

Title: Restorative justice in children

Authors: Katrin Riedl¹, Keith Jensen^{2*}, Josep Call^{1,3} & Michael Tomasello¹

Author affiliations:

¹ Max Planck Institute for Evolutionary Anthropology, Department of Developmental and Comparative Psychology, Deutscher Platz 6, D-04103, Leipzig, Germany

² University of Manchester, School of Psychological Sciences, Coupland 1 Building, Coupland Street, Oxford Road, Manchester, M13 9PL, United Kingdom

³ University of St Andrews, School of Psychology & Neuroscience, St. Andrews, Fife, KY16 9JU, United Kingdom

* corresponding author: School of Psychological Sciences, Coupland 1 Building, Coupland Street, Oxford Road, Manchester, UK, M13 9PL, UK, phone: +44 161 306 0423, email: keith.jensen@manchester.ac.uk

Summary

An important, and perhaps uniquely human, mechanism for maintaining cooperation against free-riders is third-party punishment [1, 2]. Our closest living relatives, chimpanzees, will not punish third parties even though they will do so when personally affected [3]. Until recently, little attention has been paid to how punishment and a sense of justice develop in children. Children respond to norm violations [4], they are more likely to share with a puppet that helped another individual as opposed to one who behaved harmfully and they show a preference for seeing a harmful doll rather than a victim punished [5]. By 6 years of age, children will pay a cost to punish fictional and real peers [6-8] and the threat of punishment will lead preschoolers to behave more generously [9]. However, little is known about what motivates a sense of justice in children. We gave 3- and 5-year-old children – the youngest ages yet tested – the opportunity to remove items and prevent a puppet from gaining a reward for 2nd and 3rd party violations (Experiment 1), and, we gave 3-year-olds the opportunity to restore items (Experiment 2). Children were as likely to engage in third-party interventions as they were when personally affected, yet they did not discriminate amongst the different sources of harm for the victim. When given a range of options, 3-year-olds chose restoration over removal. It appears that a sense of justice centred on harm caused to victims emerges early in childhood and highlights the value of third-party interventions for human cooperation.

Keywords

cooperation, punishment, justice, moral behaviour

Results and Discussion

To explore the emergence and early development of third-party interventions in the context of distributive justice, we tested 3- and 5-year-old children using an action-based paradigm that had been applied to chimpanzees [3, 10]. The question was whether children would selectively “punish” another individual (namely as negative reciprocity or to impose a cost to decrease future occurrences of a behaviour [11, 12]) and whether they would do so on behalf of others. Inflicting costs on others need not always be punitive: adults and children will even suffer a cost themselves to reduce another’s welfare out of a sense of fairness and even spite [13-17]. We contrasted intentional harm (theft) as a measure of punishment, with unfair outcomes (spitefulness), loss (frustration), and permitted taking (impulsive pulling). Based on previous studies, even young infants have expectations when observing reward distributions [18] and will act on these [19]. Infants will also show preferences for helpful as opposed to harmful figures [20] unless the harm is warranted [21]. As well as allowing children to respond (i.e., pull a rope causing the table to turn) in second-party conditions and intervene in third-party conditions (Experiment 1), we also allowed the 3-year-olds to give and remove rewards freely (Experiment 2). We predicted that children would respond less often as the violation diminished, namely more often in theft than unfair than loss than permitted-taking, and that they would do so more often when personally affected than when witnessing a violation.

In both Experiments 1 and 2, children were tested with a large turntable and puppet characters (Figure S1). The turntable was divided into quarters that demarcated the child's position, that of the "victim" (to the child's left), the "thief" (across from the child) and an inaccessible area called the "cave" (to the child's right). Puppets played the roles of "victim" and "thief" as well as a "stranger". The table could be turned by pulling ropes underneath it. These were only in the child's position and the "thief's"; the purpose of this was to make it clear to the child that the puppet in the victim position could not turn the table. In Experiment 1, the child had a single rope, allowing the table to be turned clockwise only; once pulled to the cave position, the apparatus was locked, preventing further movement. Children and the puppets would play with toys or eat cookies that would be present on only one quadrant of the table. Children were assigned to one of four between-subjects treatments – theft, unfairness, loss and permitted-taking – according to how the objects were taken and whom benefitted, and they participated in both 3rd and 2nd party within-subjects conditions (3 trials of each, order counterbalanced between subjects). Figure 1 presents a schematic diagram of the different treatments and conditions.

