Preschool children and chimpanzees incur costs to watch punishment of antisocial others

Natacha Mendes[1*], Nikolaus Steinbeis[2,3,4*], Nereida Bueno-Guerra[5,6], Josep Call[5,7], Tania Singer [2]

1 Max Planck Institute for Human Cognitive and Brain Sciences, Research Group Neuroanatomy & Connectivity, Stephanstraße 1A, 04103 Leipzig
2 Max Planck Institute for Human Cognitive and Brain Sciences, Department of Social Neuroscience, Stephanstraße 1A, 04103 Leipzig
3 Institute of Psychology, Leiden University, 2333 AK Leiden, The Netherlands
4 University College London, Department of Clinical, Educatinal and Health Psychology, London WC1H OAP
5 Max Planck Institute for Evolutionary Anthropology, Department of Developmental and Comparative Psychology, Deutscher Platz 6, 04103 Leipzig
6 University of Barcelona, Department of Psychology and Clinical Psychobiology, Passeig de Vall d'Hebrón 171, 08035 Barcelona
7 University of St Andrews, School of Psychology and Neuroscience, St Andrews, Fife KY16 9JP

* These authors contributed equally to the work

Corresponding author
Nikolaus Steinbeis
University College London, Department of Clinical, Educational and Health Psychology, London WC1H OAP
Email: n.steinbeis@ucl.ac.uk

Word Count: 5695

Keywords: prosocial, antisocial, motivation to watch deserved punishment, chimpanzees, children

Abstract

When misfortune befalls another, humans may feel distress, leading to a motivation to escape. When such misfortune is perceived as justified however it may be experienced as rewarding and lead to a motivation to witness such misfortune. We explored when in human ontogeny such a motivation emerges and if such a motivation is shared by chimpanzees. Chimpanzees and 4-6 year old children learned through direct interaction that an agent was either prosocial or antisocial and later saw each agent's punishment with the option to invest physical effort (chimpanzees) or monetary units (children) to continue watching. Chimpanzees and 6-year olds showed a preference for watching punishment of the antisocial agent. An additional control experiment in chimpanzees suggests that these results cannot be attributed to more generic factors such as scene coherence or informational value seeking. This indicates that both 6-year-olds and chimpanzees have a motivation to watch deserved punishment enacted.

Word Count: 149

How cooperation in societies can emerge and be maintained remains an evolutionary puzzle[1-5]. Punishment of antisocial group members is arguably one key mechanism capable of ensuring that levels of cooperation remain high in human[6] as well as other species[7,8]. It has been shown that the experience of emotions is a likely proximate cause that sustains cooperation and motivates costly punishment of antisocial others in humans[4,9-11]. Seeing others suffer can induce emotional states such as empathic distress[12] or concern[13], of which the latter is a powerful motivator for altruistic helping[10,12,14]. Along with humans, several other animal species have been tested for reactions to witnessing pain in conspecifics[15-24], providing some evidence for at least some forms of empathic responding. It has been shown in humans that empathic reactions can be radically undermined and change to feelings of pleasure, when the suffering victim was previously antisocial or perceived as an outgroup member[10,11]. Such signals of reward have been shown to be critical predictors of a subsequent absence of helping and desire for revenge and punishment[10,11]. Thus, young human infants display an early preference for prosocial compared to antisocial agents[25,26] and prefer those who are antisocial to previously antisocial others[27]. Further, preschoolers have been shown to endorse the misfortune of competitors[28,29], to think antisocial others as deserving of punishment[30] and to punish transgressions of outgroup members more than those of ingroup members[31]. Much less is known about how such mechanisms might operate in one of our closest living relatives, the chimpanzee (but see [32-35]). While it is known that chimpanzees appear to develop attitudes towards others based on previous pro- and antisocial behaviors[36-38], nothing is known about the phylogenetic origins of the motivation to watch the enactment of revenge.

We used a cross-species forced-choice behavioral paradigm to study whether chimpanzees and children aged 4-6 years differentially incur costs to continue watching the punishment of agents depending on whether these had been pro- or antisocial in a directly experienced previous interaction with them (Studies 1 and 2). The pro- or antisocial nature of the agents was operationalized by means of them offering valuable goods to children (i.e. their favorite toys) and chimpanzees (i.e. food). Whereas the prosocial agent would both offer and give the goods to the participant, the antisocial agent would offer the goods first but then withdraw the goods. The punishment procedure for all the studies entailed a punisher applying physical punishment in the form of hitting each of the two agents (i.e. either prosocial or antisocial; Figure 1A and 1B). Crucially, after a brief period of witnessing the punishment, this was rendered invisible to subjects (i.e. occurred in another part of the room for chimpanzees / was occluded by a curtain of a puppet theatre for children). Therefore to continue watching the punishment subjects had to

incur costs, which for chimpanzees entailed physical effort by operating a heavy sliding door to get to the invisible part of the room (Figure 1A) and for children entailed paying tokens or monetary units (henceforth MUs) for the curtain of the puppet theatre to be raised again (Figure 1B). As indicators of a motivation to witness punishment we used the amount of cost incurred to continue watching the punishment. We operationalized cost incurred as the expenditure of valuable monetary units (MUs) for children and physical energy and time for chimpanzees. We predicted that both chimpanzees (Study 1) and children (Study 2) would be more motivated to watch the punishment of the antisocial compared to the prosocial agent. We also predicted signs of greater positive emotions during the initial punishment of the two antisocial agent compared to the prosocial agent for the children. To measure emotional correlates, we scored facial expressions (e.g., smiles, frowns) during the punishment of the two agents. In chimpanzees no predictions for specific positive emotions were made given that happy/positive emotions in chimpanzees are very rarely observed, except in playful activities in which the ape being physically touched (tickled/chased) performs play panting vocalizations (laughter-like) [39].

We were also interested in whether, in line with previous work in humans[11], there were signs of empathic distress when witnessing prosocial agents being punished. In children, there is already a wealth of evidence for such basic empathic tendencies when watching others harm themselves[40,41], as expressed by verbalizations and facial expressions such as frowns[42]. Thus, for children we predicted that they would show greater signs of empathic distress (increased frowns) in response to the punishment of the prosocial compared to the antisocial agent. Whether chimpanzees display empathic tendencies in such situations is much less known. One key behavioral indicator of empathic distress is whether individuals have a motivation to escape the distressing situation[14]. Chimpanzees approach victims of aggression and direct agonistic behavior towards aggressors and/or affiliative behavior towards victims[20]. We were therefore interested in whether the punishment of the prosocial agent would elicit escape behavior (by operating the heavy sliding door and moving into another part of the room without visual access to the punishment of the agent) or approach behavior (i.e. by remaining in the room during the punishment). For chimpanzees, we also used their vocalizations (here defined as a compound of distress and display vocalizations, See Material and Methods section for more details) during the initial punishment as indicators of emotional arousal. The vocalizations were categorized according to their acoustic and temporal properties[43] and grouped according to the call categories suggested by Goodall[44].

