for 15 seconds, during which they waited for a possible interjection from the participant. If the participant did not speak during the first 15 seconds, they were instructed to ask, without turning around, “Are you able to open the box?”. In both cases, the experimenters were trained to respond to the exact type of utterance from the participant. If the participant mentioned anything about the number of Starbursts (e.g. “there are four in here”), the experimenter said, “Oh, okay. If there are any extra, can you give them back to me?”. The number of candies returned was recorded. If the participant mentioned anything else unrelated to the number of candies, the experimenter said, “Ok, you are all set. You can put the box in your bag to take home”. The task was considered finished at that point, unless the participant chose to return Starbursts after the experimenter’s statement. The dependent variable in the Honesty task was the number of candies returned to the experimenter.

In the conclusion section, the experimenter first asked whether or not the three paper bags belonged to other children, as a measure of the believability of the setup. They then did a simple give-N counting task with the participant, introducing them to ten wooden tokens, and asking them to first “put one token in my hand”, and then put “exactly six tokens” in their hand (see **Supplement** for results). The experimenter then asked a number of demographic questions about family composition, parental occupation, and household religious practices.

Children's Norm Battery

Participants

A separate group of children were recruited to complete the children's norm battery survey, with the exception of Ecuador, in which there was overlap between participants in the task battery and the norm battery, but a period of at least three days in between testing. This was due to a lack of additional participants in the communities tested in Ecuador, which were constrained by community size and the logistical affordances of the field season. Our participants were N = 163 children between the ages of 5 and 13 years old (M = 8.69, SD = 2.34, 76 girls) across five societies.

Protocol

The experimenter explained to the participant that they would be reading a few stories out loud and asking a few questions. The experimenter stated that there are no right or wrong answers, and that they were solely interested in what the participant believed. They then began the norm elicitation section, in which they read out loud five different vignettes that mirrored the choices that children faced in the behavioral task battery, in either forward or reversed order, counterbalanced.

In the Fairness vignette, participants were introduced to an adult who was allocating candy to two different children, gender-matched to the participant. The experimenter explained that the first child could decide whether they wanted to say "yes" or "no" to the division. There were six trials with related diagrams that corresponded with the following allocations: 1-1, 2-1, 3-1, 4-1, 4-0, and 1-4, for the first child (the actor) and the second child (the recipient), respectively. For each trial, the experimenter asked, "Should the first [child] say yes or no? What's the right thing to do?".

In the Generosity Norm vignette, participants were asked how a child should split four candies between themselves and another child. They were asked whether the first child should take zero, one, two, three, or four for themselves, and instructed to answer, "What's the right thing to do?"

In the Trustworthiness Norm vignette, the experimenter read a story in which a child has two seeds, but no garden. The child gives the seeds to a second child, who plants them in their garden. The seeds then grow into six fruits. The experimenter then asked the participant, from zero to six, how many fruits the child with the garden should share back with the child who gave them the seeds, by asking, "What's the right thing to do?".

In the Forgiveness Norm vignette, the experimenter read a story about a child who has five special pieces of paper that they wanted to make five drawings with. But, a second child accidentally uses those five pieces of paper to make their own drawings, and the papers can't be reused. The first child must now decide how many of the sheets to give back

to the other child, and how many to throw away. The experimenter asked the participant, from zero to five, how many of the drawings to return, by asking, “What’s the right thing to do?”

In the Honesty Norm vignette, the experimenter read a story in which a child did a good job in school, and as a result, their teacher said they deserved one gift for doing a good job. But when the child opened the bag from the teacher, there were actually four gifts inside. The experimenter asked the participant, on a scale of zero to four, how many of the gifts to keep, and asked “What’s the right thing to do?”

Following this block, the experimenter then moved to the Judgment section, in which they recounted each of these vignettes, and in each case, described that the child did the self-interested behavior. That is, in the Fairness vignette, they accepted an offer of 4-1. In Generosity, they decided to keep all four candies. In Trustworthiness, they kept all six fruits for themselves. In Forgiveness, they threw all five sheets away. And in Honesty, they kept all four of the gifts. The experimenter then asked the participant, for each of these cases, to imagine that either an adult or another kid their age saw the child doing this behavior. In those cases, would the child or peer think this “was a good or bad thing to do?”. Upon answering, they were instructed to describe how bad or good (e.g. “A little good or very good?”.