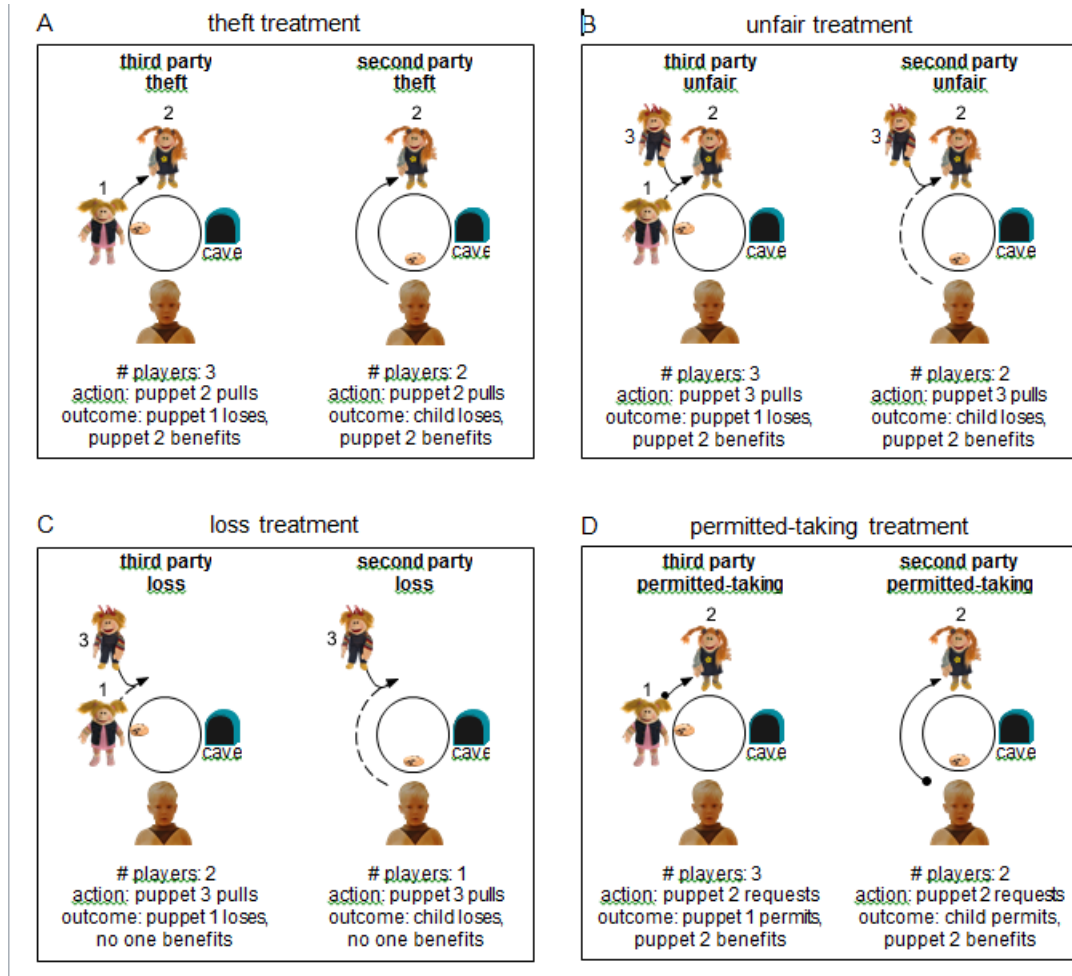


Figure 1. Illustration of (A) theft, (B) unfair, (C) loss and (D) permitted-taking treatments for second party and third party conditions. The starting positions of the objects on the turntable – either in front of the child or puppet 1 in the victim’s position – are shown. The arrows show the movement of the turntable caused by puppet 2 in the thief’s position (solid lines) or the stranger puppet 3 (dashed lines). The dot at the beginning of the arrow in the permitted-taking treatments indicates that the puppet or child allowed the objects to be moved away. See also Figure S1.