We performed an additional study with chimpanzees (Study 3) to control for the possibility that incurring a cost to watch an antisocial agent being punished merely indicates that this is seen as more socially informative or more consistent with the flow of the preceding events. The execution of Study 3 was identical to that of Study 1, with the single difference that in Study 3 chimpanzees did not directly experience but merely witnessed, how the prosocial and the antisocial agents interacted with another chimpanzee (stooge). If chimpanzees preferentially watch the punishment of antisocial agents as a function of these more superficial aspects rather than their motivational substrate (anger- and revenge-based vs. norm-based punishment), the pattern of results should be the same in both studies. Based on previous studies showing that chimpanzees do not punish others who stole food from third parties[38,45] but they preferentially beg for food from those who were prosocial to others[37,46] we predicted that chimpanzees in Study 3, unlike Study 1, would not care to watch or vocalize differentially when others (regardless of whether they were prosocial or antisocial) were being punished. Note that Study 3 differed from Study 1 only in terms of the extent to which the chimpanzee subjects were directly affected by the agents' behavior, while keeping all other aspects of the experimental set-up constant.

It is important to note that our dependent behavioral variable of opening the heavy sliding door for the chimpanzees is always the same throughout all conditions. However, we interpret it differently depending on the condition (i.e., to continue witnessing the punishment when it is invisible or to escape into another room when it is visible; see Discussion section for more details). While we tested three age groups of children, we were agnostic to any age-related changes in our variables of interest. Given our a-priori predictions one-tailed statistics were applied for the factor prosociality. All other comparisons were two-tailed. Thus, for the chimpanzees (Studies 1 and 3) this resulted in a 2x2 factorial design with factor prosociality (prosocial/antisocial) and visibility (visible/invisible) and one trial for each condition. For children (Study 2) this resulted in a design with one factor of prosociality (prosocial/antisocial) and with 4 trials for each condition.

Results

Study 1: Chimpanzees, Watching punishment following directly experienced pro- and antisocial behavior

Chimpanzees differentially operated the heavy sliding door depending on whether punishment was visible or not and whether the agent had been previously prosocial towards them or not

(Cochran's Q = 8.59, df = 3, P = 0.043, N = 16). We conducted pair-wise follow-up comparisons between the two invisible conditions to test our hypothesis of an increased motivation to witness the punishment of an agent who had been previously antisocial towards the subject. Subjects were significantly more likely to incur the physical costs to open the heavy metal door in the antisocial invisible condition (50% of the subjects) compared to the prosocial invisible condition (18.75% of the subjects) (Sign test: P = 0.032, N = 16, one-tailed; Figure 2A). We conducted another pair-wise follow-up comparison between the two visible conditions to test for the behavioral effects of empathic distress (i.e. increased opening of the door to move to another room when the punishment of the prosocial agent is visible to the subject). Here we found no significant difference in the number of subjects who opened the door during the prosocial visible condition compared to the antisocial visible condition (Sign test: P = 0.313, N = 16, one-tailed; Figure 2A).

To assess the presence of vocalizations associated with emotional arousal during the punishment of either of the agents, the testing event was divided into three periods; an initial baseline where just the agent was present; a pre-hit period where the punisher appeared but had not started to punish the agent, and a first-hit period during which the punishment actually took place. We looked at these periods separately for each of the two agents. There was a significant difference between the three periods in the duration of the vocalizations in the presence of the prosocial agent (Friedman exact test: F = 9.82, P = 0.004, N = 16; Figure 2C) but we found no such difference in the presence of the antisocial agent (F = 4.67, P = 0.107, N = 16; Figure 2C). Comparing the vocalizations in response to the presence and punishment of the prosocial and the antisocial agents, showed that chimpanzees produced longer vocalizations in the baseline period when facing the antisocial agent compared to the prosocial one (Wilcoxon exact test: T^+ = 21, P = 0.031, N_{total} = 16; corrected for the duration of each period in the Punishment phase, i.e., baseline, pre-hit, hit periods) and longer vocalizations when the prosocial agent was being punished compared to when the antisocial agent was being punished in the hit period (Wilcoxon exact test: T^+ = 21, P = 0.031, N_{total} = 16; Figure 2C).

To assess whether the prosocial/antisocial exposure procedure had been effective, we assessed the subjects' preference for the prosocial and antisocial agent upon completion of the tasks (see Materials and Methods section). This was tested by allowing the chimpanzees to beg for food from the two agents to assess whether they showed a preference for one of them. Chimpanzees showed no preference for requesting food from the prosocial over the antisocial

agent (Wilcoxon signed rank test: $T^+ = 89$, $N = 17$, $P = 0.579$). This could have been the result of the close physical proximity of both agents, which might not have allowed for a clear dissociation of the subject's behavior.

Finally, we also assessed relationships between the chimpanzees' vocalizations and their behavior. We found that chimpanzees who produced vocalizations during the punishment of the prosocial agent were more likely to open the door to continue witnessing punishment of the antisocial agent than those who did not produce any vocalizations (57% vs 12.5%; Chi-Square test: $\chi^2 = 5.402$, P = 0.041). This suggests, that those chimpanzees who signal distress in response to a prosocial agent's punishment are also more motivated to observe deserve punishment being enacted.

Study 2: Children

To test for the hypothesis that children would show an increased motivation to observe the punishment of a previously antisocial agent, we compared the number of MUs spent on continuing to watch the punishment of the prosocial and the antisocial agents. The data were normally distributed and met assumptions for parametric tests. A Repeated Measures ANOVA with agent as a within-subject and age-group as a between-subject factor, indicated a significant interaction between the factors agent and age-group in how MUs were allocated to watch the punishment ($F_{(2,62)} = 3.417$; $P = 0.039$, Figure 2B). Thus, only 6-year-olds allocated more MUs to watch the punishment of the antisocial compared to the prosocial agent ($F_{(1,20)} = 12.246$; $P = 0.002$; for 4- and 5-year olds $p > 0.2$; Figure 2B). While there was a linear increase in comprehension of the task with age ($F_{(2,62)} = 5.26$; $P = 0.007$) this did not correlate with MUs allocated either for watching punishment of the prosocial or the antisocial agent (all $rs < 0.2$; $P > 0.1$).

Coding of facial expressions while watching the initial round of punishment showed significant age-differences in number of smiles co-occurring with frowns depending on which agent was being punished ($F_{(1,62)} = 2.294$; $P = 0.03$, one-tailed; Figure 2D). Thus, only 6-year-olds showed an increased mixture of positive and negative emotions (facial expressions) while watching the punishment of the antisocial compared to the prosocial agent ($F_{(1,20)} = 3.155$; $P = 0.045$, one-tailed; Figure 2D). We assessed the number of frowns during the initial round of punishment as an indication of empathic distress in the children at seeing the punishment of the agents. Whereas children frowned for both the prosocial (one-sample t-test: $t_{(64)} = 2.408$; $P = 0.019$) and

the antisocial agent (one-sample t-test: $t_{(64)}$ = 2.644; P = 0.010), this did not differ between the two agents. Frowning during the punishment did not interact further with age (P > 0.4).

