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

nature human behaviour

Universal norm psychology leads to societal diversity in prosocial behaviour and development

Bailey R. House[1,2*], Patricia Kanngiesser[3], H. Clark Barrett[4], Tanya Broesch[5], Senay Cebioglu[5], Alyssa N. Crittenden[6], Alejandro Erut[4], Sheina Lew-Levy[7], Carla Sebastian-Enesco[8], Andrew Marcus Smith[4], Süheyla Yilmaz[3] and Joan B. Silk[2,9]

Recent studies have proposed that social norms play a key role in motivating human cooperation and in explaining the unique scale and cultural diversity of our prosociality. However, there have been few studies that directly link social norms to the form, development and variation in prosocial behaviour across societies. In a cross-cultural study of eight diverse societies, we provide evidence that (1) the prosocial behaviour of adults is predicted by what other members of their society judge to be the correct social norm, (2) the responsiveness of children to novel social norms develops similarly across societies and (3) societally variable prosocial behaviour develops concurrently with the responsiveness of children to norms in middle childhood. These data support the view that the development of prosocial behaviour is shaped by a psychology for responding to normative information, which itself develops universally across societies.

Human cooperative abilities are core to our success as a species[1,2] and differ in at least two important ways from those of other animals. First, people orchestrate group-level cooperation with large numbers of unrelated individuals. Second, cooperative behaviours vary considerably across societies[3,4], and this variation emerges during middle childhood[5–9], the developmental period between about 6 and 11 years of age. Some have suggested that the evolution of both can be explained if human social preferences are at least partly shaped by local cultural norms[5], which we acquire through an evolved psychology for learning and conforming to social norms[6,7]. According to this claim, we can explain what makes humans so successful by demonstrating (1) that our prosocial behaviour is linked to social norms, and (2) that we have a universally developing psychology for responding to these norms.

Norms are central to numerous theoretical models of human sociality and development[5,8–11], and are generally conceived as phenomena that regulate behaviour through prescriptions and proscriptions[12]. Following Bicchieri[8,13], we define a social norm as a behavioural rule that individuals conform to when they believe that: (1) a sufficiently large number of people in their community conforms to the rule (empirical expectation), and (2) a sufficiently large number of people in their community expect them to conform to the rule (normative expectations). A descriptive norm, by contrast, would focus on empirical expectations. There is already some evidence that norms underlie variation in prosociality across societies and groups[4,14]. However, most studies have only documented this variation across societies or explained it using society-level variables[3,15]. Empirical evidence that societal variation in normative expectations gives rise to variation in prosocial behaviours is needed. Such evidence would show that the prosocial behaviour of individuals is predicted by what members of their society believe to be normatively 'correct' in a particular situation (social norms). We must also distinguish the influence of social norms from that of individuals' own beliefs about what is correct (personal norms).

To connect societal variation in prosocial behaviour to the development of a universal psychology for social norms, we must also show that, across diverse societies, the tendency of children to respond to social norms is increasing during the same period that adult-like prosocial behaviour is forming. Children are sensitive to normative information as young as 1.5–4 years of age[16]. At this age, they enforce norm conformity in others[17], follow descriptive and injunctive norms[18,19], are sensitive to moral and conventional rules[20], and they know that different groups follow different norms[17]. Later, during middle childhood, children demonstrate an increasing responsiveness to novel social norms in experimental settings[18], suggesting that children of this age are becoming increasingly committed to modifying their behaviour to conform to social norms. Interestingly, this is the same age at which societal variation in children's prosocial choices seems to emerge in costly sharing tasks (that is, tasks involving a choice between outcomes that benefit oneself and outcomes that benefit others)[15,21–24]. These findings suggest that middle childhood is a particularly important period for the adoption of locally appropriate prosocial behaviours, and this could be the product of children's increasing responsiveness to social norms at this age. As children are already sensitive to norms by the time that they reach middle childhood, changes in behaviour during middle childhood may be due to developmental changes in their willingness to conform to norms, particularly their willingness to conform to norms that impose costs on them.

If societal variation in the prosocial behaviour of adults is linked to societal beliefs about correct prosocial choices, this provides evidence that prosocial behaviour is motivated by social norms above and beyond personal norms (prediction 1). If children's responsiveness to social norms is developing during childhood, then—with

[1]Department of Psychology, University of York, York, UK. [2]Institute of Human Origins, Arizona State University, Tempe, AZ, USA. [3]Faculty of Education and Psychology, Freie Universität Berlin, Berlin, Germany. [4]Department of Anthropology, University of California, Los Angeles, Los Angeles, CA, USA. [5]Department of Psychology, Simon Fraser University, Burnaby, British Columbia, Canada. [6]Department of Anthropology, University of Nevada, Las Vegas, Las Vegas, NV, USA. [7]Department of Psychology, University of Cambridge, Cambridge, UK. [8]Department of Education Science, Universidad de La Rioja, Logroño, Spain. [9]School of Human Evolution and Social Change, Arizona State University, Tempe, AZ, USA. *e-mail: bailey.house@york.ac.uk

Table 1 | Populations sampled

Population (location); description	*n* adult (female)	*n* child (female)	Child age range (year)	Norm-priming conditions received
German (Berlin, Germany); urban	32 (17)	111 (56)	4.07–13.36	Both OK, Generous and Selfish
Argentinian (La Plata, Argentina); urban	29 (13)	133 (65)	4.95–13.86	Both OK, Generous and Selfish
Wichí (Misión Chaqueña, Argentina); Rural, sedentarized hunter-gatherers	30 (19)	87 (47)	6.47–13.61	Both OK, Generous and Selfish
American (Phoenix, United States); urban	37 (19)	176 (92)	4.02–12.63	Both OK, Generous and Selfish
Indian (Pune, India); urban	30 (16)	155 (75)	4.11–13.92	Both OK, Generous and Selfish
Shuar (Amazonia, Ecuador); rural, small-scale horticulture, hunting	20 (8)	58 (27)	6.59–15.32	Both OK, Generous and Selfish
Tanna (Tafea province, Vanuatu); rural, small-scale horticulture, hunting	52 (27)	81 (43)	5.74–13.53	Both OK only
Hadza (Great Rift Valley, Tanzania); rural, foraging, hunting	25 (12)	32 (10)	7.00–17.00	Both OK only

For more details see Supplementary Table 1.

increasing age—their prosocial choices should become more adult-like and variable across societies (prediction 2), and also more strongly influenced by novel social norms (prediction 3). This would provide two independent sources of evidence for the hypothesis that social norms have an increasing influence on children's prosocial behaviour as they mature, and would be consistent with the results of previous studies. If the willingness of children to respond to norms develops similarly across a wide range of different societies, it would provide evidence for a universally developing human psychology for responding to social norms (prediction 4). If this societally common responsiveness to norms develops concurrently with the development of adult-like prosocial behaviour, it would provide evidence that a universally developing psychology for social norms can explain the emergence of societal variation in prosocial behaviour (prediction 5).

