Supporting Information (SI) Appendix for:

The Ontogeny of Prosocial Behavior across Diverse Societies

Bailey R. House * , Joan B. Silk, Joseph Henrich, H. Clark Barrett, Brooke Scelza, Adam Boyette, Barry Hewlett, and Stephen Laurence

*To whom correspondence should be addressed. E-mail: bhouse80@ucla.edu

This file includes: Page

SI Appendix 1: Additional details about methods

Instructions for Experimenters

Rewards for Subjects. The payoffs used are small food items, such as candies. Before being included in gameplay, these items are occluded from view (to avoid distracting attention). Keeping them out of sight in a bag or pocket allows them to be easily retrieved when loading the apparatus at the start of each trial. When the Actor selects a particular payoff distribution, the experimenter removes the unselected payoffs, and encourages the subjects to take their candies.

[Note: candy was not used as the reward at all sites, see Table S1(b)]

Apparatus. The choices will be presented to children on two large paper trays. On each tray there are two circles (blue on the actor's side, red on the recipient's side). Food rewards will be placed inside the circles during the Familiarization and Test trials.

Trials. Within each session, the order of trials is identical across conditions, but the sides on which the payoffs are presented are reversed between Asocial and Social (e.g., if in the Prosocial Game the 1/1 payoff is on the left in the Social condition, it will be on the right in the Asocial condition). Across sessions, the order of Test trials is counter-balanced. The order of familiarization trials is kept constant across sessions. [The order of familiarization trials is kept constant because pilot testing with adults suggested that subjects gain a clearer understanding of the instructions when Familiarization #1 comes first. This is because both Actor and Recipient could receive a payoff in Familiarization #1, whereas in Familiarization #2 the Recipient cannot receive anything. This can cause confusion.]

Participant roles. To maximize the number of pairings that we can create, most of the children will participate as an Actor and then as a Recipient (except for the Recipient in the first pairing, who will never be an Actor). Children will play the Actor role first, then become the Recipient in a pairing with a new Actor. With a set of 6 children (A, B, C, D, E, and F) 5 pairings would be constructed as follows (see Table S1[a]):

SI Table S1(a): Pairing of child participants.

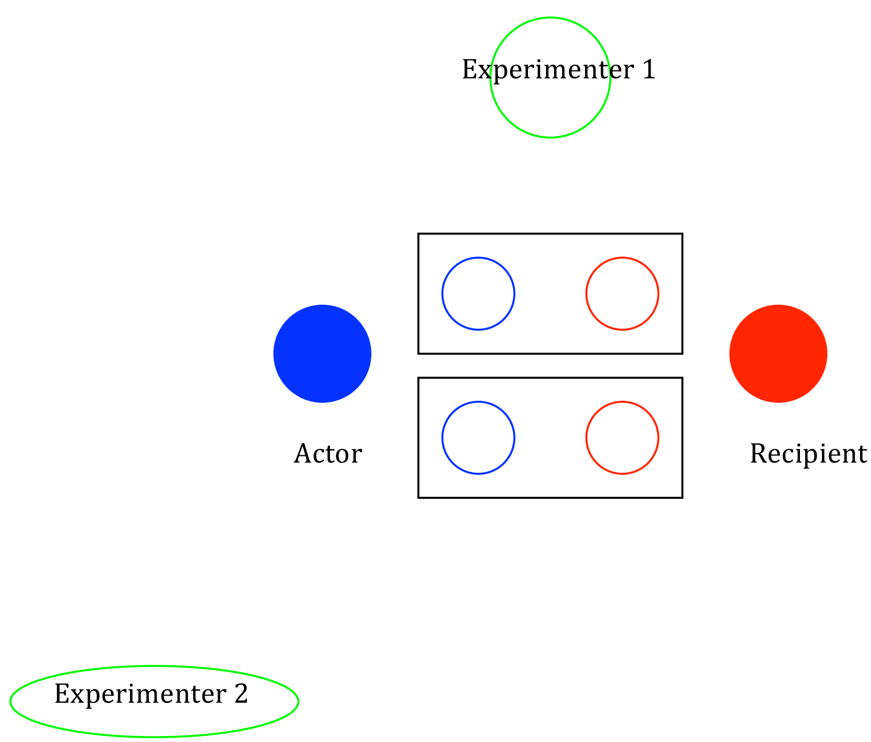
Pairing (study session #)	Actor	Recipient
1	**B**	A*
2	C	**B**
3	D	C
4	E	D
5	F**	E

*A begins as Recipient, never plays Actor role (because Actors must be naïve to the task)
**F begins as Actor, but never plays Recipient role (because there are no naïve Actors left)

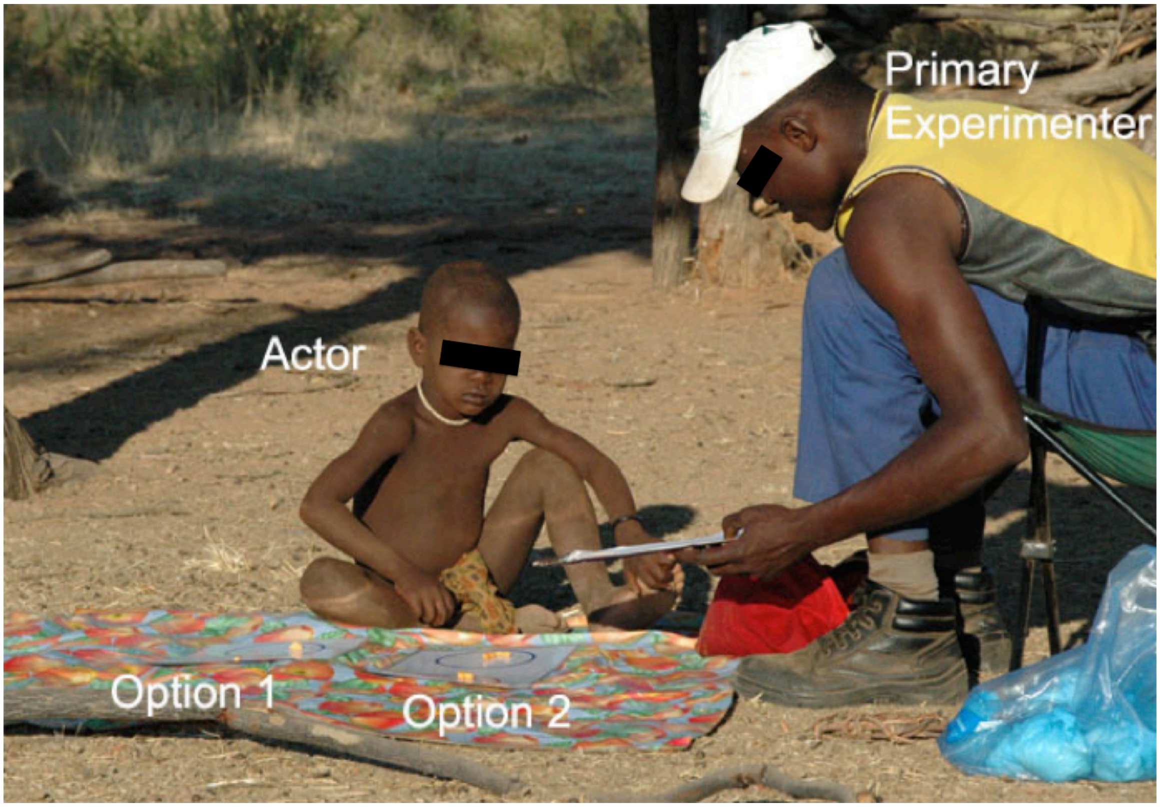
Note that the Actor in each pairing is always a naïve subject, who has not participated in the game yet. If it is a concern that the first recipient ('A' in the above example) plays the game fewer times than the other subjects, the first recipient could later play with the last Actor ('F' in the above example), who also plays fewer times than the other subjects. However, this pairing should NOT be included as data in the sample, because 'A' was not a naïve subject when playing the role of Actor. If problems with pairings occur, it is ok for a child to play Recipient in multiple pairings after he/she has been an Actor, but not vice versa. Genetically-related siblings should not be paired together, but other pairings are ok.

General Procedure/Setup. Familiarization trials and Test trials are run in a block. First the Familiarization trials are run, and then the Test trials. There will be a set of Familiarization and Test trials for the Social condition, and a separate set of Familiarization and Test trials for the Asocial condition. The ordering of the Social block of trials and Asocial block of trials will be randomized. One experimenter will explain the game & distribute the candies; if possible, a second experimenter sitting away from the children (preferentially out of Actor's line of sight) will record responses live. The children can be seated at a table or on the floor (Figure S1[a]).

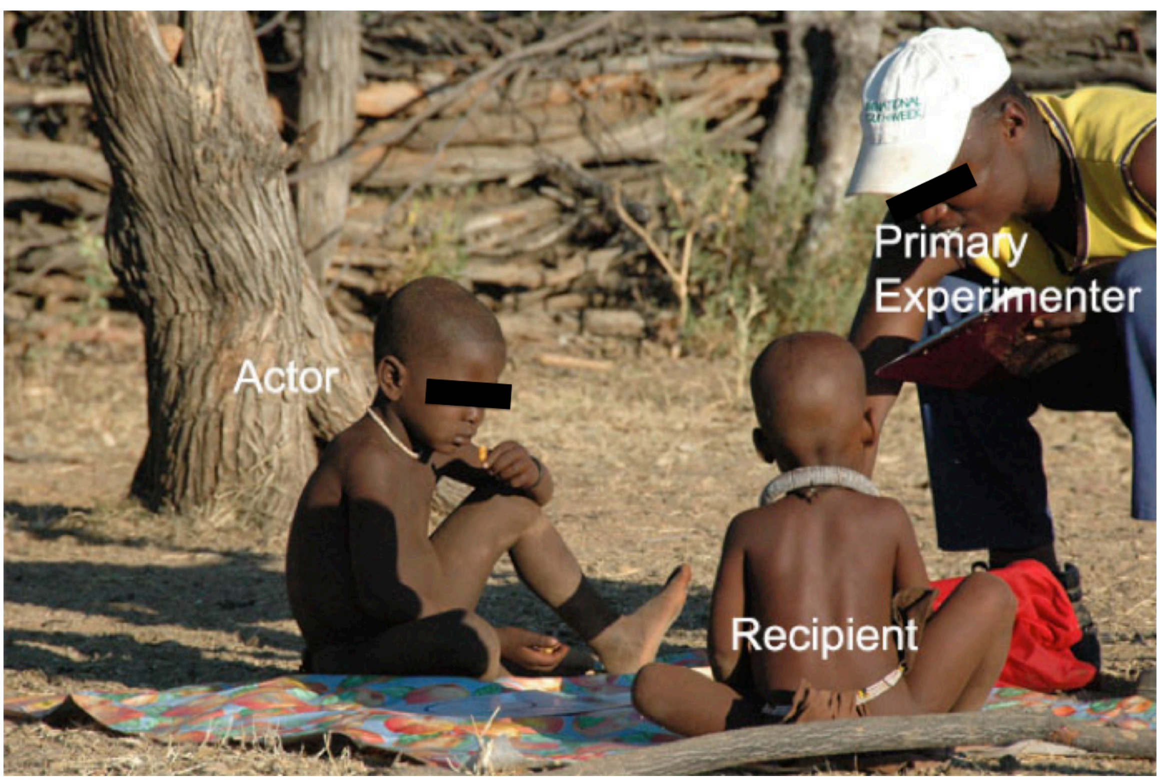
SI Figure S1(a): Experimental setup



SI Figure S1(b): Asocial condition



SI Figure S1(c): Social condition



***[Scripts:** we provide the example scripts distributed to all experimenters, which were translated as needed for each site. Text in brackets was instructions for experimenters, and was not read to participants. Where scripts refer to candies, this was modified as needed to reflect whatever rewards were used at each site (see Table S1(b) for rewards used at different sites).]*

Script for child subjects

"Okay! On each of these papers *[trays]*, there are two circles, a blue circle and a red circle." **[Point to blue circle and red circle.]**

For Familiarization #1 and #2 (both Social and Asocial conditions)

[Move Left tray in front of child.]