To test the children for a preference for either of the two agents, children were asked explicitly which of the two agents they i) considered nicer, ii) would be more willing to share with and iii) would prefer to play with (see Materials and Methods section and SI). Children of all three age groups displayed a clear preference for the prosocial over the antisocial agent (paired t-test: $t_{(64)}$ = 4.279; P < 0.001) with no age differences in this preference (One-way ANOVA; P > 0.607).

Study 3: Chimpanzees, Watching punishment following indirectly experienced pro- and antisocial behavior

This study was conducted to rule out potential alternative explanations for the outcome of Study 1 including an increased social informational value in seeing antisocial others receive punishment or finding it more coherent in terms of the unfolding of events. Unlike Study 1, we found no evidence that chimpanzees differentially opened the heavy sliding door in the four conditions (Cochran's Q = 3, df = 3, P = 0.484, N = 14).

We also analyzed the presence of vocalizations associated with emotional arousal during the punishment of each of the agents during the baseline, pre-hit and first-hit periods. There was no significant difference between the three periods in the duration of the vocalizations in the presence of the prosocial and antisocial agents (Prosocial, Friedman exact test: F = 0.125, P = 1.00, N = 14; Antisocial, F = 3.26, P = 0.218, N = 14).

Russell and colleagues[37] showed that upon witnessing an interaction between a human beggar and either a nice or a nasty agent, chimpanzees showed a preference for the former. We used Russell et al.'s paradigm to test for a potential preference between the prosocial and the antisocial agent[37]. We found that chimpanzees begged significantly more often from the prosocial than the antisocial agent (frequency of begs corrected for the amount of time spent in front of the correspondent agent, Wilcoxon signed rank test: T^+ = 82, N_{total} = 14, P = 0.008).

Discussion

Our findings demonstrate that chimpanzees and 6-year old but not 4 and 5-year old children appear to possess a motivation to watch the punishment of others who they had previously experienced as antisocial towards themselves as compared to prosocial agents. Thus,

chimpanzees endured greater physical efforts and 6-year-old children spent more valuable MUs to continue watching the punishment of an agent who had previously withheld something valuable from the subjects (i.e. food for the chimpanzees and favorite toys for the children) as compared to someone who had been prosocial and shared the valuable items. In contrast, chimpanzees spent the same effort to continue watching the punishment of a human agent regardless of the agents' social behavior towards other chimpanzees.

We observed concomitant indicators of affective responses in the children. Six-year old children showed a greater mixture of positive and negative emotions in response to watching the punishment of the antisocial agent compared to the prosocial one. The combination of these emotions, rejoicing in the misfortune of a disliked other, is also known as Schadenfreude[47]. These data suggest that in children, pleasure at seeing deserved punishment may be linked to the increased costs incurred to continue watching it. Recent studies have shown that differential punishment of selfish behaviors of in-group and out-group members already occurs from 6 years onwards[48] and that around 6 years, children are capable of experiencing such potentially conflicting emotions[49]. Thus, 6 years of age may be a critical developmental time point at which children are willing to actually sacrifice their resources to see fairness enacted[50]. Importantly, even though there were some age differences in the comprehension of the experimental procedure, comprehension scores did not correlate with our behavioral measure, suggesting that any differences in comprehension cannot account for the age-related effect in the MUs expended. Further, our MUs were made meaningful to children through a subsequent conversion to stickers, which have been shown to be valuable items for the youngest as well as the oldest children of our age groups[51-54].

Previous studies have shown that chimpanzees engage in punishment of conspecifics who had previously stolen their food by causing the thief's food to disappear[38,45]. Study 1 demonstrates that also in the absence of food, chimpanzees are motivated to watch antisocial agents being punished after directly experiencing the antisocial behavior themselves. One could argue that the chimpanzees' reaction could be driven by emotional engagement. However, chimpanzees were more aroused when they watched the punishment of the prosocial agent. Following indirectly experienced pro- and antisocial behavior, chimpanzees were equally motivated to watch punishment of the pro- and the antisocial agents. This is consistent with findings showing that chimpanzees do not punish those who stole food from third parties[45]. The results from Study 3, in which chimpanzees merely observed the prosocial and antisocial interaction prior to

306 the agents' punishment, help us to interpret the results from Study 1. In both studies all basic
307 elements were kept constant except for the degree of the chimpanzee's involvement. Thus,
308 alternative explanations such as increased social informational value or a greater coherence in
309 the unfolding of the scene can be ruled out. Instead, the most likely interpretation based on
310 these findings is that chimpanzees have an increased motivation to observe such punishment
311 because it follows a desirable action towards someone who behaved antisocially towards
312 themselves. The literature abounds with examples of animals willing to incur energy costs for
313 something they find rewarding[55-57]. It is therefore tempting to argue that watching antisocial
314 others getting harmed is rewarding and pleasurable also to chimpanzees. Suggestive of an
315 emotional antecedent to such behavior is also the finding of individual differences in the
316 relationship between vocalizations and opening the door to witness punishment. Thus,
317 chimpanzees who had vocalized distress during the punishment of the prosocial agent were
318 also more likely to incur a cost to continue witnessing the deserved punishment of an antisocial
319 other. Thus, when punishment is deserved, the experience of distress is abolished leading
320 chimpanzees to actively seek out observing such punishment. However, in the absence of direct
321 evidence, we remain cautious with an account positing the presence of actual positive emotions
322 as a driver for the observed behavior.
323
324 In addition to signs of Schadenfreude in children, we found evidence of empathic distress
325 across all three age groups. However, this was not differentially modulated by whether the
326 agent had been previously prosocial or antisocial towards them. Even though children as young
327 as 3-year old have been shown to differentiate their empathic helping between previously
328 prosocial and antisocial others[58,59] and all age groups showed a decided preference for the
329 prosocial agent, no difference in empathic responding could be found. Chimpanzees produced
330 longer vocalizations indicative of emotional arousal during the punishment of the prosocial agent
331 that had directly interacted with them but no differential vocalizations occurred when they
332 witnessed the agent being punished following the indirectly experienced pro- and antisocial
333 behavior (regardless of her social orientation). Even though in chimpanzees it is difficult to
334 clearly label the valence of such vocalizations as they can reflect conflicting emotions[44], the
335 specificity of their occurrence (longer vocalizations during the hitting of the prosocial agent
336 compared to the antisocial agent) suggests that they might reflect something akin to empathic
337 distress. However, chimpanzees did not signal distress by attempting to escape witnessing the
338 punishment of the prosocial agent nor tried to approach and console the victim of the
339 aggression as suggested by observational studies[20]. These conflicting results (distress

vocalizations vs. non-escape/non-approach behavior) make it difficult to pinpoint the underlying motivation of the chimpanzees' behaviors upon witnessing the punishment of the prosocial agent.