In the Concluding block, the experimenter conducted the Demographic Survey, which asked about household composition, parental occupation, and family religion. All children received a prize for participating.

Adult Norm Battery

Participants

A group of adults were recruited from each community to complete the adult norm battery survey. Our participants were N = 86 adults between the ages of 18 and 74 years old (M = 32.6, SD = 12.6, 35 men) across five societies.

Protocol

The procedure for adults was similar to that of the children’s norm battery, with identical instructions and vignettes. The only difference was that the first pass of each vignette all involved adults, and participants were asked about the normative behavior for adults. Then, they were asked to imagine that the participants were 10-year-old children, then 5-year-old children, to determine whether the norm changed by target age. Participants then completed the Judgment section, just as in the children’s norm battery, which was followed by a short demographic section.

4.4 Statistical Methods

We used R version 4.0.0 for all analyses (*51*). Across the behavioral tasks, we focused on several, pre-registered dependent variables. In the Fairness Task, the dependent variable was whether a participant rejected on a given trial (binary, accept or reject). In the Trustworthiness Task, the dependent measure was the number of candies shared with the trustor (numeric, 0-6). In the Forgiveness Task, the dependent variable was the number of sheets given back to the transgressor (numeric, 0-5). In the Honesty Task, the dependent variable was the number of candies returned to the experimenter (numeric, 0-4). The primary, pre-registered predictors were age (numeric), and society (factor, five levels). To address whether children's behaviors and normative judgments converged toward adult norms with age, we computed difference scores between children's behaviors and the behaviors adults endorsed as normative as our dependent variable in a number of linear probability models, where age (continuous) and society (factor, five levels) were the predictors. To assess the interrelatedness of the four cooperative behaviors, we computed Spearman correlations and conducted significant testing using the *stats* package in R (*51*).

References

1. A. P. Melis, D. Semmann, How is human cooperation different? *Philos Trans R Soc Lond B Biol Sci* **365**, 2663–2674 (2010).

2. E. Fehr, K. M. Schmidt, A theory of fairness, competition, and cooperation. *The quarterly journal of economics* **114**, 817–868 (1999).

3. D. Amir, D. Melnikoff, F. Warneken, P. R. Blake, J. Corbit, T. C. Callaghan, O. Barry, A. Bowie, L. Kleutsch, K. L. Kramer, E. Ross, H. Vongsachang, R. Wrangham, K. McAuliffe, Computational signatures of inequity aversion in children across seven societies. *Journal of Experimental Psychology: General*, 2882–2896 (2023).

4. D. Amir, W. S. Parsons, R. E. Ahl, K. McAuliffe, Trustworthiness is distinct from generosity in children. *Developmental Psychology* **57**, 1318 (2021).

5. R. Borum, The Science of Interpersonal Trust. *Mental Health Law & Policy Faculty Publications* 574 (2010).

6. H. C. Godfray, The evolution of forgiveness. *Nature* **355**, 206–207 (1992).

7. D. E. Forster, J. Billingsley, V. M. Russell, T. G. McCauley, A. Smith, J. L. Burnette, Y. Ohtsubo, J. Schug, D. Lieberman, M. E. McCullough, Forgiveness takes place on an attitudinal continuum from hostility to friendliness: Toward a closer union of forgiveness theory and measurement. *Journal of Personality and Social Psychology*, 861–880 (2019).

8. M. E. McCullough, Forgiveness: Who does it and how do they do it? *Current directions in psychological science* **10**, 194–197 (2001).

9. A. Szekely, D. Bruner, S. Steinmo, A. Todor, C. Volintiru, G. Andrighetto, Preferences for honesty can support cooperation. *Journal of Behavioral Decision Making* **36**, e2328 (2022).