"I put a CANDY in the blue circle which is next to you. Now I put another candy in the blue circle." **[Place candy on blue circle.]**

"I put a candy in the red circle which is next to [name of Recipient]. Now I put another candy in the red circle." **[Place candy on red circle.]** *{Note: in Familiarization #2 the experimenter says 'I don't put anything in the red circle next to [name of Recipient]'}*

{Note: in the Asocial condition the experimenter replaces 'next to [name of Recipient]' with 'on the other side'}

[Move Right tray in front of subject.]

"I put a CANDY in the blue circle which is next to you. "**[Place candy on blue circle.]**

"I put a CANDY in the red circle which is next to [name of Recipient]." **[Place candy on red circle.]** *{Note: in Familiarization #2 the experimenter says 'I don't put anything in the red circle next to [name of Recipient]'}*

{Note: in the Asocial condition the experimenter replaces 'next to [name of Recipient]' with 'on the other side'}

[Point to the Left tray:]

"If you choose this paper *[tray]*, you get whatever is in the blue circle and [name of Recipient] gets whatever is in the red circle." *{Note: in the Asocial condition the experimenter replaces '[name of Recipient]' with 'nobody'}*

[Point to the Right tray:]

"If you choose this paper *[tray]*, you get whatever is in the blue circle and [name of Recipient] gets whatever is in the red circle." *{Note: in the Asocial condition the experimenter replaces '[name of Recipient]' with 'nobody'}*

[Now sit back and look at the subject, not at the trays.]

"NOBODY gets what is on the other paper *[tray]*. The candy goes back into the bag. You can choose whichever paper *[tray]* you want, but you can only choose one paper *[tray]*. Now, you can make your choice. Go ahead."

"Okay, now let's do some more."

For PG and CSG (both Social and Asocial conditions)

[Left tray: Place candies on blue circle. Then, place candies on red circle.]

[Right tray: Place candy on blue circle. Then, place candy on red circle.]

"Now, you can make your choice. Go ahead."

[**Remove remaining candies and put them back in the bag. If the subjects do not take their candies, encourage them to do so.]**

Script for adult subjects

"This is a game that we usually play with children, but we're also curious about how adults will play it. Would you be willing to participate?"

"On each of these papers *[trays]*, there are two circles, a blue circle and a red circle. **[Point to blue circle and red circle.]**

For Familiarization #1 and #2 (both Social and Asocial conditions)

[Left tray: Place candies on blue circle. Then, place candies on red circle.]
[Right tray: Place candy on blue circle. Then, place candy on red circle.]

[Point to the Left tray:]

"If you choose this paper *[tray]*, you get whatever is in the blue circle and [name of Recipient] gets whatever is in the red circle." *{Note: in the Asocial condition the experimenter replaces '[name of Recipient]' with 'nobody'}*

[**Point to the Right tray:]**

"If you choose this paper *[tray]*, you get whatever is in the blue circle and [name of Recipient] gets whatever is in the red circle." *{Note: in the Asocial condition the experimenter replaces '[name of Recipient]' with 'nobody'}*

[Now sit back and look at the subject, not at the tray.]

"Remember, NOBODY gets what is on the paper *[tray]* that you don't choose. That candy goes back into the bag."

"Ok, go ahead and make your choice."

[**Remove remaining candies and put them back in the bag. If the subjects do not take their candies, encourage them to do so.]**

"Okay, now let's do some more."

For PG and CSG (both Social and Asocial conditions)

[Left tray: Place candies on blue circle. Then, place candies on red circle.]
[Right tray: Place candy on blue circle. Then, place candy on red circle.]

"Now, you can make your choice. Go ahead."

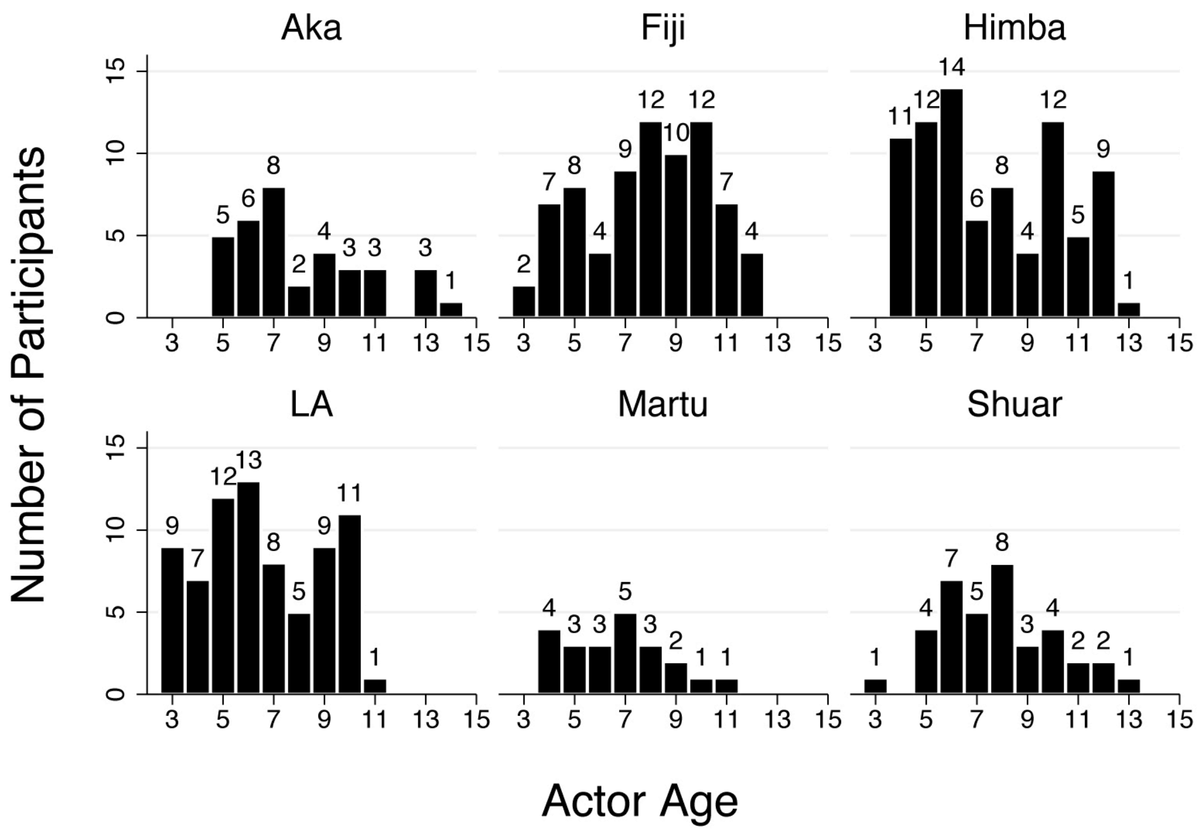
[**Remove remaining candies and put them back in the bag. If the subjects do not take their candies, encourage them to do so.**

SI Table S1(b): Methodological variations across sites

Group	Child recruitment	Child pairs familiar?	Child Rewards	Reward rare on site?	Adult recruitment	Adult pairs familiar?	Adult rewards	Reward rare on site?
Aka	On an opportunity basis in local community (forest camps).	Yes. Subjects drawn from same group of families living together. Children knew each other since infancy.	Hard candy	Yes	On an opportunity basis in local community (forest camps).	Yes, all drawn from same camp.	Hard candy	Yes
American (West Los Angeles)	At preschools and primary schools.	Yes. Subjects drawn from same school classes.	Cheese crackers	No	Undergraduate sections of Anthropology courses at UCLA.	Yes. Subjects drawn from same school classes.	Hershey's kisses	No
Fijian (Yasawa)	At schools and on an opportunity basis in three small villages.	Yes. All subjects lived in same villages; children of school age were paired with classmates.	Hard candy	Fairly rare	Residents of same villages	Yes	Hard candy	Fairly rare
Himba	At schools and on opportunity basis in extended family compounds.	Yes.	Cheese snack	Yes	Adults were drawn from the same community as child subjects.	Yes; members of the same community	Cheese snack	Yes
Martu	At community school.	Yes, drawn from same school.	Crackers	No	n/a	n/a	n/a	n/a
Shuar	At village school.	Yes, all from same village.	Animal cracker cookies	No	Residents of same village	Yes, all from same village. Some were kin.	Animal cracker cookies	No

SI Appendix 2: Additional demographic data

Figure S2: Age distributions for the different populations in our sample (N=326).



SI Table S2(a): Additional demographic information about field sites.

Group	Nation/Region	Language Family	Environment	Economic Base	Residence	Researcher
Aka	Central African Republic	DiAka (Bantu C10)	Tropical forest	Foraging	Nomadic	Boyette / Hewlett
American	United States / Los Angeles	English	Urban	Wage work	Sedentary	House
Fijian	Fiji / Yasawa Island	Oceanic	Coastal tropical island	Horticulture / marine foraging	Sedentary	Silk / Henrich
Himba	Namibia (Kaokoland)	Herero	Semi-arid mountain savanna	Agro-pastoralism	Semi-nomadic	Silk / Scelza
Martu	Western Australia	Martu Wangka	Desert	Foraging/Wage Work	Semi-nomadic	Scelza
Shuar	Ecuador/Amazonia	Jivaroan	Tropical forest	Horticulture, some wage work	Sedentary	Barrett

SI Table S2(b): Details about school attendance and change in how children interact with peers across development.

Group	Age children first attend school	Children's interactions with peers (early childhood; <5 years)	Children's interactions with peers (middle childhood; 6 years to puberty)	Children's interactions with peers (adolescence; >puberty)
Aka	Most children do not go to school. Those that do start as early as 5-years-old. None were attending school when the study was conducted.	Children stay in camp, but interact with anyone who is around. Mostly they interact with peers <5 years at this age, as older children will leave camp during the day to play and work.	After age 5, children spend on average twice as much time away from camp in the forest foraging and playing. At least half of the time spent in the forest they are out of sight of adults, with mixed-age, mixed sex groups of children (ages 7-18 years).	Aka adolescents spend a greater percentage of their time with other adolescents than younger children spend with only same-age peers, but they are also usually with adults or children younger than age 5, either working or visiting.
American	2-3 for preschool; 5 kindergarten; 6 primary school.	Children are primarily in the home, interacting with siblings or neighbors.	Children also interact with many unfamiliar peers in school settings.	Children interacting with many peers in school, and also increasingly in non-school activities.
Fijian	5: kindergarten; 6 primary school.	Children primarily in the home, interacting with siblings or neighbors in the village.	Children also interact with peers in school, which includes many of the same children from their village, in addition to children from neighboring villages whom they are likely already familiar with.	Children interacting with many peers in school. At age 14/15 many children leave the village to continue schooling on a larger island, and there will interact with unfamiliar peers.
Himba	5, but very few attend school.	Children primarily in the home, interacting with siblings or children from neighboring compounds.	Children are primarily in the home, interacting with siblings, cousins and neighboring children. Expected to perform household labor from about age 5, including collecting water, grinding maize, herding goats, and watching young siblings.	Children interact with those in their compound and neighboring compounds and may travel independently. Girls may go to live with a spouse after menarche (although marriage can occur much earlier, change in residence usually doesn't happen until menarche). First birth is in mid to late teens.
Shuar	5 years	Children visit households all over the village.	Lots of interaction across households within village; visits to other villages (e.g. for sports) starts in middle childhood.	Children interact with peers both within and outside the village.
Martu	4 years	Children primarily in the home, interacting with siblings, cousins and neighboring children.	Children at school or at home or neighboring home within the community.	Children interact with those in their household and community and often travel to other communities to see relatives or friends.

SI Appendix 3: Effect of condition and population for adults.

If Actors were prosocial then they should have chosen 1/1 more in the Social condition than in the Asocial condition. Here we explore whether adults' 1/1 choices in the Prosocial Game (PG) and the Costly Sharing Game (CSG) varied across condition and population. In SI Table S3(a) we use a model-selection procedure similar to that used with children (see Table 2 Main text, SI Appendix 4), and in SI Table S3(b) we present the best-fit models for the CSG and PG.