In Experiment 1, 3- and 5-year-old children could intervene when witnessing a third individual similarly affected as well as responding when personally affected. The children either witnessed goods being taken away from a puppet (3rd party – 3P) or had the goods taken away from them (2nd party – 2P); there was no effect of the order in which the conditions were presented (3yo $P = 0.973$; 5yo $P = 0.628$). The only action available to the children was to move objects on the table away from the thief’s position to the cave. Both 3-

and 5-year-olds turned the table as often into the cave in 3P as in 2P (3yo $T^+ = 40$, $n = 48$ (34 ties), $P = 0.448$; 5yo $T^+ = 68$, $n = 72$ (58 ties), $P = 0.319$; Figure 2). There was no difference between the paired-conditions including 2P-theft and 3P-theft (3yo and 5yo, $P > 0.25$; Table S1). We then examined the between-subjects treatments for the two conditions separately (e.g., 2P theft vs. 2P permitted-taking). When children were directly affected by the actions of puppets (2P), 5-year-olds put the objects into the cave more often in the theft treatment than in the permitted-taking treatment ($P < 0.001$), whereas 3-year-olds tended to do so ($P = 0.089$); Table S2. Furthermore, 5-year-olds ($P = 0.041$), but not 3-year-olds ($P = 0.162$), moved objects into the cave more often in response to 2P theft than 2P loss. Both 3-year-olds ($P = 0.005$) and 5-year-olds ($P = 0.015$) were significantly more likely to respond to 3P theft than 3P permitted-taking and both groups of children would make objects inaccessible when another individual lost them, even when the puppet who benefitted was not responsible for taking them (3P unfair, $P > 0.475$) or if no one else received them at all (3P loss, $P > 0.723$). In other words, the children moved objects in the unfair, loss and theft treatments and would do so as much for the sake of the victim as for themselves. The fact that they were as likely to intervene when witnessing another individual's loss – regardless of the cause – suggests that they focussed on the consequences for the victim rather than on the benefits or intentions of the thief (when present). Additionally, they protested and tattled in all of the treatments, not solely in response to theft, nor just when personally affected (see Supplemental Data).