There are some limitations to the present set of studies. One is the fact that interactions were observed between individuals that were not of the same species as the subject. However, this concern is reduced given that both chimpanzees and 6-year-olds responded differentially to the two agents. While such cross-species set-ups are common in the study of social behavior of both human and non-human primates[60-62] future work will have to assess how far these findings extend onto interactions with one's own species. Further, the different dependent variables for the chimpanzees and the children (i.e. physical energy vs. valuable MUs) makes direct inter-specific comparisons difficult. While using different dependent variables has the advantage of optimizing procedures for each species thus avoiding potential biases favoring one of the species, future work may seek to expand the findings using the same dependent variables for greater comparability of the effects. Finally, we were unable to counterbalance the administration of the direct and indirect exposures to the pro- and antisocial in chimpanzees. Our results, however, were consistent with the existing literature on the occurrence of punishment following directly and indirectly experienced transgressions in chimpanzees, which ameliorates to some extent the concerns derived from our current design.

We studied the evolutionary and ontogenetic origins of an increased motivation to watch the punishment of antisocial others and their associated emotional states. Chimpanzees and 6-year-old children showed greater motivation by incurring costs to continue watching the punishment of an antisocial over a prosocial agent. Furthermore, children displayed differential responses of mixed positive and negative emotions when they witnessed punishment of antisocial agents, which suggest that they might take some form of pleasure from this. Although such a mechanism is still uncertain in chimpanzees, vocalizations of emotional arousal produced when they witnessed the suffering of a prosocial agent, and their absence when witnessing the suffering of an antisocial agent, might indicate that affective responses such as pleasure may constitute an important motivational contributor to the exaction of revenge, with early evolutionary origins. Crucially, chimpanzees did not vocalize differentially for the two agents when seeing the two agents punished following indirectly experienced pro- and antisocial behavior. Additionally, they did not engage in differential costs to witness the punishment of the antisocial agent as compared to the prosocial agent. These findings provide

some evidence for the evolutionary origins of an increased motivation to watch punishment of antisocial behavior with - at least in children- possible links to feelings of pleasure underlying such a motivation. Such a motivation appears to develop at a protracted rate, similar to higher-level cognitive skills[63] and might emerge at an age at which children begin to care so much for justice that they are willing to pay for it.

Methods and Materials

Ethics statement

The studies reported in this manuscript were approved by the local ethics committee of the University of Leipzig and complied with all relevant regulations. Thus, the ethics committee of the University of Leipzig approved the study (Ethics Approval Number: 367-11-26092011). Caregivers provided written consent form to use the acquired data. Additionally, the chimpanzee work was approved by the MPI-EVA – Zoo Leipzig ethical committee.

Participants

Studies 1 and 3: In Study 1 we tested 17 chimpanzees (*Pan troglodytes*). There were 5 males ranging in age between 8 and 38 years (*M* = 16 years and 8 months) and 12 females ranging in age between 8 and 37 years (*M* = 22 years and 5 months). In Study 3, we tested 14 chimpanzees. There were 5 males ranging in age between 8 and 38 years (*M* = 15 years and 10 months) and 9 females ranging in age between 12 and 42 years (*M* = 27 years and 3 months). All chimpanzees were housed at the Wolfgang Koehler Primate Research Center, Leipzig Zoo, Germany. Eleven of them participated in both studies, whereas the rest could not do so because they were unavailable (see Table S1 for rearing history and detailed participation in each study). All indoor and outdoor enclosures were furnished with vegetation, climbing structures and visual barriers. Subjects were neither food- or water-deprived during the experiment.

Study 2: We tested 72 children. There were three age groups: 24 4-year-olds (*M* = 4.15, age range = 4.04-4.35), 24 5-year-olds (*M* = 5.04, age range = 4.97-5.4), and 24 6-year-olds (*M* = 6.17, age range= 5.98-6.33). In each group there were equal number of boys and girls. Seven children had to be removed from the analyses due to procedural error or fussiness. All remaining subjects received all conditions. All children were recruited from a subject database at the Max Planck Institute for Human Cognitive and Brain Sciences in Leipzig, Germany.

Experimental Procedures

Studies 1 and 3: These two studies consisted of four phases: *Training, Exposure, Preference and Punishment* (actual test). Before entering the *Punishment* phase, chimpanzees received a sequence of training stages (see *Training phase* in SI) to ensure that they understood how to open the heavy mesh sliding door that would allow them access to the adjacent room. After the training, all subjects were exposed directly (Study 1) or indirectly (Study 3, by witnessing an interaction between a human agent and a conspecific stooge) to two different human agents, one at a time. The agents either acted prosocially towards the subject/stooge (Study1/Study3), by providing food, or antisocially, by teasing and not allowing the subject/stooge to get access to the food (see *Exposure phase* in SI). Whether the agent was prosocial or antisocial was counterbalanced across subjects. To reduce carry-over effects between studies, different agents participated in Study 3 (except for the punisher), which was conducted a few months after Study 1. To test the efficacy of the *Exposure phase* a *Preference phase* was designed to test for preferential begging from the two agents (see SI for more details). In the *Punishment phase* (see SI for more details) either the prosocial or the antisocial agents entered the testing room and sat in front of the Plexiglas window in the subjects' room. After 5 seconds of being seated in front of the Plexiglas window (henceforth referred to as baseline period), a second agent, the punisher, entered the room. The punisher approached the agent from behind with a human facial expression of rage (henceforth referred to as pre-hit period) and started beating her up (henceforth referred to as hit period) with a stick for 4 sec. (i.e., 4 hits with the stick, rate: 1 Hz). While being beaten up the agent cried out in pain. After the initial punishment period (i.e., 4 seconds) the agent either: 1) remained in her initial position for the whole time of the punishment visible to the subject (10 more seconds, Figure 1A), so-called *visible* condition, or 2) left her initial position (area A, see Figure 1A) and went into another area of the room invisible to the chimp (area B, see Figure 1A) where the punishment continued for 10 more seconds, so-called *invisible* condition. If subjects wanted to continue watching the punishment in the invisible condition they had to open the heavy sliding door (learned during the Training phase) and move in front of the Plexiglas window in the new room. Similarly, if they wanted to escape from the punishment in the visible condition happening in front of them, they had to operate the door to move to another part of the room where this would then be invisible.

All sessions were videotaped and the following variables were coded from digital files: 1) opening of the heavy sliding door; 2) duration of the vocalizations associated with emotional arousal, namely: screams, whimpers, and worried hoos considered as distress vocalizations[43]

and (waa) barks and (pant) hoots considered as display vocalizations[43]. As previously mentioned, vocalizations were categorized according to their acoustic structure and temporal measures and grouped according to the call categories suggested by Goodall[44]. Distress and display vocalizations were lumped together and the combined results used for statistical analysis. The duration of the calls was analyzed with the sound analysis software Avisoft and Praat.

To assess inter-observer reliability, a second observer coded a random sample of 20% of the trials. Inter-observer reliability was high for opening the sliding door (Study 1: Pearson correlation $r = 1.000$, $P < 0.001$; Study 3: $r = 1.000$, $p < 0.001$), for duration of the vocalizations (Study 1, distress calls: $r = 1.000$, $P < 0.001$; display calls: $r = 0.900$, $P < 0.001$; Study 3, distress calls: $r = 1.000$, $P < 0.001$; display calls: $r = 1.000$, $P < 0.001$), and frequency of begs corrected for the amount of time spent in front of the correspondent human agent (Study 1: $r = 0.999$, $P = 0.028$; Study 3: $r = 0.997$, $p = 0.048$),

Study 2: Children came into the lab accompanied by at least one parent. Parents had been instructed before on the phone to bring six of their child's favorite toys, without the child noticing. These were then taken by the experimenter and used as in the two exposure phases. Children were given an initial endowment of 4 MUs. It was made clear that at the end of the experiment each of the MUs could be traded for one sticker. The experimental procedure was demonstrated using a miniature-sized puppet theatre.