SI Table S3(a): Multilevel logistic regression models used in model selection, comparable to model selection procedure used with child data. Below we provide details about Fixed Effect (FE) and Random Effect (RE) parameters included in each model, the hypothesis reflected in each model, and the DIC and DIC weight values associated with each model when it is applied to the CSG, and PG. Best-fit models are in bold.

Model	Parameters included	Hypothesis	CSG	PG
			DIC (weight)	DIC (weight)
A	**FE:** none **RE:** Actor ID	No differences across conditions or populations	335.42 (.00)	332.6 (.00)
B	**FE:** Condition **RE:** Actor ID	Differences across conditions	323.6 (.00)	**311.87 (.70)**
C	**FE:** none **RE:** Actor ID Population	Differences across populations	327.93 (.00)	334.72 (.00)
D	**FE:** Condition **RE:** Actor ID Population | Condition	Differences across conditions and populations	**302.41 (1.00)**	313.53 (.30)

As with children (Table 2 in main text, SI Appendix 4), the best-fit model for adults was Model D for the CSG, and Model B for the PG (SI Table S3[a]). Model D also had a .3 probability of being the best model for the PG, and the DIC values for Models B and D are similar, suggesting that there is a greater likelihood of differences between populations of adults than between populations of children, but it is still substantially more likely that there are population differences in the CSG than in the PG.

SI Table S3(b): Best-fit models for the CSG and PG for adults.

DV: Chose Prosocial Outcome (1/1) N=120	Model 1	Model 2
	CSG	PG
	Coef. (95% CI)	Coef. (95% CI)
Intercept	-2.11 (-4.07 : -0.56)	-1.69 (-2.59 : -0.82)
Condition	1.32 (0.20 : 2.76)	1.39 (0.83 : 1.99)
Variance parameters	**SD (95% CI)**	**SD (95% CI)**
Actor ID | Intercept	0.83 (0.47 : 1.29)	0.52 (0.41 : 1.02)
Population | Intercept	2.15 (0.43 : 3.05)	
Population | Condition	1.53 (0.45 : 2.54)	

The overall pattern is that, despite a greater likelihood of between-population differences in the CSG than in the PG, the estimate for the Condition parameter is greater than zero for both. This indicates that adults were more likely to select the 1/1 outcome in the Social condition than in the Asocial condition, for both the CSG and the PG.

SI Appendix 4: Best-fit models for PG, CSG, and FAM1 for children.

SI Table S4(a): Table 3 from the main text. N=326 for each task.

Model	Parameters included	Hypothesis	CSG	PG	FAM1 Social
			DIC (weight)	DIC (weight)	DIC (weight)
A	**FE:** CA, CA^2 **RE:** Actor ID	One developmental trajectory	765.12 (.00)	904.61 (.06)	321.53 (.05)
B	**FE:** CA, CA^2 Condition, CA*Condition, CA^2 *Condition **RE:** Actor ID	One trajectory for each condition	746.86 (.00)	**899.10 (.94)**	
C	**FE:** CA, CA^2 **RE:** Actor ID Population | CA, CA^2	One trajectory for each population	747.30 (.00)	914.97 (.00)	**315.58 (.95)**
D	**FE:** CA, CA^2 Condition, CA*Condition, CA^2 *Condition **RE:** Actor ID Population | CA, CA^2, Condition, CA*Condition, CA^2 *Condition	One trajectory for each condition, for each population	**734.49 (1.00)**	916.52 (.00)	

SI Table S4(b): Additional model selection routines, comparable to those from Table S4(a). Here we run the same procedures, but either (a) include Actor Sex as a parameter in the model (which removes three subjects whose sex was not recorded), (b) limit the dataset to those Actors 5-10 years of age, or (c) limit the dataset to those Actors who selected the 2/2 (income-maximizing) outcome in FAM1 Social. The best-fit models are the same as in Table S4(a).

		(a) Including Actor Sex as a Fixed Effect parameter (N=326)		(b) Limiting analysis to Actors who are 5-10 years of age (N=235)		(c) Limiting analysis to Actors who selected 2/2 in FAM1 Social (N=259)	
Model	Hypothesis	CSG	PG	CSG	PG	CSG	PG
		DIC (weight)	DIC (weight)	DIC (weight)	DIC (weight)	DIC (weight)	DIC (weight)
A	One developmental trajectory	761.59 (.00)	896.16 (.04)	530.17 (.00)	650.31 (.02)	560.32 (.00)	714.92 (.00)
B	One trajectory for each condition	743.87 (.00)	**889.99 (.96)**	517.66 (.02)	**642.12 (.98)**	555.07 (.00)	**696.63 (1.00)**
C	One trajectory for each population	745.97 (.00)	904.89 (.00)	521.48 (.00)	653.08 (.00)	546.72 (.01)	722.06 (.00)
D	One trajectory for each condition, for each population	**731.94 (1.00)**	907.71 (.00)	**509.54 (.98)**	653.45 (.00)	**537.34 (.99)**	708.98 (.00)

SI Table S4(c): Fully specified best-fit models from Table S4(a). The model for FAM1 focuses only on the Social condition, so does not include parameters for Condition or for Actor ID (as each Actor only contributed one observation to this dataset).

DV: Chose 1/1 N = 326	Model 3 CSG Coef. (95% CI)	Model 4 PG Coef. (95% CI)	Model 5 FAM1 Social Coef. (95% CI)
Intercept	-1.59 (-2.69 : -0.49)	-0.63 (-1.31 : 0.05)	1.53 (0.88 : 2.23)
CA	-0.71 (-1.63 : 0.15)	-0.50 (-1.04 : 0.02)	0.63 (-0.07 : 1.39)
CA^2	-0.68 (-1.57 : 0.19)	-0.24 (0.25 : 0.70)	.00 (-0.57 : 0.59)
Condition	0.14 (-0.59 : 0.89)	0.52 (0.09 : 0.95)	
CA * Condition	0.16 (-0.48 : 0.77)	0.40 (0.08 : 0.76)	
CA^2 * Condition	0.67 (0.05 : 1.29)	-0.12 (-0.41 : 0.20)	
Variance parameters	**SD (95% CI)**	**SD (95% CI)**	**SD (95% CI)**
Actor ID | Intercept	1.02 (0.70 : 1.25)	0.52 (0.28 : 0.81)	
Population | Intercept	0.66 (0.32 : 1.24)		.67 (0.33 : 1.24)
Population | CA	0.65 (0.32 : 1.17)		0.74 (0.35 : 1.34)
Population | CA^2	0.66 (0.32 : 1.21)		0.57 (0.29 : 1.01)
Population | Condition	0.60 (0.31 : 1.05)		
Population | CA * Condition	0.57 (0.30 : 1.01)		
Population | CA^2 * Condition	0.56 (0.30 : 0.98)		

SI Table S4(d): Estimated parameters for the PG, extracted from the best-fit model for this task (Model B, Table S4[a]). These are the estimates plotted in Figure 1 in the main text.

DV: Chose 1/1 N = 326	Model 6 PG Social Coef. (95% CI)	Model 7 PG Asocial Coef. (95% CI)
Intercept	0.41 (0.08 : 0.73)	-0.11 (-0.41 : 0.19)
CA	0.31 (0.08 : 0.55)	-0.10 (-0.33 : 0.14)
CA^2	0.01 (-0.21 : 0.25)	0.13 (-0.09 : 0.34)

Model 7 indicates that for the PG Asocial none of the parameters are clearly different from zero, suggesting that there was little change as a function of age in the probability that children chose 1/1 in this task. In contrast, Model 6 shows us that in the PG Social the coefficient for CA was substantially greater than zero, suggesting that children became more likely to chose 1/1 with increasing age.

SI Table S4(e): Estimated parameters for each of our six populations in the CSG Social condition, extracted from the best-fit model for the CSG (Model D, Table S4[a]). These are the estimates plotted in Figure 2(a) in the main text, and Figures S4(a-f).

DV: Chose 1/1 N = 326	**Model 8**	**Model 9**	**Model 10**	**Model 11**	**Model 12**	**Model 13**
	American	**Fijian**	**Himba**	**Shuar**	**Aka**	**Martu**
	Coef. (95% CI)	**Coef. (95% CI)**	**Coef. (95% CI)**	**Coef. (95% CI)**	**Coef. (95% CI)**	**Coef. (95% CI)**
Intercept	-1.87 (-2.81 : -1.02)	-0.80 (-1.63 : 0.00)	-2.11 (-3.07 : -1.24)	-1.82 (-2.91 : -0.85)	-0.81 (-1.79 : 0.12)	-0.61 (-1.75 : 0.59)
CA	0.96 (0.24 : 1.76)	-0.59 (-1.25 : 0.00)	-0.66 (-1.25 : -0.10)	-0.86 (-1.81 : 0.09)	-0.61 (-1.67 : 0.38)	-0.51 (-1.78 : 0.67)
CA^2	1.53 (0.76 : 2.35)	-0.17 (-0.85 : 0.46)	0.78 (0.24 : 1.38)	0.80 (0.07 : 1.59)	0.74 (0.07 : 1.50)	0.31 (-0.90 : 1.47)

All Models report a substantial positive estimate for CA^2, with the exceptions of Models 9 and 13, where the estimate for CA^2 is near zero. A near-zero estimate for CA^2 suggests a monotonic age trajectory, while substantial positive estimates for CA^2 point to a quadratic shape with a positive bend (a u-shape). The estimates for the intercept for all of the populations are substantially negative, indicating a low probability of 1/1 choices at this point on the centered age vector (i.e. 7.42 years of age), across all populations. This suggests qualitatively different age trajectories across our different populations, with some showing increasing rates of 1/1 choices as a function of age, while others show little change or decreases in the rates of 1/1 choices with age. These results are clearly discernible from Figure 2(a) in the main text (also Figure S4[a-f]).

SI Table S4(f): Estimated parameters for each of our six populations in the CSG Asocial condition, extracted from the best-fit model for the CSG (Model D, Table S4[a]). These are the estimates plotted in Figure 2(b) in the main text, and Figures S4(a-f).

DV: Chose 1/1 N = 326	Model 14	Model 15	Model 16	Model 17	Model 18	Model 19
	American	Fijian	Himba	Shuar	Aka	Martu
	Coef. (95% CI)	Coef. (95% CI)	Coef. (95% CI)	Coef. (95% CI)	Coef. (95% CI)	Coef. (95% CI)
Intercept	-1.89 (-2.69 : -1.16)	-1.18 (-1.89 : -0.51)	-1.88 (-2.68 : -1.16)	-1.77 (-2.66 : -0.99)	-1.02 (-1.83 : -0.27)	-1.03 (-1.90 : -0.14)
CA	0.34 (-0.24 : 0.93)	-0.64 (-1.26 : -0.06)	-0.69 (-1.22 : -0.19)	-0.87 (-1.61 : -0.16)	-0.70 (-1.50 : 0.04)	-0.66 (-1.61 : 0.20)
CA^2	0.41 (-0.19 : 1.01)	-0.64 (-1.33 : -0.04)	0.16 (-0.34 : 0.66)	0.13 (-0.51 : 0.78)	0.17 (-0.40 : 0.73)	-0.26 (-1.17 : 0.56)

Most of these models include estimates for CA^2 that are close to zero, with the exception of the negative estimate in the model of our Fijian sample (Model 15), which suggests an inverted u-shape. Half of the models' estimates for CA are close to zero (Model 14, 18, and 20), while the others are negative. Additionally, all of the models include a large negative estimate for the intercept. Overall, these results suggest that in the CSG Asocial there is some variation in how children's behavior develops across our populations, but across groups children are consistently unlikely to select the 1/1 outcome, and they either stay consistently unlikely to do so or become less likely to select 1/1 with increasing age.

SI Table S4(g): Estimated parameters for each of our six populations for FAM1 Social condition, extracted from the best-fit model for this dataset (Model C, Table S4[a]). These are the estimates plotted in Figure 3 in the main text, and Figures S5(a-f).