We conducted field experiments on prosocial behaviour using the dictator game (DG) as a measure of costly sharing with 255 adults (131 female) and 833 children (414 female) aged 4–17, in eight populations ranging from foragers to small-scale horticulturalists to large urban communities (Table 1). The DG provides a well-validated test of an individual's willingness to share with others at a personal cost, and its standardized design enables direct comparison across populations[25] (Fig. 1). We used a binary-choice version of the DG in which participants chose between two options—they could keep two rewards and give none to an absent anonymous peer (the '2/0' self-maximizing option), or they could keep one reward and give one to the peer (the '1/1' prosocial option). This version of the task is appropriate for children aged ≥4 years and adults[15,26,27]. All of the child and adult participants included in the dataset passed three comprehension questions confirming they understood the DG procedure, the content of the norm prime videos and that larger quantities of rewards resulted in higher payoffs (Supplementary Information, page 25).

Before each participant made their choice in the DG, they viewed a short video in which an adult model verbalized novel normative information about the two options in the DG (we refer to this normative information as the 'norm prime')[18]. Videos used a standardized script, but were recorded at each fieldsite using local translations of the script and local adults as models. Across three between-participants conditions, participants were presented with different norm primes. In the 'Generous' condition, the norm prime indicated that 1/1 was 'right' and 'good to choose', whereas 2/0 was 'wrong' and 'bad to choose'. In the 'Selfish' condition, the norm prime indicated that 2/0 was 'right' and 'good to choose' and 1/1 was 'wrong' and 'bad to choose'. Importantly, the videos did not show the model making a choice in the DG, they simply presented the norm prime as if musing about the choice between 1/1 and 2/0. In the 'Both OK' condition, the model stated that 1/1 and 2/0 were both 'OK' and 'OK to choose', language that is not strongly normative but that could arguably be at least weakly normative. Regardless, Both OK provides a reference point about participants' prosocial choices when they have been given information that does not preferentially bias them towards either 1/1 or 2/0 (as was the case for the generous and selfish conditions).

To test prediction 1, adult participants in all eight societies received only the Both OK norm prime before they made their choice in the DG. We used these data to assess variation in the probability of adults' 1/1 choices across societies. In seven of the eight societies we also elicited judgements about which norm prime was correct (practical limitations required us to use an abbreviated procedure with the Hadza society, precluding collection of data on judgements; Supplementary Information, page 28). In these seven societies, after participants had made their choice in the DG, they were presented with both the Generous and Selfish norm prime videos (randomizing the order of presentation), and the participants were asked which of the videos they believed to be 'more correct'. This judgement could be influenced both by what participants believe is correct for them to choose (personal norms), and by what they believe is correct for others to choose (social norms). If the DG choices of individuals are influenced by social norms, then their choices are expected to be predicted by the judgements of others in their society (that is, beliefs of others about what is the correct norm) in addition to their own judgements. The judgements of participants enable us to study how society-level beliefs influence prosocial behaviour, without requiring participants to explicitly report what they think other members of their community believe to be correct. This is important for a cross-cultural study, as comfort and familiarity with discussing the thoughts or mental states of others varies across societies[28].

To test prediction 2, a subset of child participants in all eight societies also received the Both OK prime, and we studied how the probability of children's 1/1 choices changes with age in the Both OK condition. To determine how adult-like prosocial behaviour develops, we explored whether children's prosocial choices were predicted by the prosocial choices of adults from their own society, and whether this relationship changed as a function of children's age.

To test prediction 3, in six of the eight societies, two additional subsets of children were presented with either the Generous or the Selfish norm primes (practical limitations prevented testing

Fig. 1 | Arrangement of the apparatus and testing area. Example setup in Pune, India (left), Phoenix, United States (middle) and La Plata, Argentina (right).

Table 2 | Model comparisons for models 1a–c using WAIC and AIC weight

Model	Model parameters	WAIC (s.e.)	dWAIC (s.e.)	AIC weight
1a	Intercept only	355.00 (1.37)	26.00 (10.22)	0.00
1b	**Society D(8)**	**329.10 (10.27)**	**0.00 (NA)**	**0.95**
1c	Society D(8), age, age × society D(8), gender, gender × society D(8)	334.80 (11.07)	5.70 (3.71)	0.05

D indicates a dummy parameter; society D(*X*) indicates that multiple dummy parameters were used for *X* number of societies. The model with the lowest WAIC provides the best fit, dWAIC indicates the difference in WAIC between the focal model and the best-fit model. Where AIC weight is substantially larger for the best-fit model, this implies that it provides a substantially better fit to the data. Where dWAIC is larger than its s.e., this also implies that the best-fit model provides a substantially better fit to the data. All of the comparisons were conducted using the 'compare' function in the R package 'rethinking', with $n = 40{,}000$ samples from the posterior for computing WAIC. Bold indicates the model that provides the best fit to the data. NA, not applicable.

these additional samples in both the Tanna and Hadza societies) in a between-participants design. We investigated whether the Generous prime increased the probability of 1/1 choices relative to the Both OK prime, and whether the Selfish prime decreased this probability relative to the Both OK prime. If the prosocial choices of participants in the DG were responsive to the normative information provided by the priming videos, then the participants were expected to be more likely to choose 1/1 in the Generous prime condition than in the Both OK prime condition, and were also expected to be more likely to choose 1/1 in the Both OK prime condition than in the Selfish condition. To test prediction 4, we then explored whether the development of children's responsiveness to the primes varied across societies.

To test prediction 5, we compared the development of adult-like DG choices in children, the development of children's responsiveness to novel social norms (for example, Generous and Selfish norm primes) and the development of children's tendency to make choices that are consistent with the beliefs of adults about social norms (that is, the probability that adults in their society judged the Generous norm prime to be most correct). If these different developmental trajectories align and if children's responsiveness to novel social norms develops similarly across societies, this will suggest that societal variation in prosocial development is linked to the development of a universal psychology for responsiveness to social norms.

Results

Prediction 1. We explored whether the DG choices of adults varied across societies by comparing regression models using widely applicable information criterion (WAIC) and applicable information criterion (AIC) weight (Table 2). Model 1a represents the hypothesis that DG choices do not vary across societies. Model 1b represents the hypothesis that DG choices vary across societies, and includes dummy parameters for each society. We had no predictions about age and gender for adult participants, but model 1c included interactions with these variables to consider whether they were important. Model 1b provides a better fit to the data than the other models (reflected by lower WAIC and higher AIC weight; Table 2), indicating that the probability of a 1/1 choice varied substantially across societies, and that this variation was not a by-product of variation in the distribution of age or gender across societies (the estimates of models 1c and 1b are similar; Supplementary Fig. 12).

The probability that adults would choose the 1/1 option varied across societies (model 1b; Fig. 2a and Supplementary Table 3). Information about mean amounts given in a continuous DG were available for three of the societies in our sample (from a previous study), and the proportion of 1/1 choices in the present binary DG (Americans, 0.54; Shuar, 0.20; Hadza, 0.20) were similar to the mean amounts given in the continuous DG[4] (Americans, 0.45; Shuar, 0.34; Hadza, 0.26; Fig. 2a, triangles). This suggests that societal variation in choices is stable across different versions of the DG, and also that our experiment elicits a form of prosocial behaviour that has been linked to cultural adaptations related to religious beliefs, market norms and norms for living in large communities[3–5,29]. This also implies that the Both OK prime does not alter the preferences of individuals in the DG.