Before the *Punishment phase*, each child was exposed consecutively to two different puppets, a prosocial and an antisocial puppet (see *Exposure phase* in SI). Exposure entailed one of two puppets to either act prosocially by returning three of the child's favorite toys, or antisocially, by keeping them for itself. The puppets would bring up a toy from behind the theatre and hold it up to the child. After telling the child that it wanted to play with them, the prosocial puppet would hold the toy towards the child and put it into the child's hands, whereas the antisocial puppet would withdraw as soon as the child reached for the toy. Similar procedures has been shown to elicit clear preferences in infants[64]. Which puppet was prosocial or antisocial was counterbalanced across subjects. Exposure and testing was performed for both puppets and fully counter-balanced across all subjects. During the *Punishment phase*, the puppet to which children had just been exposed remained on stage. After 5 seconds, another puppet appeared (different to the two agents) carrying a long stick (punisher). The punisher started beating the

other puppet (prosocial/antisocial) up with the stick for 5 seconds (i.e., 5 hits with the stick, rate: 1 Hz). After the initial punishment period (i.e., 5 seconds) the theatre curtain closed rendering both the punisher and the punished puppet invisible. The punisher puppet then returned and said to the child that they were going to continue hitting the other puppet and that if the child would like to continue watching then it should put one MU into a box to the right of the stage, whereas if it did not want to continue watching it should put a MU into a box to the left of the stage. Depending on where children placed their MU, the curtains were drawn again or not and children could continue observing the punishment or not. In case they chose not to witness the punishment, the punishment was still executed behind closed curtains. If children decided not to continue watching on the first round then the punisher puppet did not ask again whether the child cared for another round of witnessing punishment. However, if children decided to continue watching, the punisher asked again after 5 seconds of punishment if they would like to continue watching. Given that children had received 4 MUs, the maximum number of paid punishments was 4. Thus, all subjects received exposure to the first round of punishment and the first question of whether they would like to continue watching or not. Depending on whether children paid for punishment, they were asked again until they either decided to stop watching or until they had no more MUs. The final round was the pursuit and punishment behind the curtain, thus the child continued hearing the puppet crying for 10 more sec. but without visual access to the punishment.

All sessions were videotaped and the following variables were coded from digital files during the exposure phase as well as the punishment phase: 1) behaviors and verbalizations 2) pure smiles, pure frowns and given the potential ambivalence of seeing someone antisocial experience punishment, we also coded for smiles occurring jointly with frowns. Two observers coded all the videos using the Interact software.

To assess inter-observer reliability, ratings were correlated. Inter-observer reliability was high for answering the questions of the punisher ($r = 0.99$, $p < 0.0001$) as well as for occurrence of smiles, frowns and smiles with frowns during the exposure as well as the punishment phase (all $r > 0.504$, all $p < 0.0001$).

At the end of the entire Punishment phase the experimenter showed the two agents to the child and asked which puppet the child would rather play with, give a sticker to and thought was nicer. From this a composite score of preference was obtained (see SI).

All data were analyzed in SPSS 23 (SPSS Statistics Software, IBM). No attempts to replicate the findings reported in this paper have been made.

Data availability statement

The data that support the findings of this study are available from the corresponding author on reasonable request.

Author Contributions

Conceived and designed the experiments: NM, NS, JC, TS. Performed the experiments: NB, NM, NS. Analyzed the data: NM, NS. Interpretation of data and writing of the paper: NB, NM, NS, TS, JC. Funding provided by JC and TS.

References

1 Nowak, M. A. Five rules for the evolution of cooperation. *science* **314**, 1560-1563 (2006).
2 Nowak, M. A. & Sigmund, K. Evolution of indirect reciprocity. (2005).
3 Boyd, R., Gintis, H., Bowles, S. & Richerson, P. J. The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences* **100**, 3531-3535 (2003).
4 Fehr, E. & Gächter, S. Altruistic punishment in humans. *Nature* **415**, 137-140 (2002).
5 Henrich, N. & Henrich, J. P. *Why humans cooperate: A cultural and evolutionary explanation*. (Oxford University Press, 2007).
6 Henrich, J. & Boyd, R. Why people punish defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *Journal of theoretical biology* **208**, 79-89 (2001).
7 Clutton-Brock, T. H. & Parker, G. A. Punishment in animal societies. *Nature* **373**, 209 (1995).
8 Hauser, M. D. Costs of deception: cheaters are punished in rhesus monkeys (Macaca mulatta). *Proceedings of the National Academy of Sciences* **89**, 12137-12139 (1992).
9 De Quervain, D. J., Fischbacher, U., Treyer, V. & Schellhammer, M. The neural basis of altruistic punishment. *Science* **305**, 1254 (2004).
10 Hein, G., Silani, G., Preuschoff, K., Batson, C. D. & Singer, T. Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. *Neuron* **68**, 149-160 (2010).
11 Singer, T. *et al.* Empathic neural responses are modulated by the perceived fairness of others. *Nature* **439**, 466-469 (2006).
12 Batson, C. D. *The Altruism Question: Toward a Social-Psychological Answer*. (Erlbaum, Hillsdale, NJ, 1991).
13 Singer, T. *et al.* Empathy for pain involves the affective but not sensory components of pain. *Science* **303**, 1157-1162 (2004).