DV: Chose 1/1 N = 326	Model 20	Model 22	Model 23	Model 24	Model 25	Model 26
	American	Fijian	Himba	Shuar	Aka	Martu
	Coef. (95% CI)	Coef. (95% CI)	Coef. (95% CI)	Coef. (95% CI)	Coef. (95% CI)	Coef. (95% CI)
Intercept	2.08 (1.27 : 3.04)	1.24 (0.52 : 2.00)	1.07 (0.41 : 1.76)	2.03 (1.04 : 3.35)	0.74 (-0.14 : 1.59)	1.80 (0.77 : 3.06)
CA	0.83 (0.06 : 1.83)	1.01 (0.29 : 1.96)	0.28 (-0.19 : 0.77)	1.32 (0.18 : 3.13)	-0.10 (-1.07 : 0.77)	0.65 (-0.43 : 1.98)
CA^2	-0.10 (-0.82 : 0.62)	0.59 (-0.10 : 1.48)	0.04 (-0.43 : 0.54)	0.40 (-0.53 : 1.61)	0.34 (-0.28 : 1.08)	0.24 (-0.74 : 1.28)

All models report a substantial positive estimate for Centered Age (CA), with the exception of Models 23 and 25 in which the estimate for CA is close to zero. Also, all models report an estimate for CA^2 that is close to zero. Additionally, the estimates for the intercepts are positive, and all are substantially above zero with the exception of the model for the Aka sample (Model 25). This suggests a high probability of children selecting 2/2 at the intercept value of CA (i.e. 7.42 years of age), across at least most of our populations. These results suggest some variability in the age function across these populations, but all of these functions are shifted towards high probabilities of choosing 2/2, and most involve either an increasing slope to the age function (i.e. increasing change in prosociality as a function of age) or a relatively unchanging slope.

SI Appendix 5: Addressing concerns about sample size

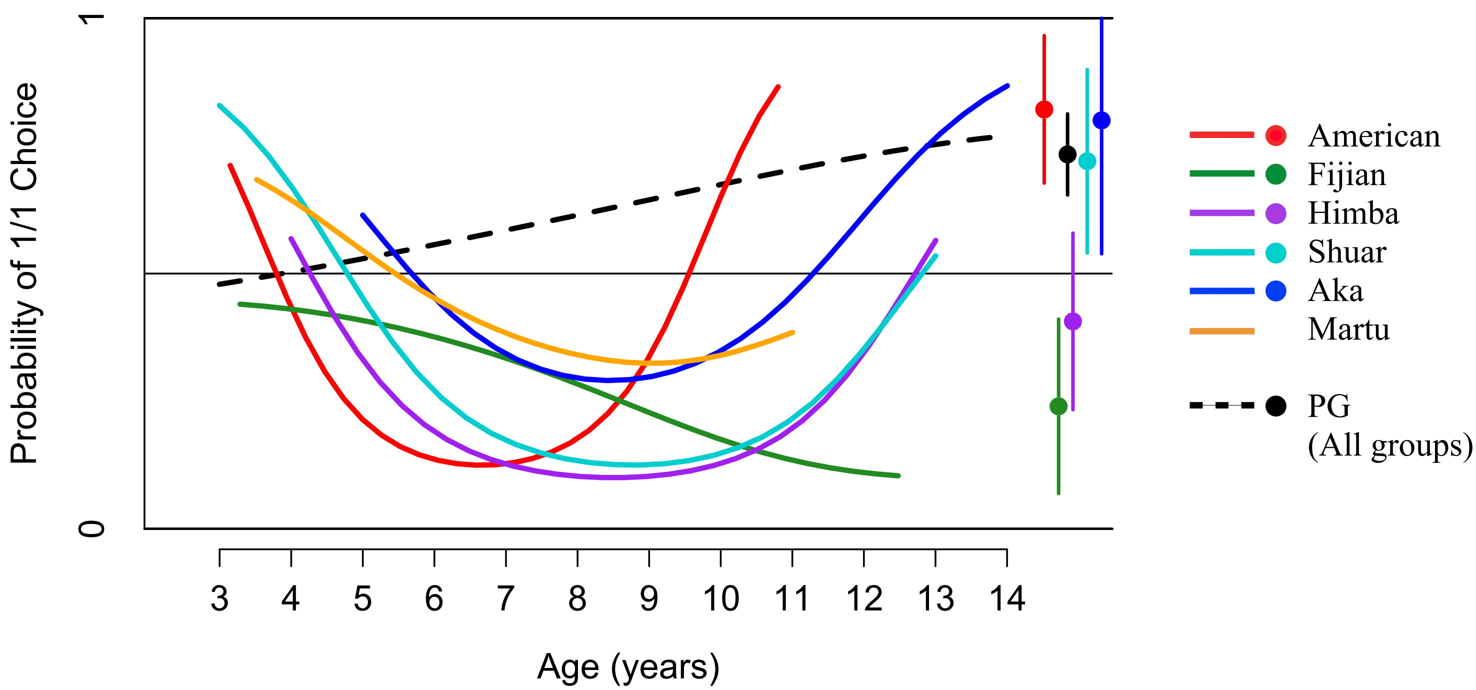
Our population samples are varied both in their total size and in their distribution across the ages of 3-14 years. This is an unavoidable feature of naturalistic samples drawn from small-scale societies, but in a number of ways our analyses are well-equipped to deal with small samples.

By utilizing age as a continuous variable we avoid much of the noise created when aggregating observations into broad age bins. For example, grouping children into 3 year-olds and 4 year-olds is problematic because a 4.1 year-old is likely to be more similar to a 3.9 year-old than to a 4.9 year-old. Using regressions also allows us to make informative predictions about the behavior of children even at the peripheries of our age distributions, where our samples are small or even missing, because the behavior of a 3 year-old is not completely independent from the behavior of a 4 year-old. Furthermore, our model structures reduce concerns about small sample sizes, by combining data across populations and estimating our age functions using interactions and random effects. By using random effects the model is considering each population as a sample drawn from a larger distribution, and observations from one population still inform the estimation of the age parameters for the other populations. This reduces concerns about differences in sample size across populations because all the observations in the sample are used when estimating parameters for each population.

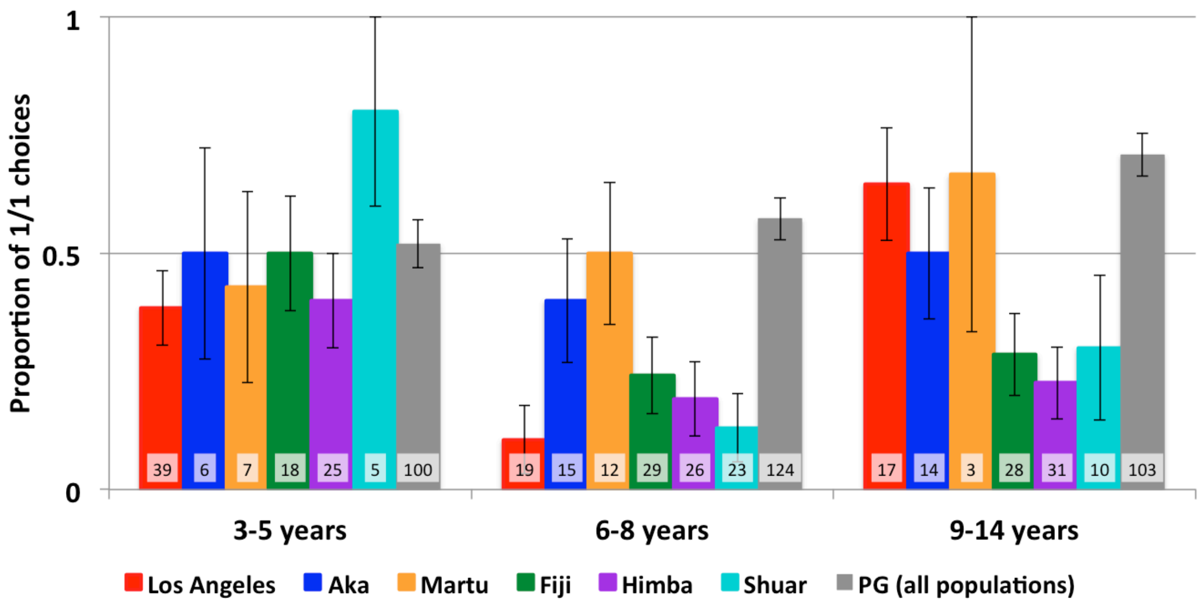
Truncated samples: We can further reduce concerns about the effect of small samples along the periphery of our age distribution by truncating our sample to all participants aged 5-10 years. This covers the densest parts of our age distributions for each of our samples, and all of our populations include participants with these ages. In Tables S4(a) and S4(b) we compare the results of our primary model selection comparison when applied to all of our subjects, and to the subset of our sample aged 5-10 years. The results are qualitatively the same.

Age bins: We can also reduce concerns about sample size by aggregating subjects into broad age categories with approximately comparable numbers of participants. Here we use three age bins: 3-5 years, 6-8 years, and 9-14 years. This approach is noisy and imprecise, but by presenting our data in a more traditional format with means and standard errors we hope to alleviate concerns that our models produce spurious results. In Figure S3(a), we re-present a version of Figure 2[a] from the main text, and in Figure S3(b) we present another version of this figure that uses means and standard errors to plot the same data (children's responses in the CSG Social and PG Social).

SI Figure S3(a): Similar to Figure 2(a) from main text. Best-fit models used to plot age functions for CSG Social (colored lines) and PG Social (black dotted line). Dots on the right-hand side of the graph represent the proportions of adults who selected 1/1 (colored dots for the CSG, solid black dot for the PG), and lines above and below dots correspond to 95% confidence intervals. The vertical axis corresponds to the estimated probability that children will choose the prosocial (1/1) outcome. Horizontal axis corresponds to children's age (in years).



SI Figure S3(b): Proportions of children who select 1/1 outcome, for three age categories that have approximately equal numbers of participants: 3-5 years, 6-8 years, and 9-14 years. Colored bars represent means for each population in the CSG Social, grey bars plot means for all populations together in the PG Social. Error bars reflect the standard error of the mean, and the values at the base of each bar reflect the number of participants included in that bar.

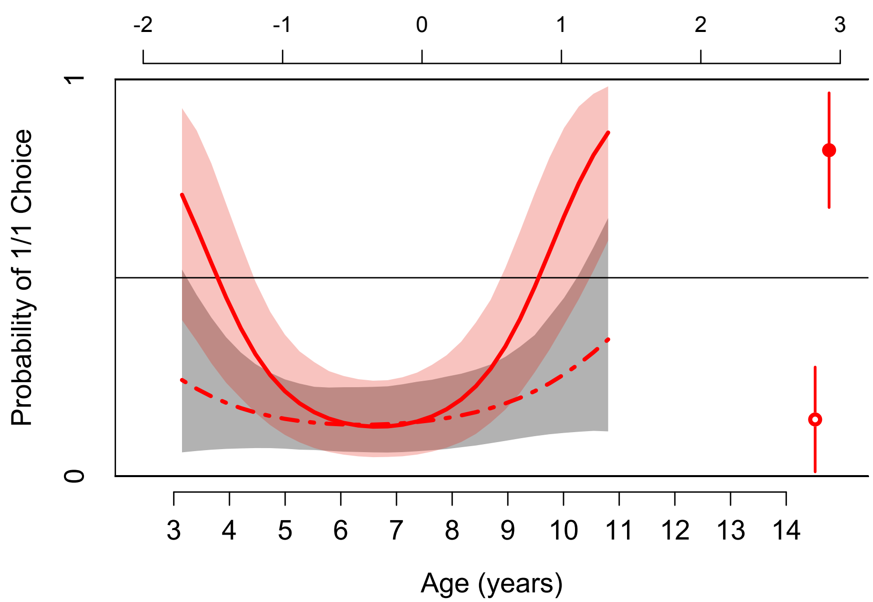


We can see similar overall developmental patterns in SI Figure S3(a) and SI Figure S3(b). Both plots indicate relative similarity across populations in early childhood (3-5 years). This is followed by lower overall probabilities/proportions of 1/1 choices and emerging population variation in middle childhood (6-8 years). Finally, in late childhood/early adolescence (9-14 years) we can see even greater population variation, along with higher probabilities/proportions of 1/1 choices in some groups (relative to 6-8 year-olds).
In general, our primary results concerning the origins of population variation hold across different methods for representing our data. This shows that our results are not an artifact of our methods. Furthermore, given the many benefits of using multilevel regressions, this strongly argues for their use when exploring development across different populations.