The probability that adults judge the Generous norm prime to be more correct also varied across societies (model 2; Fig. 2b and Supplementary Table 3). To determine whether societal differences in judgements about correct norms predicted the DG choices of participants, we constructed a two-stage model. The first stage of model 3 was equivalent to model 2, and estimated the probability that adults in each society would judge the Generous norm prime to be more correct. The second stage predicted the 1/1 choices of adults in the DG using (1) the first-stage estimates of the probability that the Generous norm prime would be judged to be more correct in an adult's society, and (2) adults' own judgements as to whether the Generous norm prime was more correct. Both of these parameters predicted the DG choices of adults. Participants were more likely to choose 1/1 if they themselves later judged the Generous norm prime to be more correct (coefficient, 1.61; s.d., 0.33; 95% confidence interval (CI), 0.96–2.27; Fig. 2c and Supplementary Table 4). Participants were also more likely to choose the 1/1 option if they lived in a society in which people were generally more likely to judge the Generous norm prime to be more correct (coefficient, 0.46; s.d., 0.22; 95% CI, 0.08–0.93; Fig. 2c and Supplementary Table 4).

Prediction 2. We explored whether children's DG choices changed with age in the Both OK condition (the same condition presented to adults) by again comparing the models using WAIC and AIC weight (Table 3). Model 4a included only society dummy parameters (the same structure as model 1b), representing the hypothesis that DG choices vary across societies but do not change with age. Model 4b represents the hypothesis that children's choices changed with age, by including interactions between society dummies and child age. Model 4c included an age^2 parameter to explore whether a U-shaped

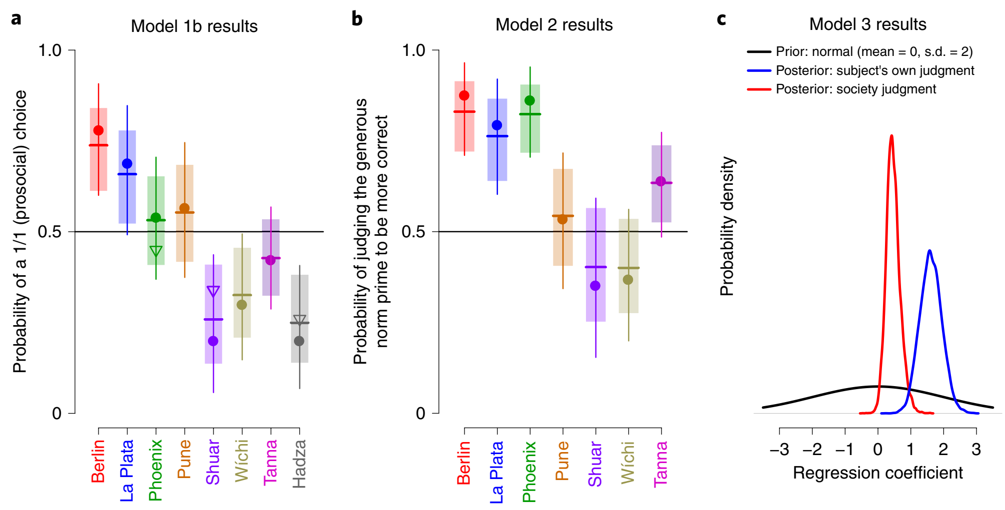

Fig. 2 | Results of models 1b, 2 and 3. a, Results of model 1b, which estimates the probability that adults chose 1/1 in the eight societies. The horizontal lines and shaded regions represent the regression estimates and 95% CIs (functions 'MAP' and 'link', R package 'rethinking'). The circles and vertical lines represent the proportions and 95% CIs of the raw data (function 'binom.confint', R package 'binom'; the exact proportions are provided in Supplementary Table 2). The triangles represent the mean DG offers in a previous study by Henrich et al.[4]. **b**, Results of model 2, which estimates the probability that adults judged the Generous norm prime to be most correct. The horizontal lines and shaded regions represent regression estimates and 95% CIs (functions 'MAP' and 'link', R package 'rethinking'; the exact proportions are provided in Supplementary Table 2). The circles and vertical lines represent the proportions and 95% CIs of the raw data (function 'binom.confint', R package 'binom'). **c**, Results of model 3, which estimates how the 1/1 choices of adults are predicted by whether they judged the Generous norm prime to be most correct, and by the estimated probability that someone in their society would judge the Generous norm prime to be most correct. The black line shows the weak prior distribution; the red line shows the 'posterior distribution for the estimated probability of society judgement' parameter (Supplementary Table 4); and the blue line shows the posterior distribution for the 'subject's own judgement parameter' (Supplementary Table 4).

Table 3 | Model comparisons for models 4a–d using WAIC and AIC weight

Model	Model parameters	WAIC (s.e.)	dWAIC (s.e.)	AIC weight
4a	Society D(8)	411.00 (16.46)	7.90 (7.97)	0.02
4b	**Society D(8), age, age × society D(8)**	**403.10 (17.62)**	**0.00 (NA)**	**0.80**
4c	Society D(8), age, age × society D(8), age^2, age^2 × society D(8)	407.40 (17.88)	4.40 (3.55)	0.09
4d	Society D(8), age, age × society D(8), gender, gender × society D(8)	407.20 (18.02)	4.20 (4.12)	0.10

All of the comparisons were conducted using the 'compare' function in the R package 'rethinking', with $n = 40{,}000$ samples from the posterior for computing WAIC. Bold indicates the model that provides the best fit to the data.

effect of age would improve model fit, and model 4d included interactions between society dummies and participant gender. Model 4b provides a better fit to the data than the other models, shown by a higher AIC weight (Table 3). There is a large s.e. for the difference in WAIC for model 4a, so we report the results for both model 4a and 4b in Supplementary Table 5. These analyses suggest that children's 1/1 choices changed with age, which is shown by the plot of model 4b (Fig. 3a). Plotting the estimates of models 4c and 4d suggests that they produce qualitatively similar results (Supplementary Figs. 13 and 14).

We explored whether children's DG choices became increasingly similar to those of adults with age using a two-stage model (model 5). The first stage was similar to model 1b, and estimated the probability that adults in each society would choose the 1/1 outcome. The second stage predicted each child's DG choice using the first-stage estimates of the probability that adults from their society would chose 1/1, and included an interaction between the first-stage estimates and child age. The interaction was reliable, indicating that children's DG choices were increasingly predicted by the DG choices of adults with increasing age (coefficient, 0.55; s.d., 0.27; 95% CI, 0.09–1.16; Supplementary Table 6). Plotting this relationship shows that the model predicts that children's choices become positively related to the choices of adults after the age of about 8, with this estimate becoming reliably different from zero after the age of about 10 (Fig. 3b).

Predictions 3 and 4. We explored whether children's DG choices were influenced by norm primes by comparing models of children's choices in all three conditions (Both OK, Generous and Selfish) in the six societies for which these data were available (excluding Tanna and Hadza), once more comparing models using WAIC and AIC weight (Table 4). Model 6b represents the hypothesis that children responded to norm primes by including dummy parameters for the Generous and Selfish conditions. Model 6a represents the hypothesis that children did not respond to norm primes by excluding these parameters (the same model structure as model 4b). Model 6c represents the hypothesis that children's responsiveness to norm primes changes with age by including interactions between child age and dummies for Generous and Selfish. Model 6d represents the hypothesis that the development of a responsiveness to norm primes varies across societies by including three-way interactions with society dummies, child age, and dummies for Generous and Selfish.