10. C. Bicchieri, R. Muldoon, A. Sontuoso, Social norms. *The Stanford encyclopedia of philosophy* (2011).

11. C. Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms* (Cambridge University Press, 2005).

12. B. R. House, How do social norms influence prosocial development? *Current Opinion in Psychology* **20**, 87–91 (2018).

13. J. Henrich, R. McElreath, A. Barr, J. Ensminger, C. Barrett, A. Bolyanatz, J. C. Cardenas, M. Gurven, E. Gwako, N. Henrich, Costly punishment across human societies. *Science* **312**, 1767–1770 (2006).

14. B. R. House, J. B. Silk, J. Henrich, H. C. Barrett, B. A. Scelza, A. H. Boyette, B. S. Hewlett, R. McElreath, S. Laurence, Ontogeny of prosocial behavior across diverse societies. *Proceedings of the National Academy of Sciences* **110**, 14586–14591 (2013).

15. B. R. House, P. Kanngiesser, H. C. Barrett, S. Yilmaz, A. M. Smith, C. Sebastian-Enesco, A. Erut, J. B. Silk, Social norms and cultural diversity in the development of third-party

punishment. *Proceedings of the Royal Society B: Biological Sciences* **287**, 20192794 (2020).

16. D. Amir, K. McAuliffe, Cross-cultural, developmental psychology: integrating approaches and key insights. *Evolution and Human Behavior* **41**, 430–444 (2020).

17. H. Kaplan, J. Lancaster, A. Robson, Embodied capital and the evolutionary economics of the human life span. *Population and Development Review* **29**, 152–182 (2003).

18. B. R. House, P. Kanngiesser, H. C. Barrett, T. Broesch, S. Cebioglu, A. N. Crittenden, A. Erut, S. Lew-Levy, C. Sebastian-Enesco, A. M. Smith, S. Yilmaz, J. B. Silk, Universal norm psychology leads to societal diversity in prosocial behaviour and development. *Nature Human Behaviour* **4**, 1–11 (2019).

19. P. Rochat, M. D. G. Dias, Guo Liping, T. Broesch, C. Passos-Ferreira, A. Winning, B. Berg, Fairness in Distributive Justice by 3- and 5-Year-Olds Across Seven Cultures. *Journal of Cross-Cultural Psychology* **40**, 416–442 (2009).

20. E. Huppert, J. M. Cowell, Y. Cheng, C. Contreras-Ibáñez, N. Gomez-Sicard, L. M. Gonzalez-Gaeda, D. Huepe, A. Ibanez, K. Lee, R. Mahasneh, S. Malcolm-Smith, N. Salas, B. Selcuk, B. Tungodden, A. Wong, X. Zhou, J. Decety, The development of children's preferences for equality and equity across 13 individualistic and collectivist cultures. *Developmental Science* **22**, e12729 (2018).

21. M. Paulus, Children's inequity aversion depends on culture: A cross-cultural comparison. *Journal of Experimental Child Psychology* **132**, 240–246 (2015).

22. M. Schäfer, D. B. M. Haun, M. Tomasello, Fair Is Not Fair Everywhere. *Psychological Science* **26**, 1252–1260 (2015).

23. H. Zeidler, E. Herrmann, D. BM Haun, M. Tomasello, Taking turns or not? Children's approach to limited resource problems in three different cultures. *Child development* **87**, 677–688 (2016).

24. P. R. Blake, K. McAuliffe, F. Warneken, The developmental origins of fairness: the knowledge–behavior gap. *Trends in Cognitive Sciences* **18**, 559–561 (2014).

25. C. A. Kallgren, R. R. Reno, R. B. Cialdini, A Focus Theory of Normative Conduct: When Norms Do and Do not Affect Behavior. *Pers Soc Psychol Bull* **26**, 1002–1012 (2000).

26. E. Fehr, I. Schurtenberger, Normative foundations of human cooperation. *Nat Hum Behav* **2**, 458–468 (2018).

27. P. R. Blake, K. McAuliffe, J. Corbit, T. C. Callaghan, O. Barry, A. Bowie, L. Kleutsch, K. L. Kramer, E. Ross, H. Vongsachang, R. Wrangham, F. Warneken, The ontogeny of fairness in seven societies. *Nature* **528**, 258–261 (2015).