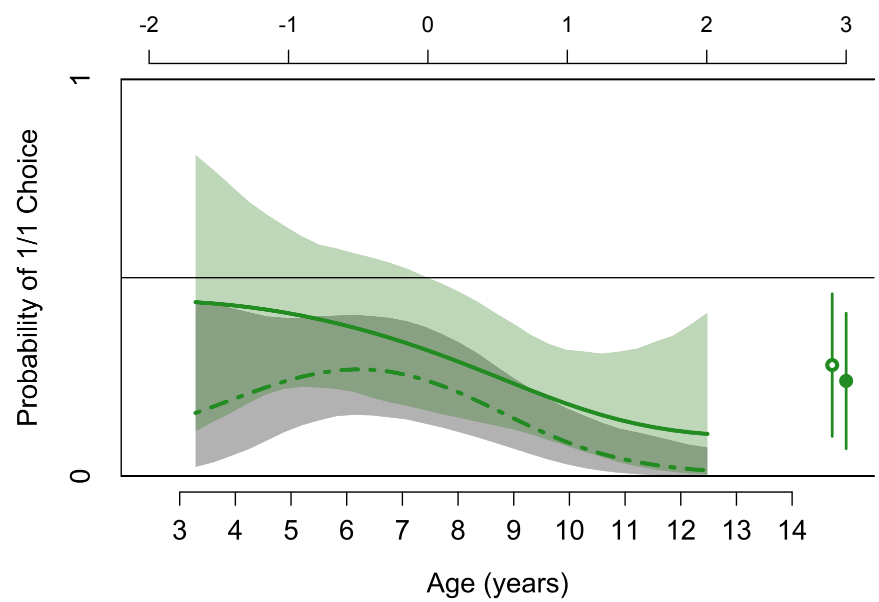
SI Appendix 6: 95% Confidence intervals for Figure 2 from main text.

SI Figures S4(a-f): See Tables S4(e) and S4(f) for associated models. The vertical axis on each plot corresponds to the estimated probability that children will choose the prosocial (1/1) outcome. The bottom horizontal axis corresponds to children's age (in years), and the top horizontal axis corresponds to equivalent values of CA. The solid line and the colored shaded area in each plot correspond to the estimated probability of 1/1 choices and the 95% confidence interval for the Social condition. The filled dot on the right side of the plot corresponds to adults' choices in the Social condition, and the hollow dot corresponds to adults' choices in the Asocial condition. The dotted line and the grey shaded area in each plot correspond to the estimated probability and 95% confidence interval for the Asocial condition. The darker shaded area is where the confidence intervals overlap.