Model 6c provides a substantially better fit to the data than model 6a or model 6d (Table 4), suggesting that children were responsive to norm primes and that this responsiveness developed similarly across societies. Model 6c had a slightly larger AIC weight than model 6b (Table 4), indicating that both models fit the data

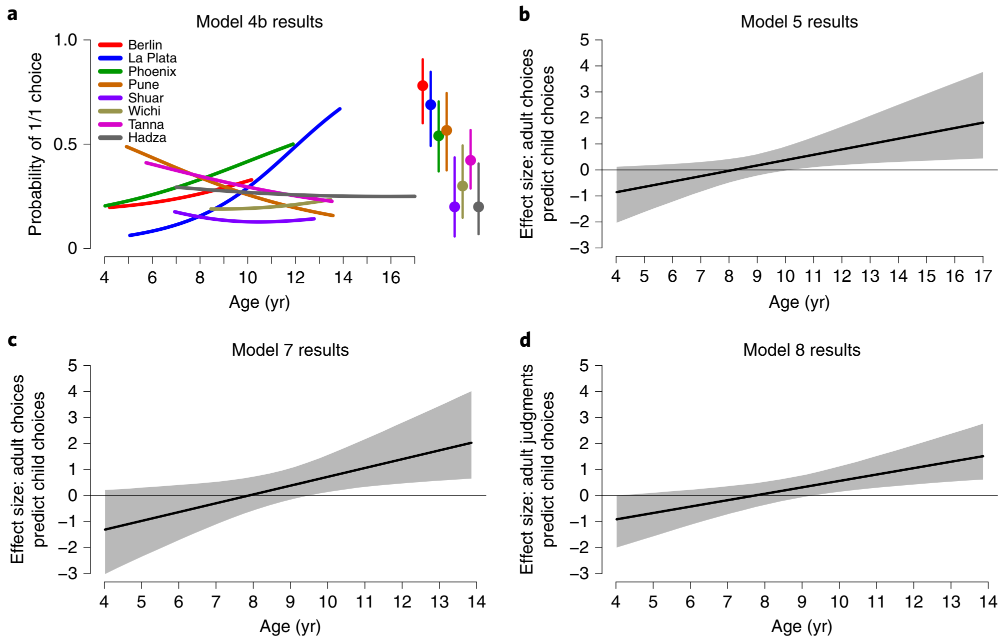

Fig. 3 | Results of models 4b, 5, 7 and 8. a, Results of model 4b. The lines represent regression estimates of the probability that children will choose 1/1 in each of the eight societies, as a function of participant age (functions 'MAP' and 'link', R package 'rethinking'). The circles and vertical bars represent the proportions and 95% CIs of adults' choices of 1/1 (function 'binom.confint', R package 'binom'). The model coefficients are provided in Supplementary Table 5. **b**, Results of model 5. The solid line plots the magnitude with the 95% CI of the estimated relationship between children's 1/1 choices and the 1/1 choices of adults, as a function of child age (the model was constructed using Rstan; link, Bernoulli_logit). This captures the emerging positive relationship between older children's DG choices and the DG choices of adults' from their society. The negative values of the effect size for the youngest children is due to young children in Berlin and La Plata being the least likely to choose 1/1 despite adults from those societies being the most likely to choose 1/1. Model coefficients are provided in Supplementary Table 6. **c**, Results of model 7. The solid line plots the magnitude with the 95% CI of the estimated relationship between children's 1/1 choices and the 1/1 choices of adults, as a function of child age (the model was constructed using Rstan; link, Bernoulli_logit). This captures the emerging positive relationship between older children's DG choices and the DG choices of adults from their society. **d**, Results of model 8. The solid line plots the magnitude with the 95% CI of the estimated relationship between children's 1/1 choices and the judgements of adults as to whether or not the generous norm prime is most correct, as a function of child age (the model was constructed in Rstan; link, Bernoulli_logit). This captures the emerging positive relationship between older children's DG choices and judgements about norms by adults from their society.

well, but the inclusion of parameters for child age improved the model fit to some extent. For both models 6b and 6c, there are reliable effects for the Generous dummy (model 6b: coefficient, 1.47; s.d., 0.20; 95% CI, 1.07–1.86; model 6c: coefficient, 1.47; s.d., 0.20; 95% CI, 1.07–1.86; note: nearly identical estimates; Supplementary Table 7) and also for the Selfish dummy (model 6b: coefficient, −1.00; s.d., 0.24; 95% CI, −1.48 to −0.52; model 6c: coefficient, −1.03; s.d., 0.25; 95% CI, −1.52 to −0.54; Supplementary Table 7). This means that children were substantially more likely to choose 1/1 when they received the Generous norm prime (relative to Both OK), and substantially less likely to choose 1/1 when they received the Selfish norm prime.

Model 6c also provides evidence of an interaction between child age and the Generous dummy that is borderline reliable, as the lower CI is zero (coefficient, 0.40; s.d., 0.21; 95% CI, 0.00–0.81; Supplementary Table 7); by contrast, the interaction between child age and the Selfish dummy was not reliable (coefficient, 0.29; s.d., 0.25; 95% CI, −0.19–0.78; Supplementary Table 7). This suggests a modest developmental increase in children's responsiveness to the Generous norm prime but not to the Selfish norm prime. Plotting these model estimates (Fig. 4a) indicates that children are more likely to choose 1/1 in the Generous norm prime condition than in the Both OK condition across the entire age range, but this responsiveness to Generous is only reliable after the age of about 6–7 and increases through middle childhood. Plotted estimates also suggest that children are less likely to choose 1/1 in the Selfish norm prime condition across the entire age range, but this responsiveness to Selfish norm primes is less pronounced, less consistently reliable and shows little sign of change with age.

The model comparison analysis in Table 4 implies that the development of children's responsiveness to norm primes did not vary substantially across societies. To confirm this, we plotted the results of model 6d separately for each of the six societies (Fig. 4b–g). All of the societies revealed a responsiveness to the norm primes. For four of the societies (La Plata, Shuar, Pune and Wichí) the results are qualitatively consistent with the overall developmental pattern in Fig. 4a—responsiveness to norm primes becomes reliable at some point around age 6–7 and increases thereafter, particularly for the Generous norm prime. The developmental pattern for responsiveness to the Selfish norm prime seems to be more inconsistent, plausibly owing to a floor effect in some of the societies in which children are unlikely to choose 1/1 in the Both OK condition. For the two other societies (Berlin and Phoenix), reliable differences between each of the conditions seem to emerge by age 4 or earlier, and children's responsiveness to the norm primes seems to change less with age (with the exception of reduced responsiveness to the Selfish norm prime in older children in Phoenix).