14 Batson, C. D., Duncan, B. D., Ackerman, P., Buckley, T. & Birch, K. Is empathic emotion a source of altruistic motivation? *Journal of personality and Social Psychology* **40**, 290 (1981).
15 Cools, A. K., Van Hout, A. J. M. & Nelissen, M. H. Canine Reconciliation and Third‐Party‐Initiated Postconflict Affiliation: Do Peacemaking Social Mechanisms in Dogs Rival Those of Higher Primates? *Ethology* **114**, 53-63 (2008).
16 Palagi, E. & Cordoni, G. Postconflict third-party affiliation in Canis lupus: do wolves share similarities with the great apes? *Animal behaviour* **78**, 979-986 (2009).
17 Seed, A. M., Clayton, N. S. & Emery, N. J. Postconflict third-party affiliation in rooks, Corvus frugilegus. *Current Biology* **17**, 152-158 (2007).
18 Byrne, R. *et al.* Do elephants show empathy? *Journal of Consciousness Studies* **15**, 204-225 (2008).
19 Clay, Z. & de Waal, F. B. Bonobos respond to distress in others: consolation across the age spectrum. *PLoS One* **8**, e55206 (2013).
20 Romero, T. & de Waal, F. Chimpanzee (Pan troglodytes) consolation: third-party identity as a window on possible function. *Journal of Comparative Psychology* **124**, 278 (2010).
21 Mallavarapu, S., Stoinski, T., Bloomsmith, M. & Maple, T. Postconflict behavior in captive western lowland gorillas (Gorilla gorilla gorilla). *American Journal of Primatology* **68**, 789-801 (2006).
22 Langford, D. J. *et al.* Social modulation of pain as evidence for empathy in mice. *Science* **312**, 1967-1970 (2006).
23 Bartal, I. B.-A., Decety, J. & Mason, P. Empathy and pro-social behavior in rats. *Science* **334**, 1427-1430 (2011).
24 Burkett, J. P. *et al.* Oxytocin-dependent consolation behavior in rodents. *Science* **351**, 375-378 (2016).
25 Hamlin, J. K., Wynn, K. & Bloom, P. Social evaluation by preverbal infants. *Nature* **450**, 557-559 (2007).
26 Hamlin, J. K. & Wynn, K. Young infants prefer prosocial to antisocial others. *Cognitive development* **26**, 30-39 (2011).
27 Hamlin, J. K., Wynn, K., Bloom, P. & Mahajan, N. How infants and toddlers react to antisocial others. *Proceedings of the national academy of sciences* **108**, 19931-19936 (2011).
28 Schulz, K., Rudolph, A., Tscharaktschiew, N. & Rudolph, U. Daniel has fallen into a muddy puddle–Schadenfreude or sympathy? *British Journal of Developmental Psychology* **31**, 363-378 (2013).
29 Shamay-Tsoory, S. G., Ahronberg-Kirschenbaum, D. & Bauminger-Zviely, N. There is no joy like malicious joy: Schadenfreude in young children. *PloS one* **9**, e100233 (2014).
30 Tisak, M. S. Preschool children's judgments of moral and personal events involving physical harm and property damage. *Merrill-Palmer Quarterly (1982-)*, 375-390 (1993).
31 Jordan, J. J., McAuliffe, K. & Warneken, F. Development of in-group favoritism in children's third-party punishment of selfishness. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 12710-12715, doi:10.1073/pnas.1402280111 (2014).
32 De Waal, F. B. *Good natured*. (Harvard University Press, 1996).
33 De Waal, F. B. & Luttrell, L. M. Mechanisms of social reciprocity in three primate species: symmetrical relationship characteristics or cognition? *Ethology and Sociobiology* **9**, 101-118 (1988).
34 Jensen, K., Call, J. & Tomasello, M. Chimpanzees are rational maximizers in an ultimatum game. *science* **318**, 107-109 (2007).

35 Suchak, M. *et al.* How chimpanzees cooperate in a competitive world. *Proceedings of the National Academy of Sciences*, 201611826 (2016).
36 Herrmann, E., Keupp, S., Hare, B., Vaish, A. & Tomasello, M. Direct and indirect reputation formation in nonhuman great apes (Pan paniscus, Pan troglodytes, Gorilla gorilla, Pongo pygmaeus) and human children (Homo sapiens). *Journal of Comparative Psychology* **127**, 63 (2013).
37 Russell, Y. I., Call, J. & Dunbar, R. I. Image scoring in great apes. *Behavioural Processes* **78**, 108-111 (2008).
38 Jensen, K., Call, J. & Tomasello, M. Chimpanzees are vengeful but not spiteful. *Proceedings of the National Academy of Sciences* **104**, 13046-13050 (2007).
39 Matsusaka, T. When does play panting occur during social play in wild chimpanzees? *Primates* **45**, 221-229 (2004).
40 Fabes, R. A., Eisenberg, N. & Eisenbud, L. Behavioral and physiological correlates of children's reactions to others in distress. *Developmental Psychology* **29**, 655 (1993).
41 Eisenberg, N. *et al.* The relations of children's dispositional empathy-related responding to their emotionality, regulation, and social functioning. *Developmental psychology* **32**, 195 (1996).
42 Valiente, C. *et al.* Prediction of children's empathy-related responding from their effortful control and parents' expressivity. *Developmental psychology* **40**, 911 (2004).
43 Nishida, T., Zamma, K., Matsusaka, T., Inaba, A. & McGrew, W. C. *Chimpanzee behavior in the wild: an audio-visual encyclopedia*. (Springer Science & Business Media, 2010).
44 Goodall, J. *The chimpanzees of Gombe: Patterns of behavior*. (MA: Belknap Press of Harvard University Press, Cambridge, 1986).
45 Riedl, K., Jensen, K., Call, J. & Tomasello, M. No third-party punishment in chimpanzees. *Proceedings of the National Academy of Sciences* **109**, 14824-14829 (2012).
46 Anderson, J. R., Takimoto, A., Kuroshima, H. & Fujita, K. Capuchin monkeys judge third-party reciprocity. *Cognition* **127**, 140-146 (2013).
47 Smith, R. H. *et al.* Envy and schadenfreude. *Personality and Social Psychology Bulletin* **22**, 158-168 (1996).
48 Jordan, J. J., McAuliffe, K. & Warneken, F. Development of in-group favoritism in children's third-party punishment of selfishness. *Proceedings of the National Academy of Sciences* **111**, 12710-12715 (2014).
49 Steinbeis, N. & Singer, T. The effects of social comparison on social emotions and behavior during childhood: The ontogeny of envy and Schadenfreude predicts developmental changes in equity-related decisions. *Journal of Experimental Child Psychology* **115**, 198-209 (2013).
50 McAuliffe, K., Jordan, J. J. & Warneken, F. Costly third-party punishment in young children. *Cognition* **134**, 1-10 (2015).
51 Engelmann, J. M., Over, H., Herrmann, E. & Tomasello, M. Young children care more about their reputation with ingroup members and potential reciprocators. *Developmental Science* **16**, 952-958 (2013).
52 Blake, P. R., Piovesan, M., Montinari, N., Warneken, F. & Gino, F. Prosocial norms in the classroom: The role of self-regulation in following norms of giving. *Journal of Economic Behavior & Organization* **115**, 18-29 (2015).
53 Smith, C. E., Blake, P. R. & Harris, P. L. I should but I won't: Why young children endorse norms of fair sharing but do not follow them. *PloS one* **8**, e59510 (2013).
54 Dunfield, K., Kuhlmeier, V. A., O'Connell, L. & Kelley, E. Examining the diversity of prosocial behavior: Helping, sharing, and comforting in infancy. *Infancy* **16**, 227-247 (2011).