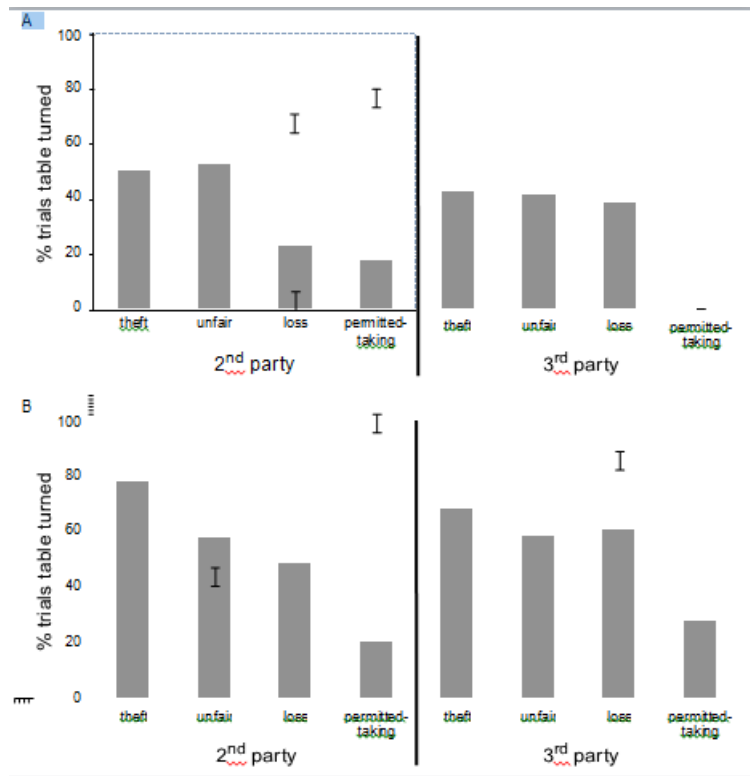


Figure 2. Proportion of trials in which (A) 3-year-old children and (B) 5-year-old children turned the table to move objects from the thief's position to an inaccessible cave. Values are expressed as means and error bars as bootstrapped 95% confidence intervals.

In Experiment 2, we wanted to determine whether children would prefer to restore items by returning them to the original possessor or intervene by making them inaccessible. As in Experiment 1, 3-year-old children and a puppet either had items taken away from them; the key difference was that children had access to two rope ends, allowing them to turn the table freely in either direction. In addition to allowing the children to move objects into the cave, they could take the objects for themselves or move them to the victim's position. There was again no effect of order for 2P and 3P ($P = 0.433$). Children were more likely to pull objects away from the thief position in the second-party conditions than the third-party conditions (69% and 58% respectively; $T^+ = 43.50$, $n = 48$ (29 ties), $P = 0.031$), but the only difference was a trend for children to turn the table more in 2P theft than 3P theft ($P = 0.063$; all other $P \geq 0.438$; Figure 3; Table S3). Across second party treatments, they were more

likely to turn the table when the thief took the objects away from them than when they consented to having them taken (2P theft vs. 2P permitted-taking; $P = 0.014$; Table S4). There was no difference, however, between 3P theft and 3P permitted-taking, or any of the other treatments (all $P > 0.10$). The predominant response was to return the objects to the original owner, though they were more likely to do this for themselves in the 2P trials than for the victim in 3P trials ($T^+ = 47.00$, $n = 48$ (25 ties), $P = 0.004$). They restored the objects more often when personally affected by theft than when a third-party was affected ($P = 0.016$) and they showed a tendency to do so when consent was given (2P permitted-taking vs. 3P permitted-taking; $P = 0.094$; Table S5). There was no difference in restoration between 2P unfair and 3P unfair, nor between 2P loss and 3P loss (all $P \geq 0.5$). It is worth noting that stealing (pulling to self) was uncommon in all but the permitted-taking treatment: this occurred only three times in 3P theft, twice in 3P unfair and seven times in 3P permitted-taking). Children predominantly chose the "self" option in the 3P permitted-taking treatment; this was the only situation in which objects were not restored to the original owners. Removal – moving the object into the cave where no one could get the goods – was the most infrequent choice, occurring only twice in the 2P condition and only once in the 3P condition. As in experiment 1, the children protested and tattled across treatments in in 3P as well as 2P (see Supplemental Data).



Figure 3. Frequency with which children turned the table to the different positions in both 2P and 3P conditions. Children could either move the objects to themselves (white bars), to the victim (grey bars) or to the cave (black bars). Note that percentages do not add up to 100%: trials in which the children did not pull the table are also included.

Three-year-old children pulled the rope causing the table to turn at a very high rate in all treatments with the exception of the permitted-taking treatment. They were far more likely to return items to the original owner – either themselves (2P) or the victim puppet (3P) – than they were to do nothing, make the items inaccessible, give them away or steal them for themselves. The children were more likely to return things to themselves in the second-party conditions than in the third-party, but they still reacted at a surprisingly high rate when they were not directly affected. The only group of children that did not turn the table as often as the others were those in the permitted-taking treatment, namely when the puppet in the victim’s position or the child gave permission to the third puppet to take the objects away. The children in this study appeared to focus on how the outcomes affected the original owners and less, perhaps, on the consequences for the puppet who only secondarily received them.