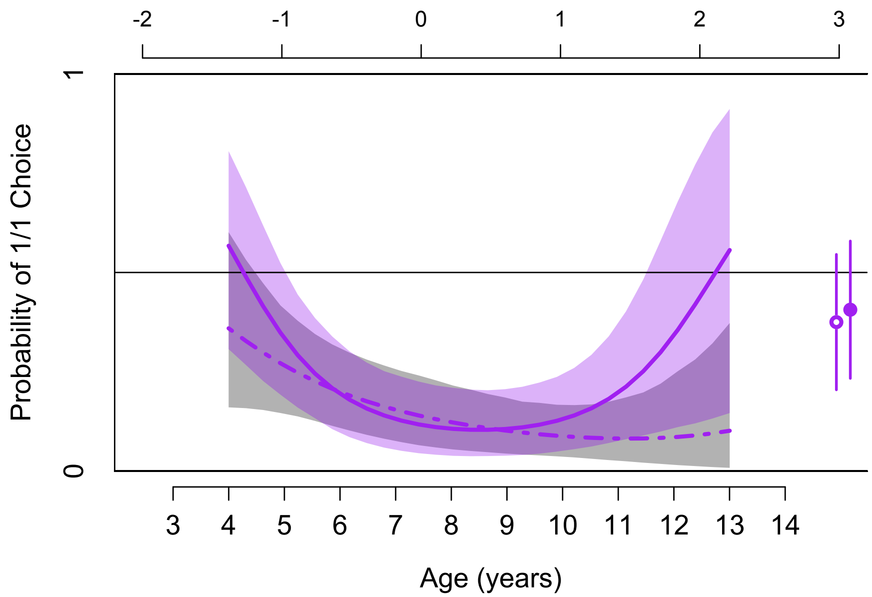
SI Figure S4(a): Los Angeles



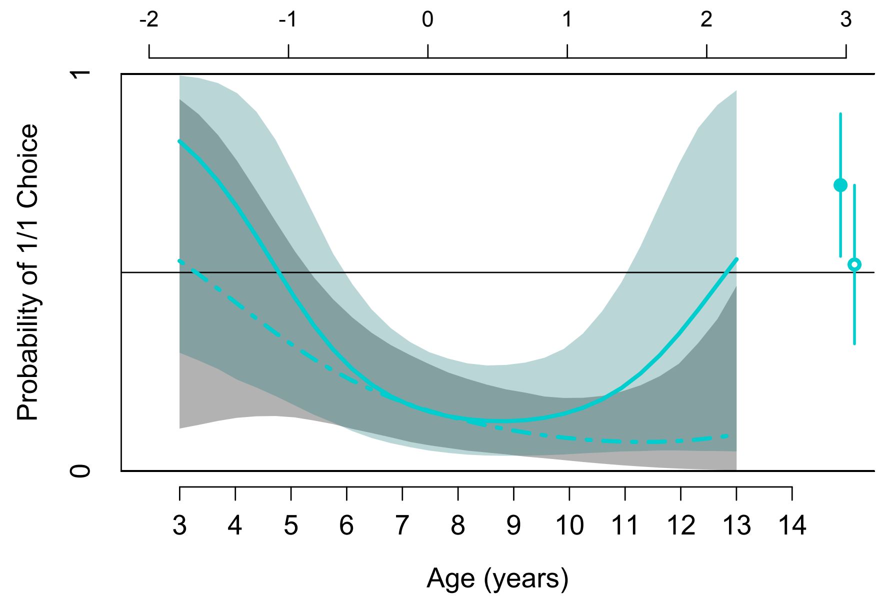
SI Figure S4(b): Fiji



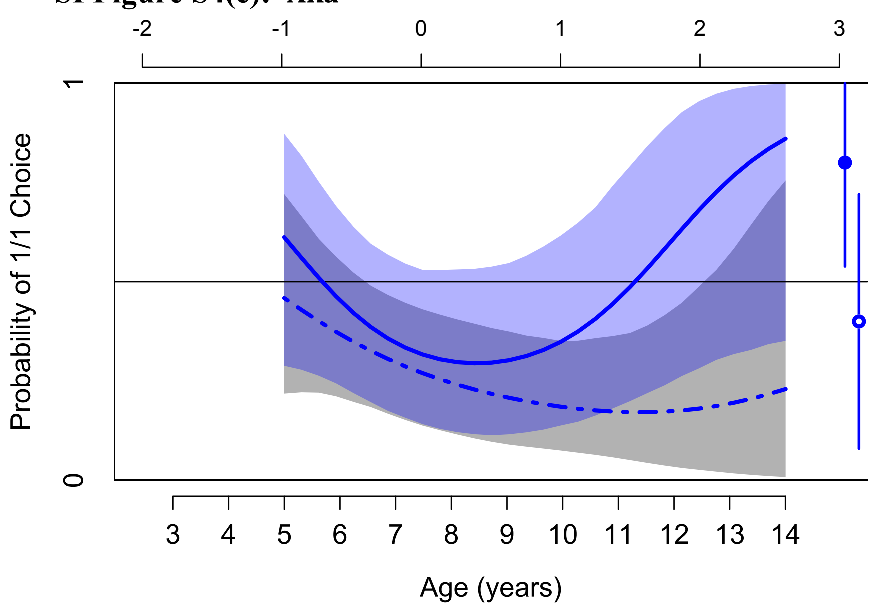
SI Figure S4(c): Himba



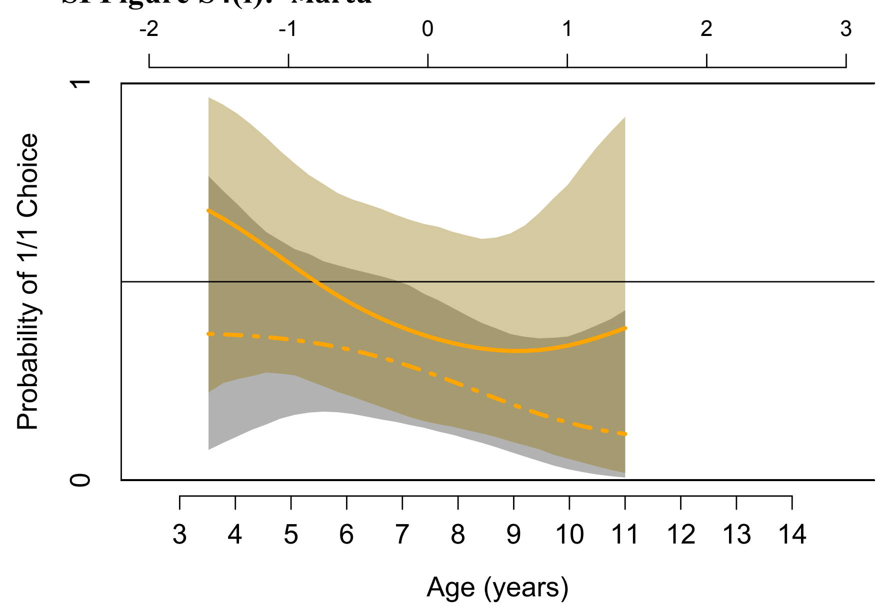
SI Figure S4(d): Shuar



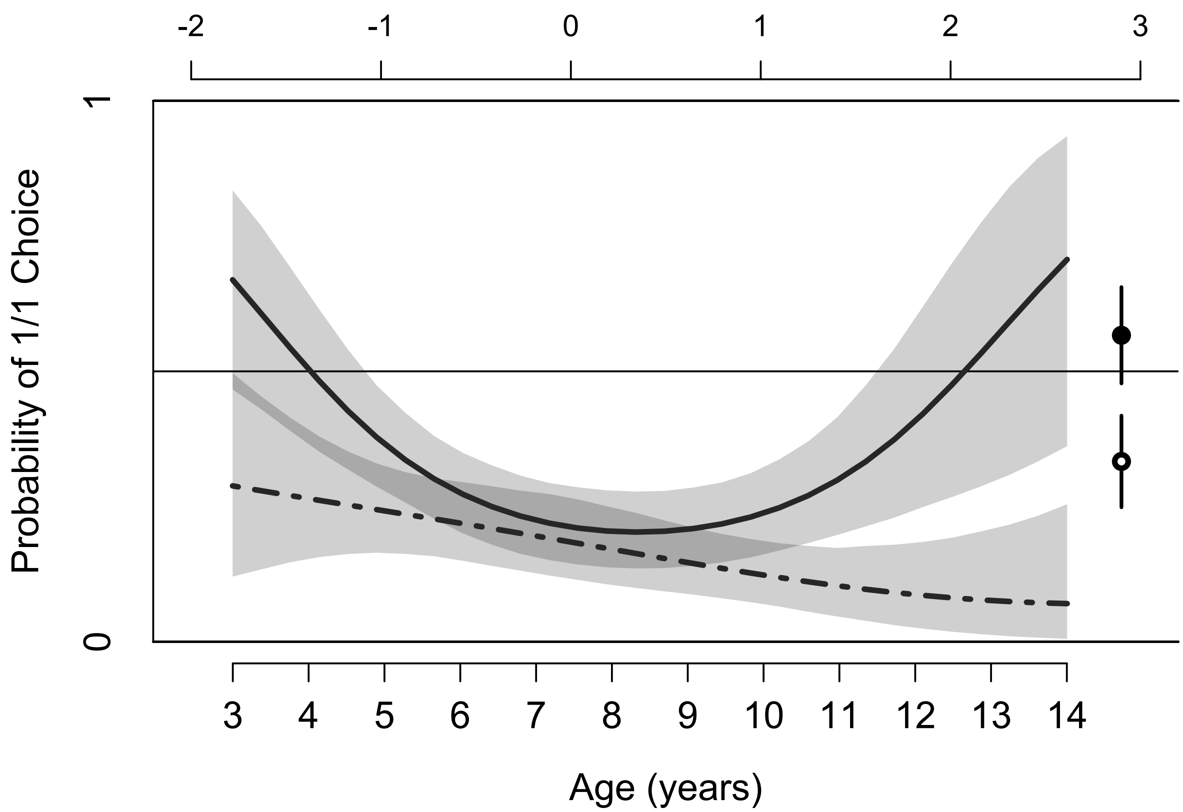
SI Figure S4(e): Aka



SI Figure S4(f): Martu



SI Figure S4(g): Age functions for the CSG, collapsing across all populations. Vertical axis is the estimated probability that children will choose the prosocial (1/1) outcome. Bottom horizontal axis is children's age (in years), and top horizontal axis is the equivalent value of CA. Age functions capture the estimated probability that children will select the 1/1 outcome as a function of age, with estimates extracted from Model 2 for the CSG (Table 4 main text, Table S4[a] above) for both the Social condition (solid line) and the Asocial condition (dotted line). Dots on the right side of the graph reflect the proportion of 1/1 choices actually made by adults.



SI Appendix 7: Analysis of Familiarization Trials.

The choices that participants made during familiarization trials provide some insight about their comprehension of the experimental task. However, some of these trials allow multiple interpretations that could lead Actors to select either outcome, even if they correctly understand the task. In particular, for three of the four trials a participant might choose one outcome if they wish to maximize personal payoff, but they might choose the other outcome if they wish to minimize inequity in payoff between themselves and the recipient. Only one trial avoids this confound between payoff maximization and inequity aversion (Familiarization #1 Social).

We included four familiarization trials across the Social and Asocial conditions:

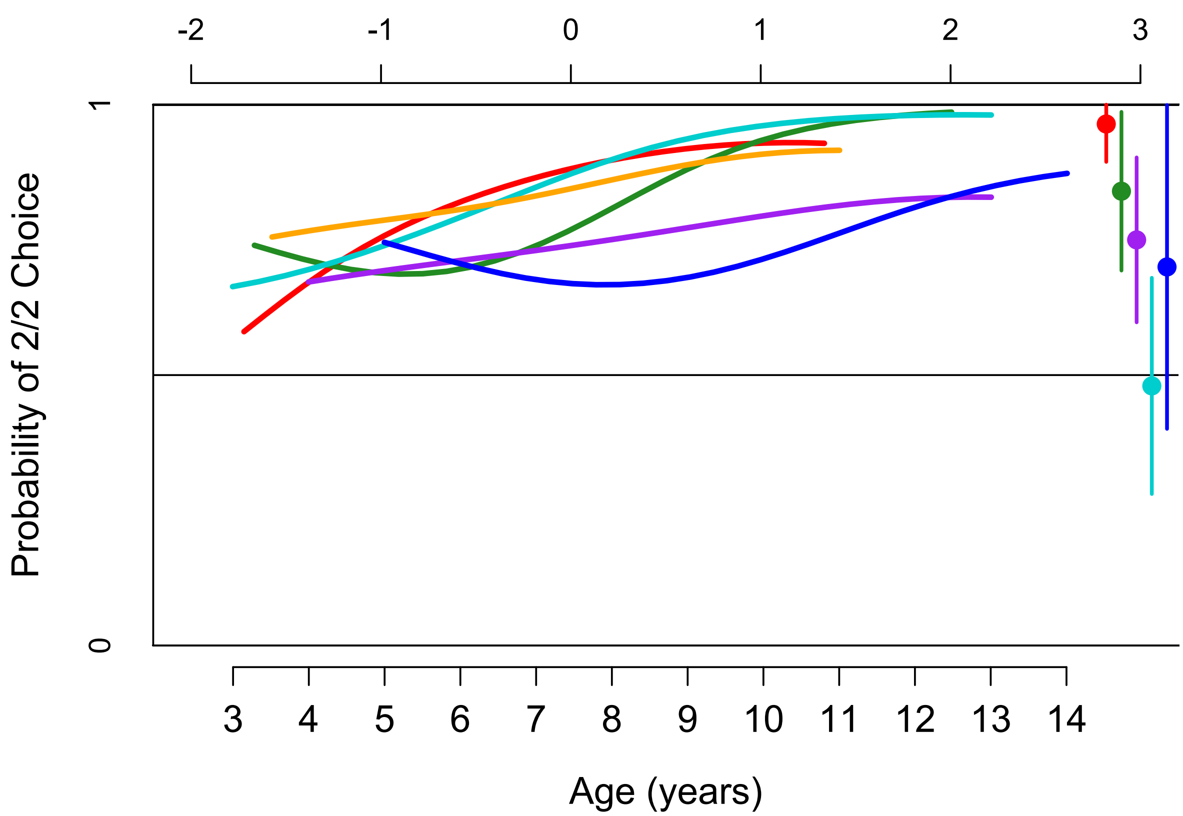
SI Table S5: Details about Familiarization trials.

Trial	Payoffs	Rationale for choice
Familiarization #1 (Social Condition)	1/1 vs. 2/2	2/2 maximizes payoff to self *without increasing inequity*
Familiarization #2 (Social Condition)	1/0 vs. 2/0	2/0 maximizes payoff to self, 1/0 minimizes inequity
Familiarization #1 (Asocial Condition)	1/1 vs. 2/2	2/2 maximizes payoff to self, 1/1 minimizes inequity
Familiarization #2 (Asocial Condition)	1/0 vs. 2/0	2/0 maximizes payoff to self, 1/0 minimizes inequity

Only the first trial, Familiarization #1 Social, affords an income-maximizing trial that doesn't increase inequity. All three other trials result in the Actor receiving a higher payoff than the Recipient. If the Actor and Recipient are both present (Social condition), then choosing 2/0 over 1/0 increases inequity between the Actor and the Recipient, because at the end of the trial the Actor's payoff has been increased by two while the Recipient's payoff has not increased. If the Actor had selected 1/0 over 2/0 then there would be less inequity, as the Actor's payoff would only increase by one. This also applies if the Actor chooses 2/0 over 1/0 when no recipient is present (Asocial condition), because the Actor's payoff increases by two while the Recipient's payoff does not increase (the physical presence of the Recipient being irrelevant in this case). This argument also applies to the 2/2 vs. 1/1 choice in the Asocial condition. To avoid this potential confound between self-interest and inequity aversion, we focus only on Familiarization #1 (FAM1) Social.

In Figure S5(a) (Figure 3 in main text) we plot the age functions for this trial, with the outcome measure reflecting the probability that participants will select 2/2 (coded as '1') over 1/1 (coded as '0'). It is clear from Figure S5(a) that children generally prefer 2/2 over 1/1, and that this tendency generally increases with age across the six populations. Figures S7(b-g) plot these functions separately for each population, including estimated 95% confidence intervals for the functions. These confidence intervals reveal that for all of the populations, by about age 5 children show evidence of selecting 2/2 with a probability above 0.5. This indicates that they show a systematic bias towards the self-maximizing outcome, and provides strong evidence that they comprehend the task. The age function for the Aka (one of our smaller samples) only briefly deviates above 0.5 probability between ages 5-7 and 11-13 years, but the pattern is very consistent with the other populations.

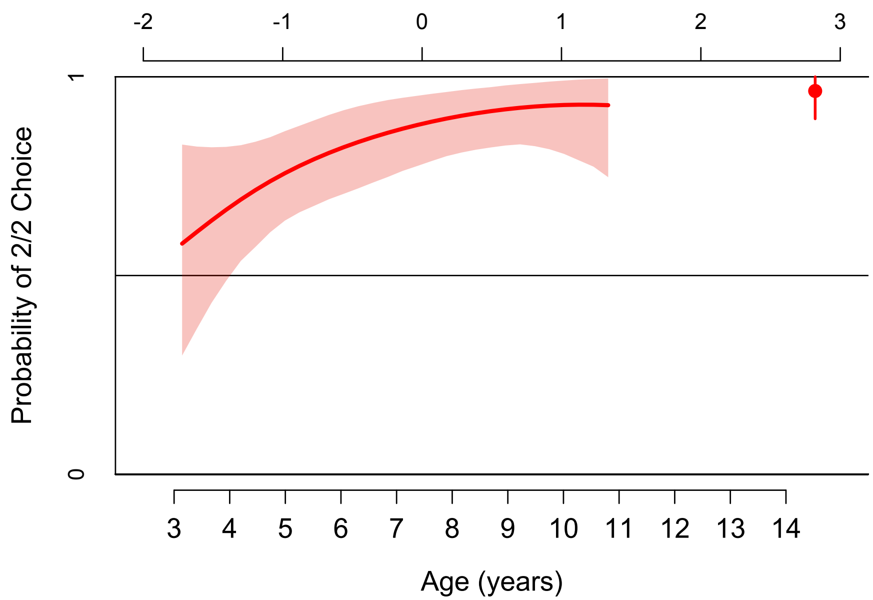
Figure S5(a): Age functions for all six populations for the FAM1 Social trial (same as Figure 3 from main text). Each colored line corresponds to children from a separate population, and the dots on the right hand side correspond to the mean choices of adults from each population in the same task (lines above and below each dot corresponds to 95% confidence intervals). The vertical axis on each plot corresponds to the estimated probability that children will choose the income-maximizing (2/2) outcome. The bottom horizontal axis corresponds to children's age (in years), and the top horizontal axis corresponds to equivalent values of CA.



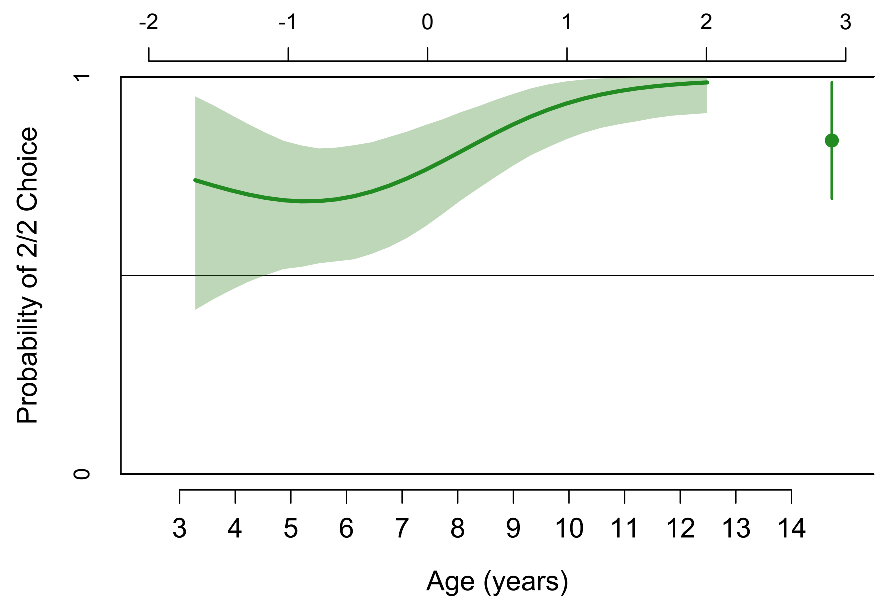
It is worth highlighting the fact that adults do not show a perfect pattern of selecting 2/2 over 1/1. It is highly unlikely that adults were unable to comprehend the nature of the task, meaning that we must not require perfect rates of choosing the income-maximizing outcome for deciding whether or not children understood the experiment. It is possible that during these trials participants may sometimes "overthink" the task (plausible as the task is very simple especially for adults and adolescents), or they may be testing what happens when they select an obviously 'incorrect' outcome. They may also simply be exploring the range of available choices.