55 Salamone, J. D., Correa, M., Farrar, A. & Mingote, S. M. Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology* **191**, 461-482 (2007).
56 Grossbard, C. L. & Mazur, J. E. A comparison of delays and ratio requirements in self-control choice. *Journal of the Experimental Analysis of Behavior* **45**, 305-315 (1986).
57 Beran, M. J. & Evans, T. A. Delay of gratification by chimpanzees (Pan troglodytes) in working and waiting situations. *Behavioural processes* **80**, 177-181 (2009).
58 Eisenberg, N. *et al.* The relations of emotionality and regulation to dispositional and situational empathy-related responding. *Journal of personality and social psychology* **66**, 776-797 (1994).
59 Vaish, A., Carpenter, M. & Tomasello, M. Young children selectively avoid helping people with harmful intentions. *Child development* **81**, 1661-1669 (2010).
60 Custance, D. M., Whiten, A. & Bard, K. A. Can young chimpanzees (Pan troglodytes) imitate arbitrary actions? Hayes & Hayes (1952) revisited. *Behaviour* **132**, 837-859 (1995).
61 Whiten, A., Custance, D. M., Gomez, J.-C., Teixidor, P. & Bard, K. A. Imitative learning of artificial fruit processing in children (Homo sapiens) and chimpanzees (Pan troglodytes). *Journal of comparative psychology* **110**, 3 (1996).
62 Call, J., Hare, B., Carpenter, M. & Tomasello, M. 'Unwilling'versus 'unable': chimpanzees' understanding of human intentional action. *Developmental science* **7**, 488-498 (2004).
63 Hanus, D., Mendes, N., Tennie, C. & Call, J. Comparing the performances of apes (Gorilla gorilla, Pan troglodytes, Pongo pygmaeus) and human children (Homo sapiens) in the floating peanut task. *PloS one* **6**, e19555 (2011).
64 Dunfield, K. A. & Kuhlmeier, V. A. Intention-mediated selective helping in infancy. *Psychological science* **21**, 523-527 (2010).

Competing interest statement

The authors declare no competing interests.

Acknowledgments

We are grateful to Mike Tomasello for early input into the study design and to Matthias Allritz, Vera Ehrich, Kerstin Esau, Elisa Felsche, Johannes Grossmann, Susan Hunger, Saskia Lorenz, Julia Steinhardt, Katrin Schumann, Katja Waldherr and Katharina Wenig for helping with the training phase and data collection with the chimpanzees at the Wolfgang Köhler Primate Research Centre; to Yseult Hejja-Brichard and Katrin Schumann for analyzing the chimpanzee vocalizations, Katrin Schumann for analyzing part of the chimpanzee behavioral data, Markus Neuschulz and Anja Hutschenreiter for reliability with the chimpanzee data, and to Christine Brenner, Katharina Mueller, Charlotte Hoecker, and Jessica Buergel for the data collection with the children. Thank you to Thibaud Gruber, Catherine Crockford, and Ammie Kalan for helping us identifying some of chimpanzee vocalizations as well as Ammie Kalan for help with the

software Avisoft and Praat. Special thanks go to Henrik Grunert and Raik Pieszek for their help in constructing the experimental apparatus. Thanks also to the zookeepers at the Leipzig Zoo for their help with the chimpanzees. Salaries of N.S., N.M., T.S., as well as testing of the children were supported by a Max Planck budget granted to T.S. as director of the Department of Social Neuroscience. NS was supported by the European Research Council (ERC grant agreement n° 715282, project DEVBRAINTRAIN), as well as a Jacobs Research Fellowship. JC was supported in part by the European Research Council (ERC grant agreement n° 609819, project SOMICS). NBG was supported by an FPU scholarship from the Spanish Ministry of Education (Ref. FPU12/00409). With the exception of the Max Planck Society, none of the funders played a role in study design, data collection and analysis, decision to publish, or preparation of the manuscript

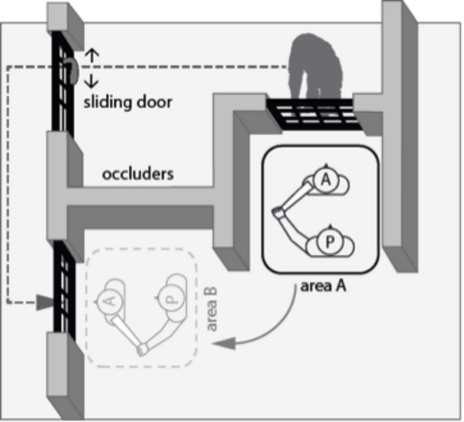
Figure Legends

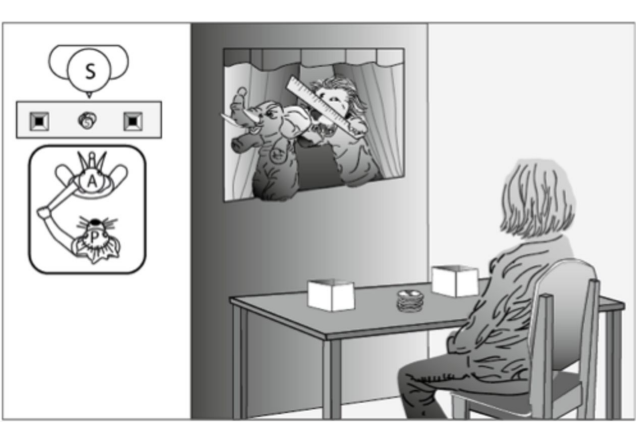
Figure 1. Experimental Design for (A) chimpanzees and (B) children. Subjects (S) watch the punishment of a previously either prosocial or antisocial agent (A) by a punisher (P). For the chimpanzees in the visible conditions, the punishment took place outside the cage of the chimpanzee. For the invisible conditions, the punishment moved to a part of the room out of sight from the chimpanzee. For the children the punishment was visible until a curtain fell and children were asked to put their MUs into the box on the right in order to continue watching the punishment.