Discussion

The first experiment demonstrated that both 3- and 5-year-old children will intervene against third-party violations, and they will do so as much as when personally affected. Three-year-olds did not appear to punish theft, in that they were as likely to respond to unfairness and loss; 5-year-olds did not treat an unfair outcome differently than theft, but they did punish theft in that they responded more often than when there was no thief (loss), but only when personally affected. Overall, the children appeared to be responding more to the consequences for the victim, rather than for the act itself or the outcomes for the beneficiary (in the loss condition, there were none). Spiteful motivations might have explained responses to unfairness, and for loss, children might have been responding out of frustration (2nd party) and empathic concern (3rd party). One could speculate on other motives, such as punishing for the sake of reputation [22], a point that can be addressed in future studies. In the second experiment, 3-year-olds did not show a taste for removing rewards, but preferred to restore objects to original owners. Children did not match theft with theft unless the “victim” willingly gave up objects, setting up a turn-taking game; they restored objects to the victim as much as to themselves. They enforced consequences as often in the third-party as in the second-party conditions (although there was a tendency to be more self-regarding), and as in Experiment 1, they did not discriminate between theft, unfairness and loss.

At the preschool age, children do not appear to inflict harm on others out of a sense of justice, deterrence or revenge (e.g., [23]), but out of a concern for the welfare of the victim. This is especially striking in that they do so as much for another individual as for themselves even before the milestone age at which they demonstrate perspective-taking on the basis of false-belief understanding (theory of mind [24]). The third-party responses of the children are likely due to a combination of affective perspective-taking, namely responding to the distress

of the victims [25, 26], and to an already established norm of ownership [27]. In our studies, children intervened when the victim was present because the victim had no recourse for action; only the child was in the position to act. It is possible that the protests of the victims cued the children to act, just as expression of distress elicits looks of concern [26].

The two studies presented here are the clearest – and earliest – demonstrations of third-party interventions in young children, and the first to attempt to disentangle punishment from other third-party motivations. Whether children use these interventions as deterrence or as a form of just deserts is a question that requires further work; at least by the age of 7, children do respond to the threat of punishment by behaving more cooperatively [9] and they will even pay a cost to impose a cost on a norm violator [10]. As well, it would be important to address the role of signalling on part of the victim; children of the age of 18 months will show concern for others even in the absence of emotional cues [25], but requests at this age are important for eliciting sharing [28]; furthermore, children might signal their role as enforcers [22]. To determine whether concern for victims motivates punishments is a universal part of a child's development, future studies could examine children in other cultures, since adults in different societies show patterns of third-party and altruistic punishment [29-31]. What is clear is that already by 3 years of age, children are capable of intervening on the behalf of others, quite unlike our closest living relatives tested in a comparable situation [3]. It appears that in humans, intervening on the behalf of others begins with a concern for the victim before becoming focussed on consequences for the perpetrator.

Experimental Procedures

Experiment 1

We tested 58 3-year-old children (age range = 3;3-3;9, mean = 3;6, 24 boys) and 79 5-year-olds (age range = 5;3-5;9, mean = 5;6, 36 boys) in a mid-sized German city. Ten 3-year-olds were excluded because they showed signs of distress and did not complete testing. Seven 5-year-olds were excluded due to experimenter error (one), unwillingness to participate in the test (three) and interference with the test apparatus (three). A pilot study on adults showed that they understood the apparatus and the procedure and interpreted the conditions appropriately. They were brought to the testing room individually by an experimenter (moderator) who introduced them to the two puppets who were controlled by two other experimenters. They were then familiarised with the apparatus by first being asked to pull the rope to observe the consequences, namely that the table could turn in only one direction and would then become locked at the inaccessible cave where objects could not be recovered. They then used the table and toys with the puppets. The experimenter then asked the child to move the object after the last puppet played with it so that it would come to rest in an inaccessible cave. This was done so that the child would know that he or she could prevent the puppets from accessing the objects, and that the puppets would also know this (in effect, “trashing” the objects [32]). Normative terms such as “punish” were not used and none of the puppets protested, implying consent. Children were assigned to one of four treatments – theft, unfairness, loss and permitted-taking – according to how the objects were taken and whom benefitted (between subjects), and they participated in both 3rd and 2nd party conditions (within subjects, 3 trials of each, counterbalanced for order in a blocked design between subjects). In “theft”, the thief puppet pulled the rope, moving the turntable and claiming the objects (marble game, stamps or cookies) on it. In 2P, she pulled the objects from the child; in 3P, these were taken from the victim puppet. In “unfair”, another puppet (stranger) would enter the room and turn the objects from the child to the thief puppet (2P) or from the victim puppet to the thief puppet (unfair). The “loss” treatment was similar, except that there was no