Prediction 5. We explored the relationship between children's DG choices and adults' DG choices (as for prediction 2) in the six societies in which we investigated responsiveness to norm primes. This afforded the closest comparison between the development of adult-like DG choices and the development of a responsiveness to norm primes. Model 7 used the same two-stage structure as model 5 and

Table 4 | Model comparisons for models 6a–d using WAIC and AIC weight

Model	Model parameters	WAIC (s.e.)	dWAIC (s.e.)	AIC weight
6a	Society D(6), age, age × society D(6)	883.20 (21.77)	136.60 (22.40)	0.00
6b	**Society D(6), age, age × society D(6), Generous D, Selfish D**	**747.00 (27.90)**	**0.30 (4.05)**	**0.45**
6c	**Society D(6), age, age × society D(6), Generous D, Selfish D, age × Generous D, age × Selfish D**	**746.60 (28.16)**	**0.00 (NA)**	**0.53**
6d	Society D(6), age, age × society D(6), Generous D, Selfish D, age × Generous D, age × Selfish D, age × Generous D × society D[6], age × Selfish D × society D(6)	754.00 (28.69)	7.40 (6.91)	0.01

All of the comparisons were conducted using the 'compare' function in the R package 'rethinking', with $n = 40{,}000$ samples from the posterior for computing WAIC. Bold indicates the models that provide the best fit to the data.

produced the same result—convergence between the 1/1 choices of adults and children increased with age (coefficient, 0.85; s.d., 0.42; 95% CI, 0.15–1.81; Supplementary Table 8). Using the same approach and model structure, model 8 analysed the relationship between children's DG choices and the judgements of adults. This model shows that, with age, children's 1/1 choices were increasingly predicted by the estimated probability that adults from their society would judge the Generous norm prime to be more correct (coefficient, 0.61; s.d., 0.24; 95% CI, 0.20–1.16; Supplementary Table 9). Plotting both of these results shows that from the age of about 8, children's choices are positively related to both the DG choices and the judgements of adults, and this relationship was reliable from the age of about 9–10 (Fig. 3c,d). These analyses reveal that both the DG choices and the judgements of adults predict children's choices, but not whether these are independent effects (when both parameters are included in a single model, neither effect is reliable; Supplementary Table 10).

Discussion

Here we present three main findings: (1) cross-cultural variation in the prosocial behaviour of adults is related to what members of their society judge to be the correct prosocial norm (prediction 1); (2) in middle childhood and early adolescence, children's prosocial behaviour becomes increasingly similar to the prosocial behaviour of adults (prediction 2), and also increasingly similar to adults' judgements about the correct prosocial norm; and (3) by middle childhood, children in very different societies develop a uniform tendency to respond to novel social norms about prosocial behaviour, and this coincides with the development of adult-like societal variation in that behaviour (predictions 3, 4 and 5). Together, these findings link societal variation in prosociality to the development of a universal psychology for responding to social norms.

The DG choices of adults were predicted by the probability that members of their society would judge the generous norm prime to be more correct. This effect was independent of the influence of the personal norms of individuals, indicating that individuals' prosocial choices were related to local social norms (that is, society-level beliefs about what is correct). We note that this need not have been the case—the personal norms of individuals could have been the only factor predicting decisions, and other differences across communities (for example, relatedness, community size and migration rates) could have created enough societal variation in prosocial choices to swamp the influence of societal-level norms.

During middle childhood, children's prosocial choices became increasingly predicted by the prosocial choices of adults from their own societies, with this relationship emerging by the age of about 8–10 at the latest. This is consistent with findings from previous studies showing that societal variation in prosociality and fairness emerges during middle childhood and early adolescence[15,21,22]. We extend this work to show that, during middle childhood (by age 8–10), children's choices become increasingly predicted by the probability that adults from their society would judge generous norm primes to be more correct. This is consistent with our finding that adults' own prosocial choices were predicted by local beliefs about what is correct, and it reinforces the idea that children's prosocial choices are becoming both more adult-like and more attuned to local prosocial norms during this developmental period. Future studies should explore whether these are independent developmental phenomena, and whether children's prosocial behaviour is better predicted by adults' prosocial behaviour or judgements about local norms.

Although children's prosocial choices generally became more adult-like with age, there were exceptions to this pattern. For example, in Pune and Tanna, older children were less likely to choose 1/1 than adults. At both of these sites, adults chose 1/1 with a probability close to 0.5 and they also judged the Generous norm prime to be more correct with a probability close to 0.5. This suggests that adults in these communities held a variety of beliefs about correct norms for behaviour in the DG, and this heterogeneity could complicate children's attempts to navigate towards adult-like patterns of behaviour. This interpretation is supported by the pattern of variation in the Phoenix sample. In Phoenix, adults chose 1/1 with a probability close to 0.5, but they were much more likely to judge the Generous norm prime to be more correct. This suggests that there was a greater consensus in beliefs about prosocial norms in the DG in Phoenix than in Pune and Tanna, and it may explain why children in Phoenix seemed to follow the overall trend towards adult-like behaviour. The lack of clearly adult-like choices for children in Berlin may be an artefact of the composition of the sample. In Berlin, the oldest children in our Both OK norm prime sample were only around 10 years of age, the age at which reliably adult-like choices begin to emerge.

In previous studies with the DG, we found that children were more generous in early childhood than in middle childhood[15,27], a U-shaped pattern that we did not replicate here. This may be due to different experimental procedures. Previously, participants were face-to-face with their partners, whereas in the present study, participants were alone and anonymous. By the age of about 5, children are more likely to be selfish when they are unobserved[30,31], but it is unlikely that children younger than 5 use anonymity strategically as they are not very good at managing their reputation[32]. It is more plausible that the lack of face-to-face contact with a partner in our study reduced social factors, such as empathy[33] and a desire to interact with others[34], factors that are more likely to motivate prosociality at this age. Future work should directly compare the influence of these factors (as well as motivations, such as strategic reciprocity[23,35–37], kin biases[38] and group biases[39]) with the influence of norms on costly prosocial behaviour in early and middle childhood.

Our experiments show that novel social norms influenced children's prosocial choices. Children's responsiveness to novel norms developed similarly across societies, generally increasing with age and becoming a reliable effect by the age of about 6–8. This suggests that children's sensitivity to novel norms is growing at the same age