SI Figures S5(b-g): 95% Confidence intervals for age functions in Figure 3 from main text. See Table S4(g) for associated models. The vertical axis on each plot corresponds to the estimated probability that children will choose the income-maximizing (2/2) outcome. The bottom horizontal axis corresponds to children's age (in years), and the top horizontal axis corresponds to equivalent values of CA. The solid line and the colored shaded area in each plot correspond to the estimated probability and 95% confidence interval. The dots on the right side of each plot correspond to adults' choices, and the lines above and below them correspond to 95% confidence intervals.

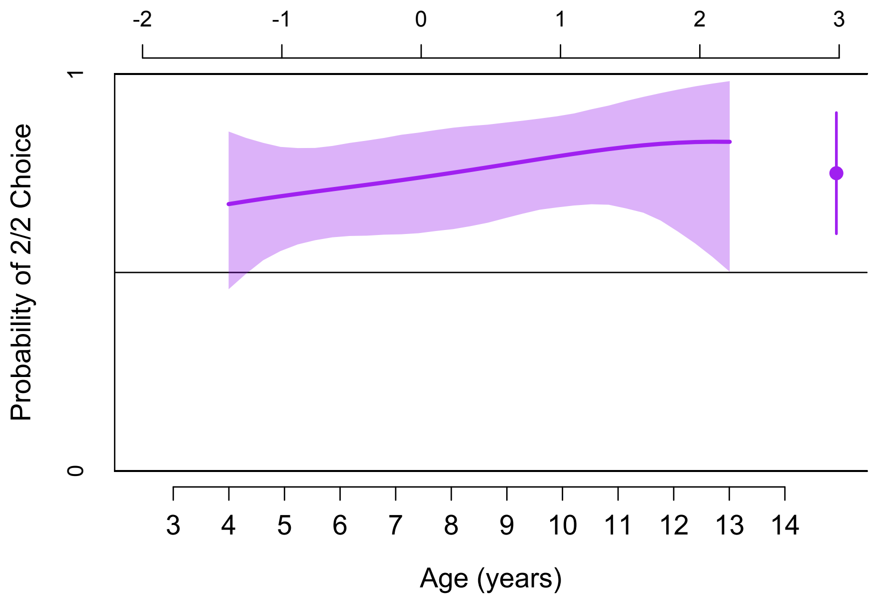
SI Figure S5(b): Los Angeles



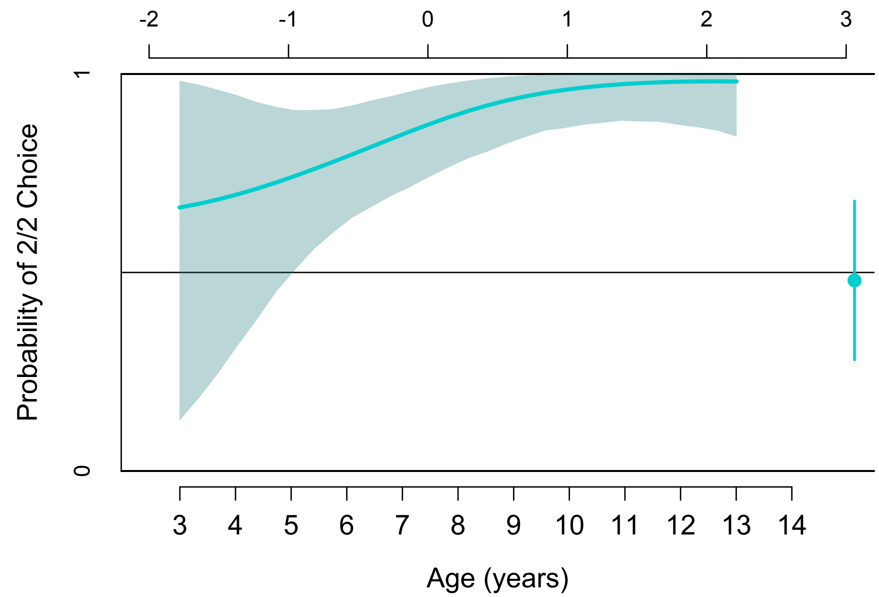
SI Figure S5(c): Fiji



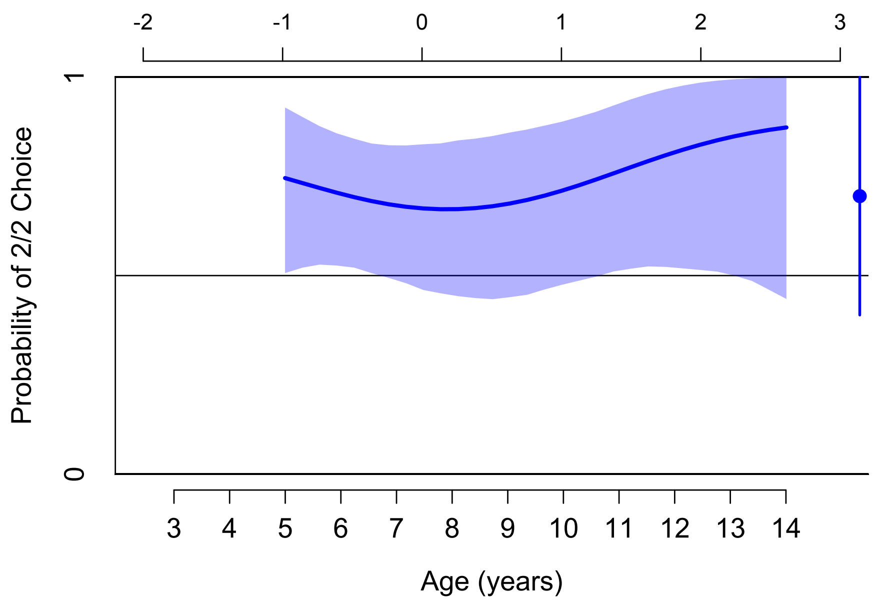
SI Figure S5(d): Himba



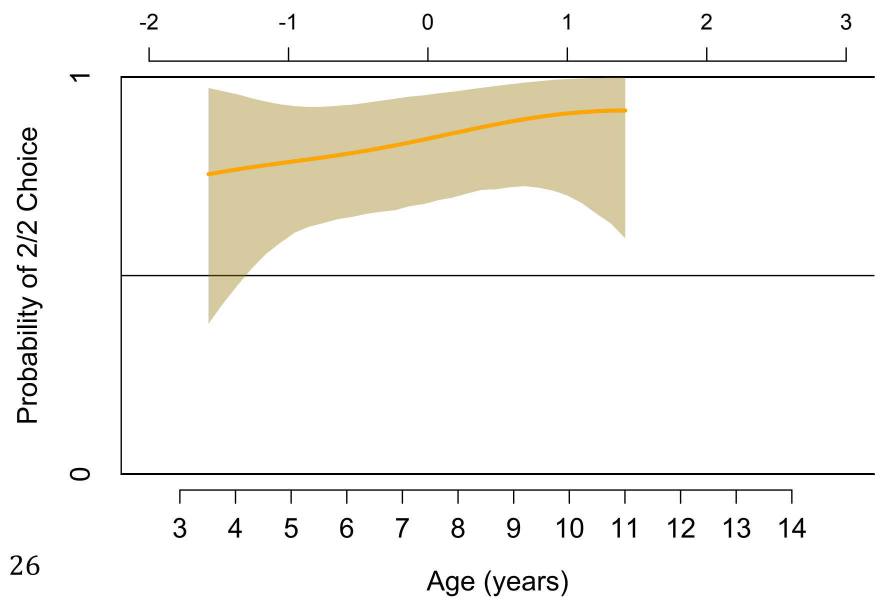
SI Figure S5(e): Shuar



SI Figure S5(f): Aka



SI Figure S5(g): Martu

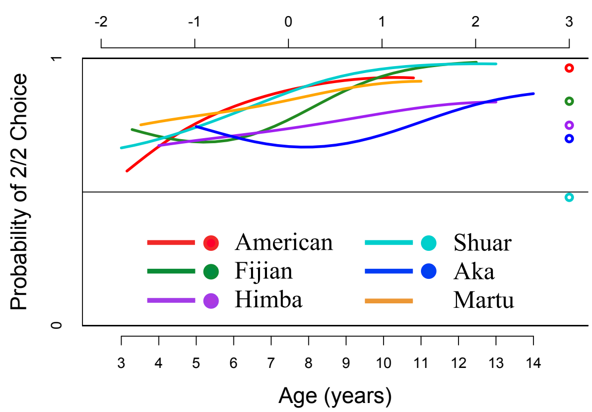


Comparison across all Familiarization Trials: For completeness, we also compare the patterns of development across our societies in the three other familiarization trials: Familiarization #2 Social, Familiarization #1 Asocial, and Familiarization #2 Asocial.

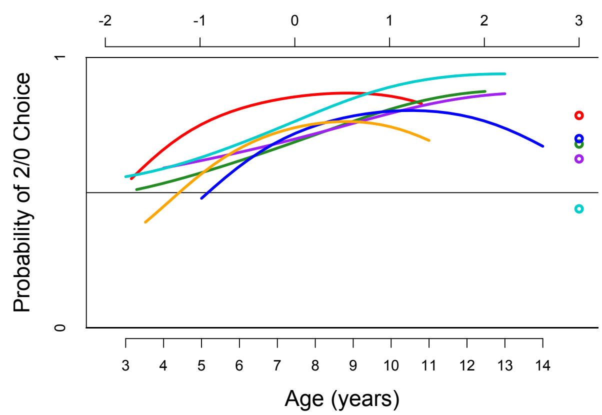
As already noted, we hesitate to draw conclusions from these trials because they are confounded: participants might select the income-maximizing outcome because it resulted in the highest payoff for them, or they might select the income-minimizing outcome because it reduced inequity between the participant and the recipient. If subjects were motivated by inequity aversion on these trials, then the probability of selecting income-maximizing outcomes (2/2 or 2/0) should fall among older children. This is because a growing number of studies have found increasing rates of inequity aversion among children aged about 8 years and older (22, 23, 24; references from main text). However, we should not see such a pattern in Familiarization #1 Social, as inequity aversion does not come into play here. Additionally, we might expect to see a greater drop in the probability of choosing 2/2 or 2/0 for those populations that are most prosocial in the CSG (behavior that may also be driven partly by inequity aversion).

SI Figures S5(h-k): The vertical axis on each plot below corresponds to the estimated probability that children will choose the income-maximizing (2/2) outcome. The bottom horizontal axis corresponds to children's age (in years), and the top horizontal axis corresponds to equivalent values of CA.