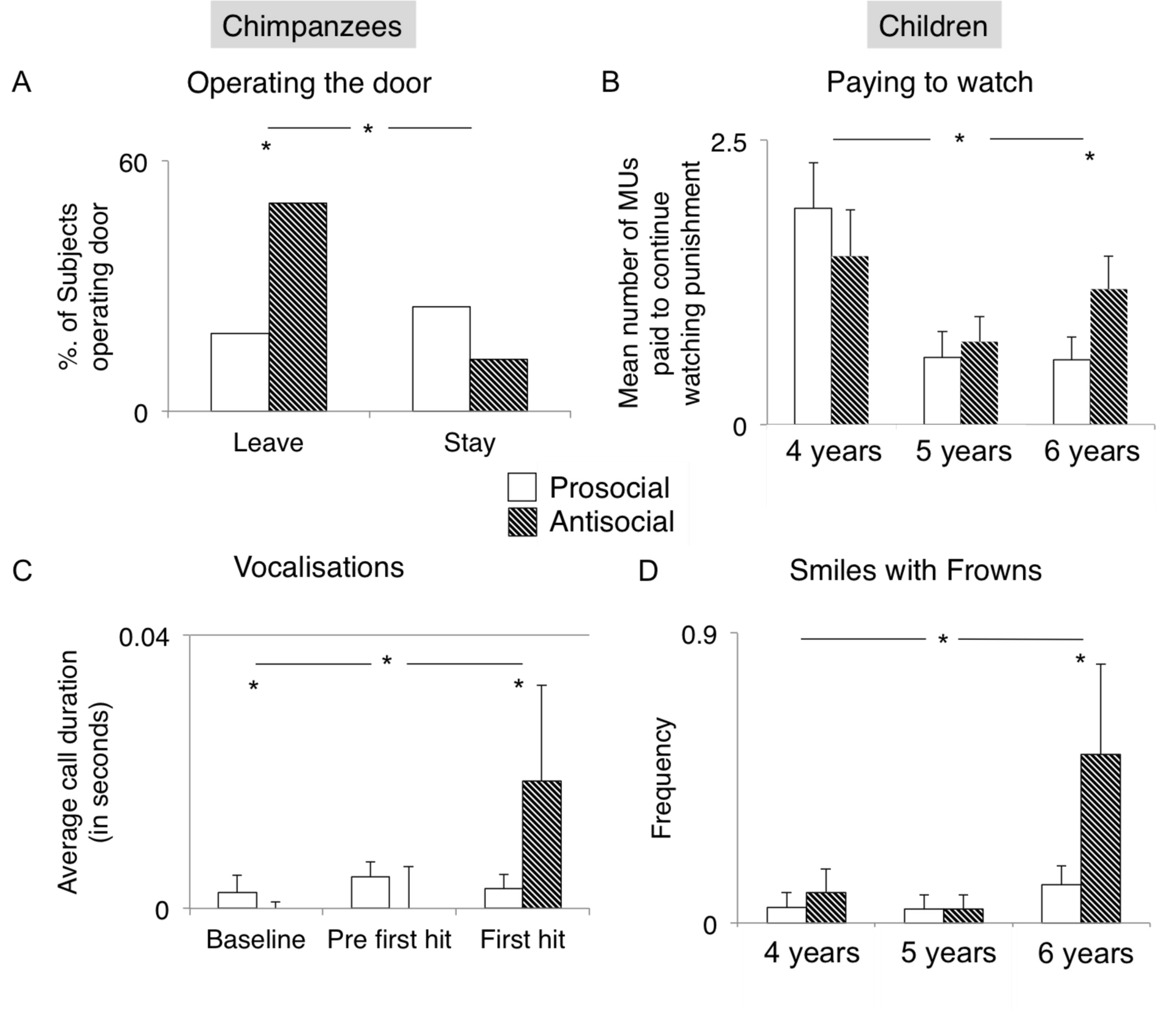
Figure. 2 Behavioral data and emotional indicators for chimpanzees (Study 1; N = 17) and children (Study 2; N = 65). (A) More chimpanzees opened the heavy sliding door to continue watching the punishment in the invisible antisocial (i.e. when punished and human agent left to move to an invisible part of the room) compared to the invisible prosocial condition. Note, that not all the chimpanzees opened the door. (B) All children paid to continue watching some of the punishment, but only 6-year olds paid more to watch the antisocial agent being punished compared to the prosocial agent. (C) Chimpanzees expressed greater distress vocalizations when watching the punishment of the prosocial human agent. (D) Only 6-year old children

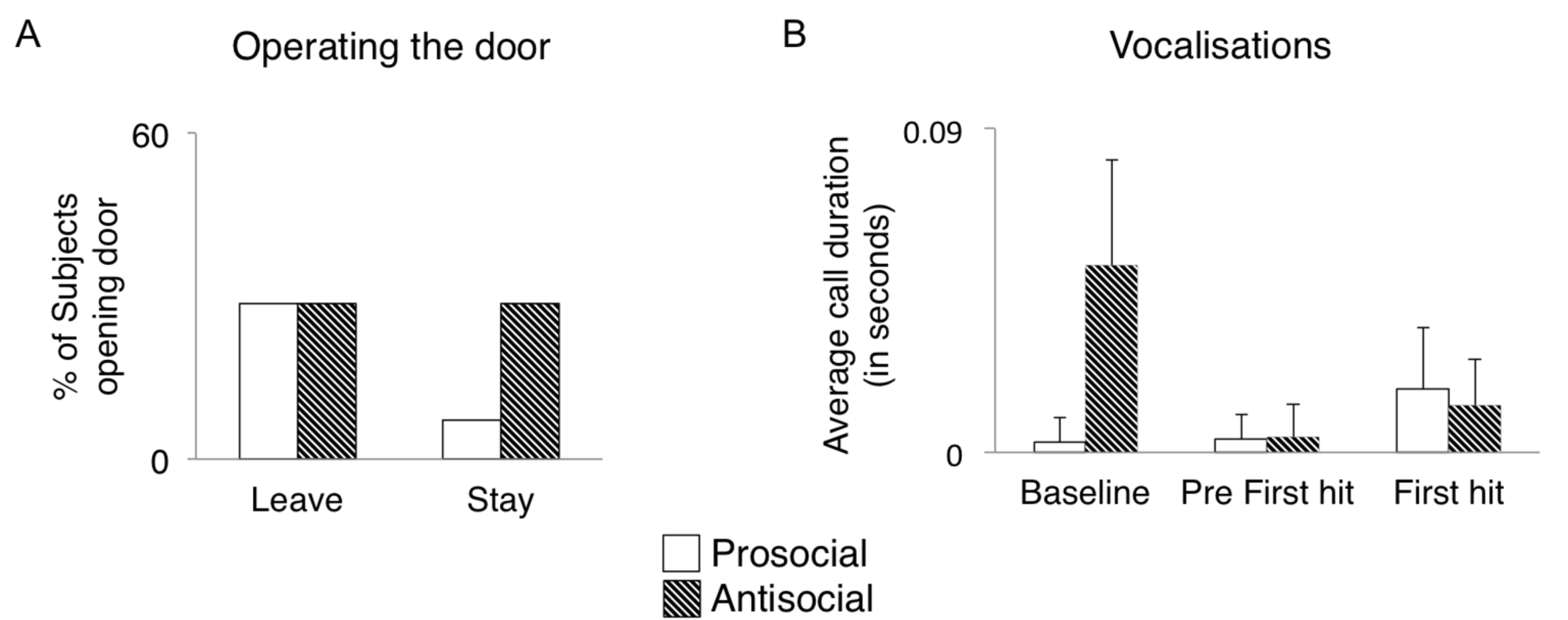
displayed more frequent smiles coupled with frowns during the punishment of the antisocial compared to the prosocial agent. The error bars show s.e.m.

Figure. 3 Behavioral data and emotional indicators for chimpanzees in Study 3 (N = 14). (A) There were no differences in the chimpanzees' (A) behavior or (B) vocalizations between any of the conditions. The error bars show s.e.m.


sliding door
occluders
A
P
area A
area B
A
P


S
A
P


Chimpanzees
Children
A
Operating the door
%. of Subjects operating door
60
0
Leave
Stay
*
*
B
Paying to watch
Mean number of MUs paid to continue watching punishment
2.5
0
4 years
5 years
6 years
*
*
Prosocial
Antisocial
C
Vocalisations
Average call duration (in seconds)
0.04
0
Baseline
Pre first hit
First hit
*
*
*
D
Smiles with Frowns
Frequency
0.9
0
4 years
5 years
6 years
*
*


A
Operating the door
60
0
% of Subjects opening door
Leave
Stay
B
Vocalisations
0.09
0
Average call duration (in seconds)
Baseline
Pre First hit
First hit
Prosocial
Antisocial