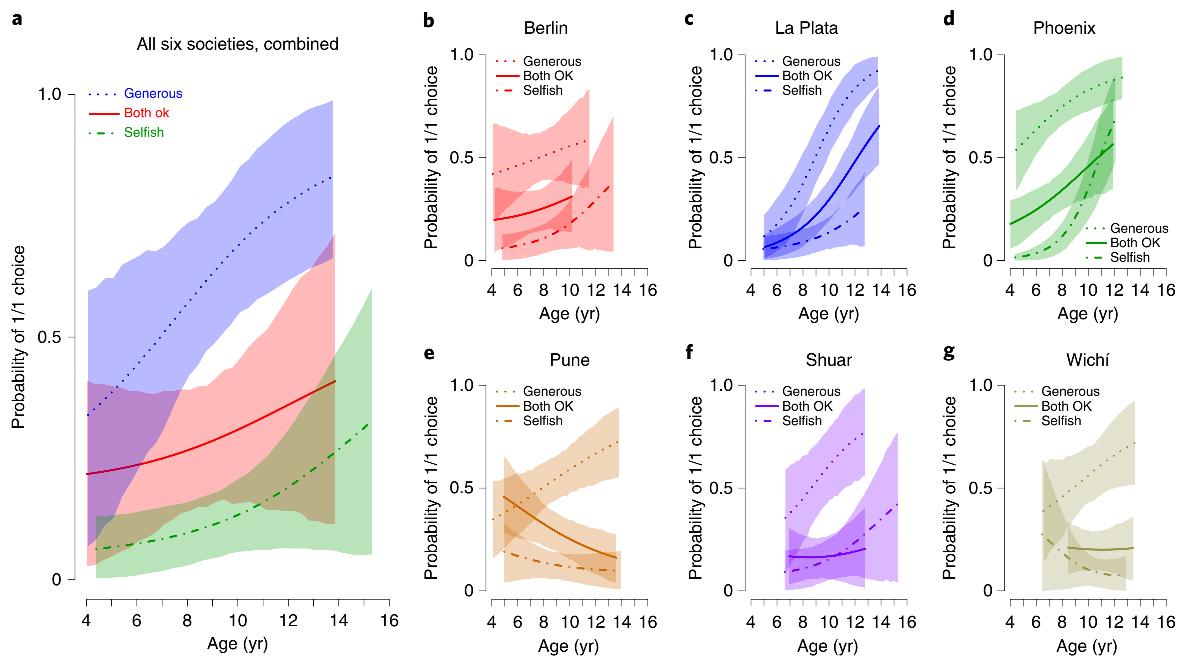

Fig. 4 | Results of models 6c and 6d. a, The lines and shaded regions represent regression estimates and 95% CIs for the probability that children will choose 1/1 in the Generous, Both OK and Selfish norm prime conditions, combining samples from the six different societies (functions 'MAP' and 'link', R package 'rethinking'). **b–g**, The lines and shaded regions represent regression estimates and 95% CIs for the probability that children will choose 1/1 in the Generous, Both OK and Selfish norm prime conditions for the Berlin (**b**), La Plata (**c**), Phoenix (**d**), Pune (**e**), Shuar (**f**) and Wichí (**g**) societies (functions 'MAP' and 'link', R package 'rethinking').

at which their choices are also becoming more adult-like and more consistent with the judgements of adults about correct behaviour. Evidence for a developmental increase in children's responsiveness to Generous norm primes was modest, but this, nonetheless, provides evidence for developmental changes in children's willingness to pay a cost to conform to a norm.

Developmental changes may have been obscured by children's responsiveness to norms emerging earlier in some societies than in others. For four societies (La Plata, Shuar, Pune and Wichí), responsiveness to generous norm primes became reliable by the age of about 6–8, and seemed to increase through middle childhood. However, for two societies (Berlin and Phoenix), reliable responsiveness emerged by age 4. This implies not only broad cross-cultural similarity in development, but also some variation in timing. This is consistent with previous studies showing that the foundations of moral evaluation[40], prosocial behaviour[41] and normative behaviour are present early in childhood[16–20]. It also supports the proposal that adult-like prosocial behaviour emerges due to increases in children's responsiveness to normative information rather than fundamental changes in their ability to conform to norms (given that, in at least some societies, this is present earlier).

Our studies suggest that the emergence of adult-like prosocial behaviour is linked to the development of children's responsiveness to normative information. Future research should explore in detail how children's willingness to respond to norms changes during middle childhood, and how the development of this willingness predicts children's tendency to behave like adults. In Phoenix and Berlin, children younger than age 6–8 were willing to pay a cost to conform to norms, but they did not make very adult-like choices in the Both OK condition. In these societies, children's responsiveness to norms in early childhood may be based less on a general interest in behaving normatively, and more on a tendency to interpret normative information as 'what adults want them to do'. If adults in these societies tend to strongly encourage and enforce normative behaviour at young ages, children may have learned to simply do whatever adults say the right thing to do is. This highlights that the critical developmental change in middle childhood is likely an increasing willingness to pay a cost to behave normatively, and it will be crucial for future studies to investigate how this is shaped by other aspects of psychological development, such as increases in perspective taking or mental state reasoning, emotional development and cognitive inhibition[42].

Equally important will be understanding the role of social environment, which has a crucial influence on prosocial behaviour in infancy[43], and may also affect prosocial behaviour later in childhood. For example, children's choices in costly sharing tasks become markedly more egalitarian if they had been exposed to civil warfare between the ages of 7 and 12, but not if the exposure occurred earlier in development (age 3–6)[44], and these effects seem to persist across the lifespan. Although the results of the present study are most informative about the development of children's responsiveness to normative information in personally costly cooperative dilemmas, it will also be important to explore how children's responsiveness develops differently across domains or contexts.

Future work should also explore other strategies for modelling the nature of social norms within communities. Our strategy was based on the estimated probability with which individuals in a society judged generous norms to be most correct, an approach that is similar to what has been used in previous studies. In a study of costly punishment in 15 societies, the decisions of individuals about whether to punish selfishness in third parties were predicted by the mean amount that members of their society gave in a DG[4]. Similarly, in a study of cooperation across camps of Hadza foragers, the contributions of individuals in a public goods game were predicted by the mean contribution of members of their camp in the same game[14]. This suggests that modelling norms using the frequency of a behaviour (or the probability of particular normative judgements) is an effective strategy, but other approaches may be even better, for example, a more conformist approach[6,45]. Future experiments should also explore the content of norms in other ways, for example, by eliciting judgements from participants about what others in their society do or expect them to do, or judgements about how similar the game is to real-world situations.

Our findings show that societal variation in prosocial behaviour is linked to beliefs about correct social norms. They also demonstrate that prosocial behaviour becomes increasingly adult-like and normative during middle childhood and that, during this same period, children across societies develop a tendency to respond (at a personal cost) to social norms about prosocial behaviour. In doing so, this project explains how the development of a universal norm psychology can lead to the emergence of societal variation in prosociality, and it adds to the growing evidence that the unique forms of cooperation in humans are highly dependent on acquired cultural norms and institutions.

Methods

All research and consent procedures were approved by the appropriate university ethical review boards at Arizona State University (IRB ID, STUDY00001591), Cambridge Psychology Research Ethics Committee (PRE.2016.026), Simon Fraser University Office of Research Ethics (study number, 2013s0335; study title, 'Prosocial development in Vanuatu and Canada'). All of the appropriate national and community bodies also gave consent for the research at all of our fieldsites. Research participants provided written informed consent for publication of the images in Fig. 1. Images of participants were obtained by experimenters from video recordings of the experimental trials. Research participants provided written informed consent for the video recordings.

Participants. Details about the participants are provided in Table 1 and Supplementary Table 1.

DG. Children participated in a binary choice DG in which the experimental participant decided between two predetermined payoff distributions, referred to below as ratios. Test ratio 1:1 for the participant, 1 for the recipient (that is, 1/1). Test ratio 2:2 for the participant, 0 for the recipient (that is, 2/0).

Apparatus and procedure. In cases in which between-participants conditions were used, participants were randomly assigned to conditions. Data were collected by fieldworkers familiar with the research design and hypotheses, and was therefore not blind. The apparatus consisted of two laminated paper trays, each with a red and a blue circle on them, which were placed in front of the participant (Supplementary Information, page 23). Each tray corresponded to one of the two DG test ratios, with tokens placed in the red circle going to the participant, and tokens placed in the blue circle going to the recipient. The recipient was not present during the study, but was represented by a small wooden person-shaped figurine. Recipients and participants were anonymous. The experimenter placed tokens on the trays, and the participant then selected one of the trays. Recipients were real, and rewards were delivered to them at a later time. For every choice during the study, different coloured figurines were used to indicate that the choices impacted a different recipient. The procedure was modified for the Hadza society owing to the practical need to shorten the study for participants; details of the full procedure and the modified Hadza procedure are provided in the Supplementary Information (pages 25, 28). All scripts were translated and then back-translated.