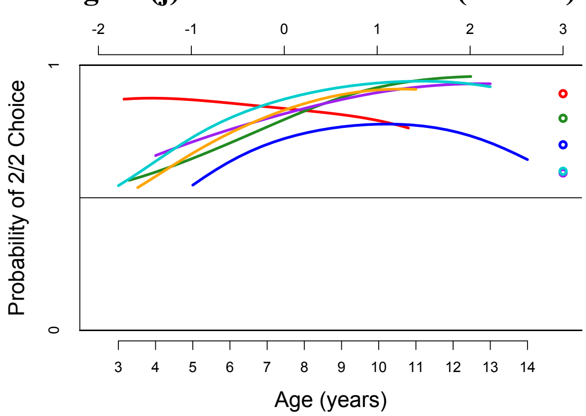
SI Fig. S5(h) Familiarization #1 (Social)



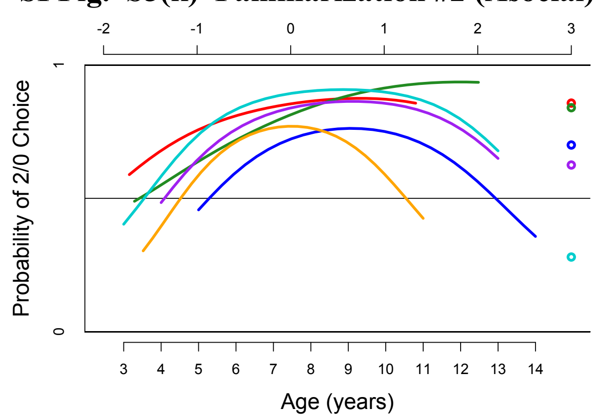
SI Fig. S5(i) Familiarization #2 (Social)



SI Fig. S5(j) Familiarization #1 (Asocial)



SI Fig. S5(k) Familiarization #2 (Asocial)



We do see some evidence of this overall pattern. We only observe relative drops in the probability of income-maximizing choices among older children in Figures S7 (i-k), where we would have predicted it. We also see larger drops among the populations that show a stronger shift towards prosociality among older children in the CSG: the U.S. and Aka, with some evidence of drops also among the Martu, Himba, and Shuar. Notably, the only population with no obvious drop in the probability of 2/2 and 2/0 choices is Fiji, the population that showed little to no evidence of a shift towards prosocial choices in older children in the CSG.

Finally, we again note that the distribution of adults' choices in these familiarization trials generally reflect the distribution of predicted choices among adolescents, though there is more variability than in the CSG. This suggests that older children and adolescents were not simply confused, and it is plausible that they were making purposeful choices, just not ones focused exclusively on income-maximization. However, the fact that adults also made variable choices in the familiarization trials reinforces the need for caution when interpreting behavior such as that represented in Figures S7(h-k), and we thus hesitate to draw strong inferences from these trials.

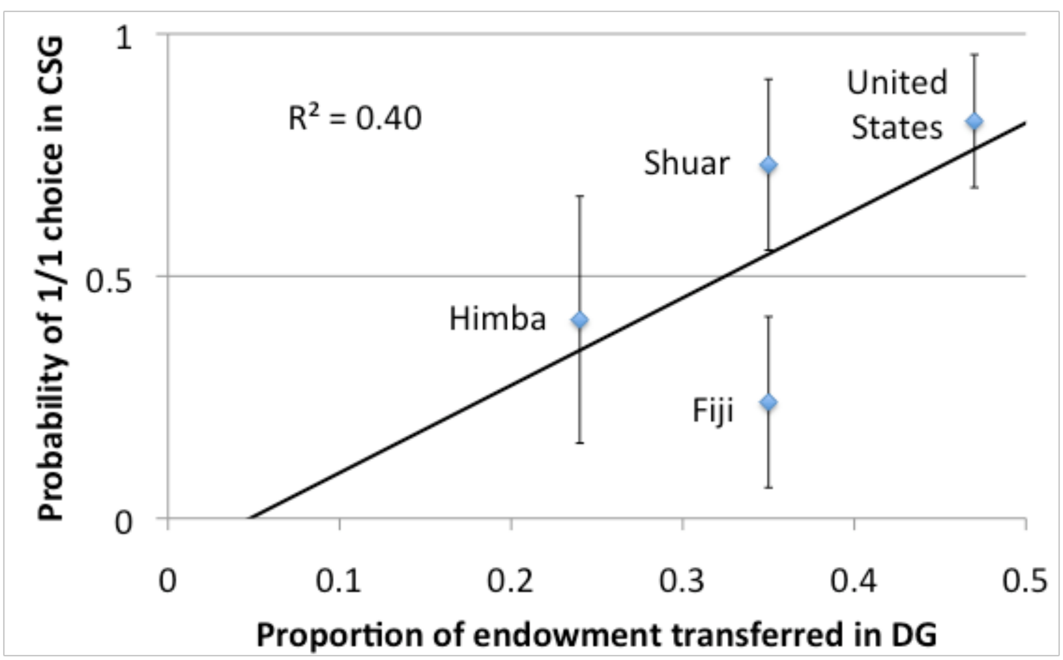
SI Appendix 8: The behavior of adults in the Costly Sharing Game (CSG) and the anonymous Dictator Game (DG).

SI Table S6: Adults' behavior in the CSG and the anonymous DG (* data drawn from (6; reference from main text); ‡ data collected as part of the current study, see SI Appendix 9 for methods).

	Proportion of 1/1 choices in CSG (Social Condition)	Proportion of endowment shared in anonymous Dictator Game
	Mean (SE) N	Mean (SE) N
United States	.82 (.07) 28	.47 (.03) 15 *
Shuar	.73 (.09) 26	.35 (.04) 21 *
Himba	.41 (.13) 32	.24 (.04) 32 ‡
Fiji	.24 (.09) 25	.35 (.03) 35 *

SI Figure S6: The proportion of adults' 1/1 choices in the CSG (vertical axis) are plotted against the behavior of adults from the same populations in the Dictator Game (horizontal axis). Values are drawn from SI Table S6.

SI Figure S6:



The positive relationship in SI Figure S6 between offers in the DG and the probability of a 1/1 choice in the CSG suggests that both tasks are measuring similar underlying prosocial tendencies (and similar variation in those tendencies across populations.

However, a better approach to exploring the relationship between behavior in the DG and CSG would use a regression model that limits estimates to possible outcomes (e.g., doesn't estimate probabilities of 1/1 choices in the CSG of <0 or >1), and which accounts for uncertainty in the measures of both the DG and the CSG. To do this we use regressions to first estimate behavior in the DG within each of our four populations (SI Table S7). We then use the posterior distributions of the mean DG offers to predict CSG behavior in each population (SI Table S8). This procedure provides a conservative estimate of the strength of association between DG offers and CSG behavior, by accounting for uncertainty in the average DG offer (SI Figure S7 for model structure).

SI Figure S7: Model structure. In this model, $\boldsymbol{y_{ik}}$ is the i-th DG offer from population k, and $\boldsymbol{x_{jk}}$ is the j-th CSG choice from population k ($\boldsymbol{\pi_{jk}}$ is the expected value of $\boldsymbol{x_{jk}}$). Fitting this model using Hamiltonian Monte Carlo allows us to integrate over uncertainty in $\boldsymbol{\mu_k}$ (the mean DG offer for each population) when estimating $\boldsymbol{\beta}$.

$$\boldsymbol{y_{ik} \sim \mathbf{Normal}(\mu_k, \sigma)}$$

$$\boldsymbol{x_{jk} \sim \mathbf{Bernoulli}(\pi_{jk})}$$

$$\boldsymbol{\mathbf{log}\frac{\pi_{jk}}{1 - \pi_{jk}} = \alpha + \beta\mu_k}$$

SI Table S7: Estimated DG Offers for each population. These largely correspond to the mean DG offers in SI Table S6, though the Monte Carlo sampling means that the estimated values here differ somewhat from the actual values. The actual values fall within the confidence intervals of our estimates.

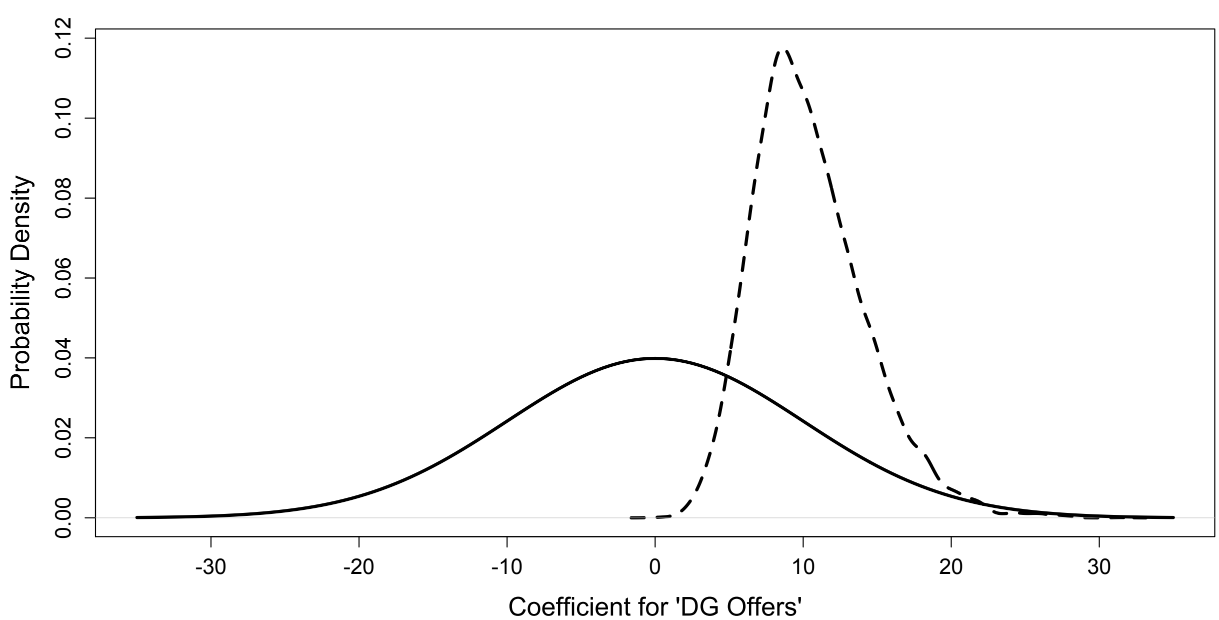
DV: Proportion of Endowment Given in Dictator Game N=103	**Coef. (95% CI)**
Intercept: United States	.48 (.40 : .57)
Intercept: Shuar	.39 (.32 : .45)
Intercept: Himba	.27 (.20 : .33)
Intercept: Fiji	.30 (.25 : .36)

SI Table S8: Estimate for *DG Offers* as a fixed effect parameter, predicting choices in the CSG.

DV: Probability of a 1/1 choice in the Costly Sharing Game N=120	**Coef. (95% CI)**
Intercept	-3.48 (-6.55: -1.28)
DG Offers	10.43 (4.32 : 19.10)

The parameter *DG Offers* has a positive estimate likely to be above zero, again suggesting a positive relationship between offers in the DG and choices in the CSG, but now incorporating uncertainty from both the DG and CSG measurements. Model results are based on 2000 samples from four chains, and we can plot both the assumed prior distribution (solid line, SI Figure S8) and the actual observed distribution (dashed line, SI Figure S8) of coefficient values for *DG Offers* across the 8000 samples modeled. SI Figure S8 shows that the actual distribution of coefficients for DG Offers is likely to be positive, as the density of the probability distribution that is near and below 0 is very small.

SI Figure S8: Prior distribution (solid line) and Posterior/observed distribution (dashed line) for the parameter *DG Offers* in SI Table S7.



SI Appendix 9: Methods for the anonymous Dictator Game conducted with the Himba.

Participants: 32 individuals participated (16 women). Subjects received a small gift (500 kg sugar) as a show-up fee, in addition to payoffs obtained as part of the study.

Experimental task: Subjects were informed that they would be allocated an endowment of 10 NAD (worth $1.30), and could keep the full amount or give some amount to another anonymous individual within their community (the recipient). One investigator laid 10 coins down on a wooden pallet, counting as they laid them down one by one. In Familiarization trials, subjects were asked a series of questions about hypothetical allocation decisions to make sure participants understood the game. In the Test trials, 10 coins were counted out on the pallet, and once the coins were in place, the subject was allowed to make a decision about how to allocate the coins to themselves and the recipient.

Anonymity: Both subjects and recipients were kept fully anonymous. Additionally, investigators were blind to the allocation decisions made by particular subjects. Before subjects made their choices, the investigators moved away from the pallet and stood with their back to the subject at a distance of about 5-8 meters. The subject then placed the coins that they wanted to keep in their pocket, and placed coins that they wanted to allocate to the anonymous recipient in an opaque plastic bag, which was half-filled with popcorn or rice to muffle the sound and obscure the weight of coins being placed in the bag. The subject tied up the bag and deposited it in a box beside the pallet. No identifying information, except the subject's sex, was placed on the bag. Recipient payoffs, transferred in the same opaque bags, were distributed three weeks later to community members at a public gathering open to all those living in the area.