28. D. Amir, M. R. Jordan, K. McAuliffe, C. R. Valeggia, L. S. Sugiyama, R. G. Bribiescas, J. J. Snodgrass, Y. Dunham, The developmental origins of risk and time preferences across diverse societies. *Journal of Experimental Psychology: General* **149**, 651–661 (2019).

29. P. R. Blake, K. McAuliffe, “I had so much it didn’t seem fair”: Eight-year-olds reject two forms of inequity. *Cognition* **120**, 215–224 (2011).

30. K. McAuliffe, P. R. Blake, F. Warneken, Costly fairness in children is influenced by who is watching. *Developmental Psychology* **56**, 773–782 (2020).

31. A. Shaw, N. Montinari, M. Piovesan, K. R. Olson, F. Gino, M. I. Norton, Children develop a veil of fairness. *Journal of Experimental Psychology: General* **143**, 363–375 (2014).

32. J. Berg, J. Dickhaut, K. McCabe, Trust, reciprocity, and social history. *Games and economic behavior* **10**, 122–142 (1995).

33. D. Amir, R. E. Ahl, W. S. Parsons, K. McAuliffe, Children are more forgiving of accidental harms across development. *Journal of Experimental Child Psychology* **205**, 105081 (2021).

34. R. Gomila, Logistic or linear? Estimating causal effects of experimental treatments on binary outcomes using regression analysis. *Journal of Experimental Psychology: General* **150**, 700–709 (2021).

35. D. Amir, C. Valeggia, M. Srinivasan, L. S. Sugiyama, Y. Dunham, Measuring subjective social status in children of diverse societies. *PloS one* **14**, e0226550 (2019).

36. A. Shaw, K. R. Olson, Children discard a resource to avoid inequity. *Journal of Experimental Psychology: General* **141**, 382 (2012).

37. I. Almås, A. W. Cappelen, E. Ø. Sørensen, B. Tungodden, Fairness and the Development of Inequality Acceptance. *Science* **328**, 1176–1178 (2010).

38. E. Fehr, H. Bernhard, B. Rockenbach, Egalitarianism in young children. *Nature* **454**, 1079–1083 (2008).

39. J. Henrich, S. J. Heine, A. Norenzayan, The weirdest people in the world? *Behav Brain Sci* **33**, 61–135 (2010).

40. M. Nielsen, D. Haun, J. Kärtner, C. H. Legare, The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology* **162**, 31–38 (2017).

41. A. Peysakhovich, M. A. Nowak, D. G. Rand, Humans display a ‘cooperative phenotype’ that is domain general and temporally stable. *Nature communications* **5**, 4939 (2014).

42. K. A. Dunfield, A construct divided: prosocial behavior as helping, sharing, and comforting subtypes. *Front. Psychol.* **5**, 958 (2014).

43. J. F. Schulz, Kin networks and institutional development. *The Economic Journal* **132**, 2578–2613 (2022).

44. B. Enke, Kinship, Cooperation, and the Evolution of Moral Systems*. *The Quarterly Journal of Economics* **134**, 953–1019 (2019).

45. J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, R. McElreath, M. Alvard, A.

Barr, J. Ensminger, N. S. Henrich, K. Hill, F. Gil-White, M. Gurven, F. W. Marlowe, J. Q. Patton, D. Tracer, "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences* **28**, 795–815 (2005).

46. M. Fafchamps, F. Gubert, The formation of risk sharing networks. *Journal of Development Economics* **83**, 326–350 (2007).

47. M. M. Gervais, RICH Economic Games for Networked Relationships and Communities: Development and Preliminary Validation in Yasawa, Fiji. *Field Methods* **29**, 113–129 (2017).

48. D. Bahrami-Rad, J. Beauchamp, J. Henrich, J. Schulz, Kin-based institutions and economic development. SSRN. [Preprint] (2022). https://dx.doi.org/10.2139/ssrn.4200629.

49. R. E. Ahl, K. Hannan, D. Amir, A. Baker, M. Sheskin, K. McAuliffe, Tokens of virtue: Replicating incentivized measures of children's prosocial behavior with online methods and virtual resources. *Cognitive Development* **66**, 101313 (2023).