one in the thief's position. Finally, in the "permitted-taking" treatment, the child or the victim puppet consented to the "thief" pulling the rope to take the objects (Figure 1). Children could pull the rope, causing the objects to move from the thief's position to the cave, or do nothing. Trials ended after 1 minute if they did nothing, or when the objects came to rest in the cave. The puppets never communicated with or made direct "eye-contact" with the children (nor did the experimenters controlling the puppets), and protests by the puppet in the victim's position were never directed at the children. None of the puppets or experimenters in any way commented on the choices made by the children.

Experiment 2

Participants were 54 3-year-olds (age range = 3;3-3;9, mean = 3;5, 30 boys). None of the children had participated in Experiment 1. Six children were excluded from the analysis due to experimenter error (one), fear of the puppets (two), failing to meet criteria in familiarisation (one), and feeling uncomfortable with the test (two). Children were randomly assigned to one of the four treatments as before, and given both 2P and 3P conditions. They were introduced to an apparatus that had two ropes that they could pull, and the table could move freely. Importantly, the table would not stop at the cave position; if children wanted the objects to stop in that position, they had to do so manually. Trials lasted 1 minute if the children did nothing, or ended once the child had moved the apparatus and stopped it in one of the four positions. All trials in both experiments were videotaped and coded for reliability (choice: Cohen's $\kappa = 1.000$).

Supplemental Information

Supplemental Information includes Supplemental Data, Supplemental Experimental Procedures, 1 figure and 5 tables.

Acknowledgements

We would like to thank Angela Loose research assistance, Georg Keller, Monique Horstmann, Nadin Bobovnikov, Marte Krüger and Diana Paschenda for their involvement with the experiments, Manfred Ulrich for building the apparatus, Roger Mundry for statistical advice, Eric Jensen for photography, Petra Jahn and Marlen Sureck for multimedia support 3D-imaging, and the children and their parents for participating in these studies.