Comprehension checks. At the start of the study, participants demonstrated that they understood that a larger quantity of tokens would produce the most rewards, and all participants answered questions to indicate that when watching the videos they attended to the location of tokens, and remembered the content of the norm primes (for example, the experimenter pointed to 1/1 and asked 'is this right or wrong?', then pointed to 2/0 and asked 'is this right or wrong?'; Supplementary Information, pages 24, 25). No participants who passed these comprehension questions were excluded from the sample.

Rewards. Within the study, rewards were represented as tokens (such as glass beads and stones). The participants were informed that 'the more tokens they received, the more rewards they would receive', but the precise nature of the rewards or the exchange rate was not communicated to child participants. The exception to this was for the Hadza, in which the use of tokens was not understood by participants, and small sweets were used directly within the study in place of tokens (see the descriptions of the modified Hadza procedure below). For children, rewards were sourced locally, and usually consisted of sweets or small food items, or small items, such as stickers, glow-in-the-dark bracelets, or pens and pencils. Adult participants were in most cases told what the nature of the rewards would be (for example, money), and the general amount usually obtained by participants, but they also understood that the exact amount would be determined by their choices in the study. At one site (Pune), adults were not told what the reward would be, they were simply told that they would be obtaining 'prizes'. After the study, tokens were exchanged for rewards. This either occurred immediately for each participant, or it occurred later after each person had participated, with the rewards being distributed to all participants at the same time.

Statistical modelling approach. All data were binary choices taking the form of 0 (choice of 2/0) or 1 (choice of 1/1), so we modelled the choices of participants using regression with a binomial link function. For multilevel models, the posterior distribution of the model can be most easily estimated using Markov chain Monte Carlo. Using Markov chain Monte Carlo, we generated model predictions by processing many samples from the posterior distribution of the model. Each sample of parameter values from the posterior can be plugged into the model, producing a predicted value for any observable variable. As the distribution of the samples approximates the posterior distribution of the parameters, the distribution of predictions generated from a large number of samples will approximate the target predictive distribution. Examples of this approach have been described previously[46].

Data were analysed using the R Environment for Statistical Computing[47], with most of the specified models using the function 'MAP' (R package 'rethinking')[46]—a convenience tool for fitting a large number of different regression models. Multilevel models were specified and run using a variant of Hamiltonian Monte Carlo (an algorithm that works well with high dimension models) implemented in RStan[48]. Models were specified using weakly informative priors, which reduce overfitting and also help the Markov chain to converge to the posterior distribution more effectively than flat priors. The posterior distribution we present here is based on 5,000 samples from three chains (after 1,000 adaptation steps), for a total of 12,000 samples. These samples were sufficient to establish convergence to the target posterior distribution. We assessed convergence through the R-hat Gelman and Rubin statistic (R-hat values greater than 1.01 can indicate that the chain did not converge), and the effective number of samples for all of the parameters was substantial (an effective number of samples that is much smaller than the actual number of samples can suggest that the chain was not efficient). Further information about diagnosing chain convergence can be found in chapter 8 of ref. [46].

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data Availability

All data supporting the findings of this study are provided in the Supplementary Code and data.

Code Availability

All code supporting the findings of this study are provided in the Supplementary Code and data.

Received: 3 August 2018; Accepted: 18 August 2019;
Published online: 23 September 2019

References

1. Tomasello, M. *Why We Cooperate* (MIT Press, 2009).
2. Henrich, J. *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter* (Princeton Univ. Press, 2015).
3. Henrich, J. et al. Markets, religion, community size, and the evolution of fairness and punishment. *Science* **327**, 1480–1484 (2010).
4. Henrich, J. et al. Costly punishment across human societies. *Science* **312**, 1767–1770 (2006).
5. Richerson, P. et al. Cultural group selection plays an essential role in explaining human cooperation: a sketch of the evidence. *Behav. Brain Sci.* **39**, e30 (2016).
6. Chudek, M. & Henrich, J. Culture–gene coevolution, norm-psychology and the emergence of human prosociality. *Trends Cogn. Sci.* **15**, 218–226 (2011).
7. Barrett, H. C. *The Shape of Thought: How Mental Adaptations Evolve* (Oxford Univ. Press, 2014).
8. Bicchieri, C. *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms* (Oxford Univ. Press, 2016).
9. Keller, H. *Cultures of Infancy* (Psychology Press, 2013).
10. Rutland, A., Killen, M. & Abrams, D. A new social-cognitive developmental perspective on prejudice: the interplay between morality and group identity. *Perspect. Psychol. Sci.* **5**, 279–291 (2010).
11. Tomasello, M. & Vaish, A. Origins of human cooperation and morality. *Annu. Rev. Psychol.* **64**, 231–255 (2013).
12. Opp, K.-D. & Hechter, M. *Social Norms* (Russell Sage Foundation, 2001).
13. Bicchieri, C. *The Grammar of Society: The Nature and Dynamics of Social Norms* (Cambridge Univ. Press, 2006).