50. S. S. Urlacher, M. A. Liebert, J. Josh Snodgrass, A. D. Blackwell, T. J. Cepon-Robins, T. E. Gildner, F. C. Madimenos, D. Amir, R. G. Bribiescas, L. S. Sugiyama, Heterogeneous effects of market integration on sub-adult body size and nutritional status among the Shuar of Amazonian Ecuador. *Annals of human biology* **43**, 316–329 (2016).

51. R Core Team (2020) R: A language and environment for statistical computing. R: Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

52. A. D. Blackwell, J. J. Snodgrass, F. C. Madimenos, L. S. Sugiyama, Life history, immune function, and intestinal helminths: Trade-offs among immunoglobulin E, C-reactive protein, and growth in an Amazonian population. *American Journal of Human Biology* **22**, 836–848 (2010).

53. T. Callaghan, H. Moll, H. Rakoczy, F. Warneken, U. Liszkowski, T. Behne, M. Tomasello, W. A. Collins, Early social cognition in three cultural contexts. *Monographs of the Society for Research in Child Development* **76**, 1–142 (2011).

54. B. D. Jokisch, K. McSweeney, Assessing the potential of indigenous-run demographic/health surveys: The 2005 Shuar Survey, Ecuador. *Human Ecology* **39**, 683–698 (2011).

55. M. Leeuwen, I. Maas, A. Miles, *HISCO: Historical International Standard Classification of Occupations* (Leuven: Leuven University Press, 2002).

56. M. A. Liebert, J. J. Snodgrass, F. C. Madimenos, T. J. Cepon, A. D. Blackwell, L. S. Sugiyama, Implications of market integration for cardiovascular and metabolic health among an indigenous Amazonian Ecuadorian population. *Annals of human biology* **40**, 228–242 (2013).

57. F. C. Madimenos, J. J. Snodgrass, A. D. Blackwell, M. A. Liebert, T. J. Cepon, L. S. Sugiyama, Normative calcaneal quantitative ultrasound data for the indigenous Shuar and non-Shuar Colonos of the Ecuadorian Amazon. *Archives of osteoporosis* **6**, 39–49 (2011).

58. F. C. Madimenos, J. J. Snodgrass, M. A. Liebert, T. J. Cepon, L. S. Sugiyama, Reproductive effects on skeletal health in Shuar women of Amazonian Ecuador: A life history perspective. *American Journal of Human Biology* **24**, 841–852 (2012).

59. S. Rubenstein, Colonialism, the Shuar Federation, and the Ecuadorian state. *Environment and Planning D* **19**, 263–294 (2001).

Acknowledgments

We would first and foremost like to acknowledge the families, schools, and communities that participated in this research. We're grateful for the support of the Shuar Health and Life History Project, the Shuar Federation, the Kibale Forest Schools Program, and the Kasiisi Project. We would also like to thank Aaron Baker for creating the new occupational coding scheme found in the Supplement, and for his efforts and support, more generally. We would also like to thank several colleagues for helpful feedback, support, and commentary: Elizabeth Ross, Peter Blake, and members of the Boston College Cooperation Lab and Morality Lab.

Funding

This work was primarily funded through the John Templeton Foundation (#61138, PIs: Katherine McAuliffe & Liane Young, Boston College).

Author Contributions

Conceptualization — DA, REA, KM
Methodology — DA, REA, MRJ, KM
Formal Analysis — DA, MRJ, KM
Investigation — DA, REA, MRJ, HB, MG, GTG, TC, LSS, EO, PT, SB, JJS, KM
Writing – Original Draft — DA, MRJ, KM
Writing – Review & Editing — DA, MRJ, REA, KM
Supervision — DA, HB, MB, GTG, TC, LSS, EO, PT, SB, JJS, KM
Funding Acquisition — KM

Competing Interests

All authors declare they have no competing interests.

Data and Materials Availability:

All materials, original data, pre-registration file, and analysis scripts are available here: https://osf.io/d64qn/ The exact scripts and protocol can be found in the Supplement.