References

1. Boyd, R., Gintis, H., Bowles, S., and Richerson, P.J. (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 3531-3535.
2. Fehr, E., and Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior* **25**, 63-87.
3. Riedl, K., Jensen, K., Call, J., and Tomasello, M. (2012). No third-party punishment in chimpanzees. *Proceedings of the National Academy of Sciences* **109**, 14824–14829.
4. Rakoczy, H., Warneken, F., and Tomasello, M. (2008). The sources of normativity: Young children's awareness of the normative structure of games. *Developmental Psychology* **44**, 875-881.
5. Kenward, B., and Dahl, M. (2011). Preschoolers distribute resources according to recipients' moral status. *Developmental Psychology* **47**, 1054-1064.
6. Jordan, J.J., McAuliffe, K., and Warneken, F. (2014). Development of in-group favoritism in children's third-party punishment of selfishness. *Proceedings of the National Academy of Sciences Early Edition*, 1-6.
7. McAuliffe, K., Jordan, J.J., and Warneken, F. (2014). Costly third-party punishment in young children. *Cognition* **134C**, 1-10.
8. Salali, G.D., Juda, M., and Henrich, J. (2014). Transmission and development of costly punishment in children. *Evolution and Human Behavior*.
9. Lergetporer, P., Angerer, S., Glatzle-Rutzler, D., and Sutter, M. (2014). Third-party punishment increases cooperation in children through (misaligned) expectations and conditional cooperation. *Proceedings of the National Academy of Sciences* **111**, 6916-6921.
10. Riedl, K., Jensen, K., Call, J., and Tomasello, M. (2011). Restorative justice and third party punishment in preschool children. In Society for Research in Child Development. (Montreal).
11. Jensen, K., and Tomasello, M. (2010). Punishment. In Encyclopedia of Animal Behavior, M.D. Breed and J. Moore, eds. (Oxford: Academic Press), pp. 800-805.
12. Clutton-Brock, T.H., and Parker, G.A. (1995). Punishment in animal societies. *Nature* **373**, 209-216.
13. Jensen, K. (2010). Punishment and spite, the dark side of cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 2635-2650.
14. McAuliffe, K., Blake, P.R., and Warneken, F. (2014). Children reject inequity out of spite. *Biol Lett* **10**.
15. Wittig, M., Jensen, K., and Tomasello, M. (2013). Five-year-olds understand fair as equal in a mini-ultimatum game. *Journal of Experimental Child Psychology* **116**, 324-337.
16. Zizzo, D.J., and Oswald, A.J. (2001). Are people willing to pay to reduce others' incomes? *Annales d'économie et de statistique* **63-64**, 39-65.
17. Falk, A., Fehr, E., and Fischbacher, U. (2003). On the nature of fair behavior. *Economic Inquiry* **41**, 20-26.
18. Sloane, S., Baillargeon, R., and Premack, D. (2012). Do infants have a sense of fairness? *Psychological Science* **23**, 196-204.
19. Schmidt, M.F.H., and Sommerville, J.A. (2011). Fairness expectations and altruistic sharing in 15-month-old human infants. *PLoS ONE* **6**, e23223.

20. Hamlin, J.K., Wynn, K., and Bloom, P. (2007). Social evaluation by preverbal infants. *Nature* **450**, 557-559.
21. Hamlin, J.K., Wynn, K., Bloom, P., and Mahajan, N. (2011). How infants and toddlers react to antisocial others. *Proceedings of the National Academy of Sciences* **108**, 19931–19936.
22. Shaw, A., Montinari, N., Piovesan, M., Olson, K. R., Gino, F., & Norton, M. I. (2014). Children develop a veil of fairness. *Journal of Experimental Psychology: General* **143**, 363-375.
23. McCullough, M.E., Kurzban, R., and Tabak, B.A. (2013). Cognitive systems for revenge and forgiveness. *Behavioral and Brain Sciences* **36**, 1-15.
24. Doherty, M.J. (2008). *Theory of Mind: How Children Understand Others' Thoughts and Feelings*, (Hove, UK: Psychology Press).
25. Vaish, A., Carpenter, M., and Tomasello, M. (2009). Sympathy through affective perspective taking and its relation to prosocial behavior in toddlers. *Developmental Psychology* **45**, 534-543.
26. Hobson, J.A., Harris, R., García-Pérez, R., and Hobson, P. (2009). Anticipatory concern: A study in autism. *Developmental Science* **12**, 249-263.
27. Rossano, F., Rakoczy, H., and Tomasello, M. (2011). Young children's understanding of violations of property rights. *Cognition* **121**, 219-227.
28. Brownell, C.A., Svetlova, M., and Nichols, S. (2009). To share or not to share: When do toddlers respond to another's needs? *Infancy* **14**, 117-130.
29. Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J.C., Gurven, M., Gwako, E., Henrich, N., et al. (2006). Costly punishment across human societies. *Science* **312**, 1767-1770.
30. Marlowe, F.W., and Berbesque, J.C. (2008). More 'altruistic' punishment in larger societies. *Proceedings of the Royal Society of London B Biological Sciences* **275**, 587-590.
31. Herrmann, B., Thöni, C., and Gächter, S. (2008). Antisocial punishment across societies. *Science* **319**, 1362-1367.
32. Shaw, A., and Olson, K.R. (2012). Children discard a resource to avoid inequity. *Journal of Experimental Psychology: General* **141**, 382-395.