14. Smith, K. M., Larroucau, T., Mabulla, I. A. & Apicella, C. L. Hunter-gatherers maintain assortativity in cooperation despite high levels of residential change and mixing. *Curr. Biol.* **28**, 3152–3157 (2018).
15. House, B. R. et al. Ontogeny of prosocial behavior across diverse societies. *Proc. Natl Acad. Sci. USA* **110**, 14586–14591 (2013).
16. Rakoczy, H. & Schmidt, M. F. H. The early ontogeny of social norms. *Child Dev. Perspect.* **7**, 17–21 (2013).
17. Schmidt, M. F. H., Rakoczy, H. & Tomasello, M. Young children enforce social norms selectively depending on the violator's group affiliation. *Cognition* **124**, 325–333 (2012).
18. House, B. R. & Tomasello, M. Modeling social norms increasingly influences costly sharing in middle childhood. *J. Exp. Child Psychol.* **171**, 84–98 (2018).
19. McAuliffe, K., Raihani, N. J. & Dunham, Y. Children are sensitive to norms of giving. *Cognition* **167**, 151–159 (2017).
20. Smetana, J. G. & Braeges, J. L. The development of toddlers' moral and conventional judgments. *Merrill-Palmer Q.* **36**, 329–346 (1990).
21. Blake, P. R. et al. The ontogeny of fairness in seven societies. *Nature* **528**, 258–261 (2015).
22. Rochat, P. et al. Fairness in distributive justice by 3- and 5-year-olds across seven cultures. *J. Cross Cult. Psychol.* **40**, 416–442 (2009).
23. House, B. R. Diverse ontogenies of reciprocal and prosocial behavior: cooperative development in Fiji and the United States. *Dev. Sci.* **20**, e12466 (2017).
24. Cowell, J. M. et al. The development of generosity and moral cognition across five cultures. *Dev. Sci.* **20**, e12403 (2017).
25. Engel, C. Dictator games: a meta study. *Exp. Econ.* **14**, 583–610 (2011).
26. Fehr, E., Bernhard, H. & Rockenbach, B. Egalitarianism in young children. *Nature* **454**, 1079–1083 (2008).
27. House, B. R., Henrich, J., Brosnan, S. F. & Silk, J. B. The ontogeny of human prosociality: behavioral experiments with children aged 3 to 8. *Evol. Hum. Behav.* **33**, 291–308 (2012).
28. Robbins, J. & Rumsey, A. Introduction: cultural and linguistic anthropology and the opacity of other minds. *Anthropol. Q.* **81**, 407–420 (2008).
29. Purzycki, B. G. et al. Moralistic gods, supernatural punishment and the expansion of human sociality. *Nature* **530**, 327–330 (2016).
30. Leimgruber, K. L., Shaw, A., Santos, L. R. & Olson, K. R. Young children are more generous when others are aware of their actions. *PLoS One* **7**, e48292 (2012).
31. Engelmann, J. M., Herrmann, E. & Tomasello, M. Five-year olds, but not chimpanzees, attempt to manage their reputations. *PLoS One* **7**, e48433 (2012).
32. Engelmann, J. M. & Rapp, D. J. The influence of reputational concerns on children's prosociality. *Curr. Opin. Psychol.* **20**, 92–95 (2018).
33. Decety, J. & Svetlova, M. Putting together phylogenetic and ontogenetic perspectives on empathy. *Dev. Cogn. Neurosci.* **2**, 1–24 (2012).
34. Paulus, M. The emergence of prosocial behavior: why do infants and toddlers help, comfort, and share? *Child Dev. Perspect.* **8**, 77–81 (2014).
35. House, B. R., Henrich, J., Sarnecka, B. & Silk, J. B. The development of contingent reciprocity in children. *Evol. Hum. Behav.* **34**, 86–93 (2013).
36. Sebastián-Enesco, C. & Warneken, F. The shadow of the future: 5-Year-olds, but not 3-year-olds, adjust their sharing in anticipation of reciprocation. *J. Exp. Child Psychol.* **129**, 40–54 (2015).
37. Warneken, F. & Tomasello, M. The emergence of contingent reciprocity in young children. *J. Exp. Child Psychol.* **116**, 338–350 (2013).
38. Olson, K. R. & Spelke, E. S. Foundations of cooperation in young children. *Cognition* **108**, 222–231 (2008).
39. Over, H. The influence of group membership on young children's prosocial behaviour. *Curr. Opin. Psychol.* **20**, 17–20 (2018).
40. Hamlin, J. K. Moral judgment and action in preverbal infants and toddlers: evidence for an innate moral core. *Curr. Dir. Psychol. Sci.* **22**, 186–193 (2013).
41. Hepach, R. & Warneken, F. Editorial overview: early development of prosocial behavior: revealing the foundation of human prosociality. *Curr. Opin. Psychol.* **20**, iv–viii (2018).
42. House, B. R. How do social norms influence prosocial development? *Curr. Opin. Psychol.* **20**, 87–91 (2018).
43. Kärtner, J. Beyond dichotomies—(m)others' structuring and the development of toddlers' prosocial behavior across cultures. *Curr. Opin. Psychol.* **20**, 6–10 (2018).
44. Bauer, M., Cassar, A., Chytilová, J. & Henrich, J. War's enduring effects on the development of egalitarian motivations and in-group biases. *Psychol. Sci.* **25**, 47–57 (2014).
45. Henrich, J. & Boyd, R. The evolution of conformist transmission and the emergence of between-group differences. *Evol. Hum. Behav.* **19**, 215–241 (1998).
46. McElreath, R. *Statistical Rethinking* (Chapman and Hall/CRC, 2015).
47. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2018).
48. Stan Development Team. *RStan: The R Interface to Stan.* http://mc-stan.org/ (2018).

Acknowledgements
We thank C. Ross for assistance with statistical methods, and the staff of the Institute of Human Origins at Arizona State University for support. This research was funded by a grant (grant no. 48952) from the John Templeton Foundation to the Institute of Human Origins, Arizona State University. The opinions expressed in this publication do not necessarily reflect the views of the John Templeton Foundation. This research was also funded by a Freigeist Fellowship from the Volkswagen Foundation to P.K. (grant no. 89611) and a SSHRC Doctoral Scholarship to S.L.-L. (award no. 752-2016-0555). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions
B.R.H. and J.B.S. conceived the project and designed the study. P.K., H.C.B., T.B., A.E., S.L.-L., C.S.-E. and A.M.S. also contributed to study design. B.R.H., P.K., H.C.B., T.B., S.C., A.E., S.L.-L., C.S.-E., A.M.S., S.Y. and A.N.C. collected data. B.R.H. analysed the data. B.R.H. and J.B.S. wrote the manuscript and P.K. and H.C.B. also contributed substantially to writing. All of the authors contributed to writing the Supplementary Information.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41562-019-0734-z.

Correspondence and requests for materials should be addressed to B.R.H.

Peer review information Primary Handling Editor: Stavroula Kousta.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): Bailey House

Last updated by author(s): 18/07/2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	ONA Collect
Data analysis	R environment for statistical computing, RStan

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The authors declare that the data supporting the findings of this study are available within the paper's supplementary information files.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Quantitative; a cross cultural study of children's and adult's choices in a binary choice experimental task, measuring cooperative behavior.
Research sample	Children and adults in 8 different populations, both rural and urban, Western and non-Western (United States, Germany, Argentina, India, Ecuador, Tanzania, Vanuatu). The sites were chosen from those already studied by collaborators, to provide a diverse subset of human societies.
Sampling strategy	Our sampling strategy was based on convenience, participants were selected by availability in the locations where we collected data. Minimum sample sizes were chosen based on the smallest usable samples that from prior studies.
Data collection	Data was collected on a tablet computer, with forms set up using ONA Collect. Usually, no one was present besides the participant, the experimenter, and when appropriate a translator. The researcher was not blind to the experimental conditions.
Timing	Data collection began in September 2016 and completed in October 2017.
Data exclusions	We excluded data from participants who were not able to pass comprehension questions about numerical quantities, or who weren't able to report the content of videos shown to them as part of experimental manipulations. These criteria were conservative and set in advance, and resulted in 52 children (.057%) and 7 adults (.027%) being excluded.
Non-participation	Very few participants declined participation once the study session had begun, but exact drop-out rates for all field populations are not available. There were the standard field problems with faulty equipment, and inconsistent or unreliable data about age, which led to 34 participants' data being lost.
Randomization	Participants were allocated to experimental conditions based on a scheme of ensuring that relatively equal numbers of children of each age were included in each condition. This necessarily wasn't entirely random, but the only criteria employed was equalizing age distributions across conditions.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study	
☒	☐	Antibodies
☒	☐	Eukaryotic cell lines
☒	☐	Palaeontology
☒	☐	Animals and other organisms
☐	☒	Human research participants
☒	☐	Clinical data

Methods

n/a	Involved in the study	
☒	☐	ChIP-seq
☒	☐	Flow cytometry
☒	☐	MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	Both male and female participants aged 4-5 to adulthood (sample goes up to ages in the 60s).
Recruitment	Participants were recruited by being approached at a museum or coffee shop or at their place of residence in their village, or by sending home flyers/consent forms through their school.
Ethics oversight	Arizona State University, University of Cambridge, Simon Fraser University

Note that full information on the approval of the study protocol must also be provided in the manuscript.