Instrumental helping and short-term reciprocity in chimpanzees and human children

Hagen Knofe[1], Jan M. Engelmann[*2], Sebastian Grueneisen[3], and Esther Herrmann[1,4]

[1]Minerva Research Group on the Origins of Human Self-Regulation, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany

[2]Department of Psychology, University of California, Berkeley, Berkeley, CA 94720-1650

[3]Faculty of Education, Leipzig University, 04109, Leipzig, Germany

[4]Department of Psychology, Centre for Comparative and Evolutionary Psychology, University of Portsmouth, PO1 2DY Portsmouth, UK

[*] Corresponding author: jan_engelmann@berkeley.edu; ORCID ID of the corresponding author: https://orcid.org/0000-0001-5139-0289

Acknowledgements

We thank Martin Mulama, Richard Vigne, George Paul, the board members, all the staff of Sweetwaters Chimpanzee Sanctuary in Kenya and Isabelle de Gaillande-Mustoe for their support during this research. We also thank the chimpanzee caregivers, especially David Mundia, Josep Mayo, and Anthony Kamau, the day care centers, chimpanzees and children, especially the stooges, for their help and cooperation. Thanks go out to Susanne Buhl, Benjamin Brückner, Moritz Kaplonek, Ken Krempler, Julia Siemann, Kristina Kellermann for assistance in testing the children; to Diotima Rapp, Rebecca Koomen for helpful discussions; Juliane Bojahr for her greatest support; Daniil Serko for statistical advice; Ronny Barr for apparatus figures; Ramona Frickel for reliability coding; and Anna-Claire Schneider for helpful comments on a previous version of the manuscript. We also thank Ol Pejeta Conservancy, Kenya Wildlife Service (KWS) and National Council for Science and Technology (NCST) for approving our research.

We have no competing interests to declare.

S.G. was supported by an ERC Starting Grant (grant agreement no. 948748)

Supporting information including the data accompanies this manuscript.

Author contribution

Hagen Knofe and Esther Herrmann conceptualized the original research idea and developed the methodology with substantial contributions of Jan Engelmann and Sebastian Grueneisen during later stages of the intellectual process. Hagen Knofe built the apparatuses, conducted the research and coded the data for both studies under the supervision of Esther Herrmann, who managed and coordinated the research. Hagen Knofe curated and analysed the data for the original version of the manuscript. Sebastian Grueneisen and Hagen Knofe created the models, analysed and visualised the data for the revised version of the manuscript. Hagen Knofe, Jan Engelmann and Esther Herrmann wrote and edited the initial draft of the paper. Hagen Knofe, Sebastian Grueneisen, Jan Engelmann and Esther Herrmann wrote and edited the revised version of the manuscript and the response letter.

ABSTRACT

Chimpanzees (*Pan troglodytes*) and humans cooperate in reciprocal patterns, but it is unclear whether these interactions are based on the same psychological foundations. While there is evidence suggesting that both species engage in long-term forms of reciprocity, there is very little work exploring their short-term behavioural contingencies with suitable methods. Here we present a direct comparative study on short-term reciprocity in chimpanzees and human children using a novel, low-cost instrumental helping task. We investigated whether participants help a conspecific partner to obtain a tool for accessing a reward, and whether the level of helping depends on the partner's previous *helpful* or *unhelpful* behaviour. In line with prior research, both chimpanzees and children demonstrated helping behavior toward their partner. However, the extent to which the two species showed short-term reciprocity differed considerably. After receiving help, tested children always helped in return. They helped substantially less when interacting with an unhelpful partner. Chimpanzees showed a higher tendency to help when interacting with a helpful compared to an unhelpful partner only in the first half the experiment. With increasing trial number, chimpanzees stopped discriminating between helpful and unhelpful partners. This study provides evidence for short-term reciprocity in human children and, to a lesser extent, in our closest living relatives. Our findings demonstrate that helping paradigms provide a useful context to investigate reciprocal motives in humans and chimpanzees alike.

Keywords: Short-term reciprocity, helping, prosocial behaviour, chimpanzees, children, instrumental helping

INTRODUCTION

The question of how natural selection can explain the existence of agents that deliver benefits to others at a cost to themselves has attracted much interest (Henrich & Henrich, 2007; Tomasello, 2016; Raihani, 2021). Depending on the social relationship in question, different evolutionary pathways to explaining such prosocial behaviours have been proposed. Kin selection has emerged as the dominant explanatory framework for explaining cooperation among biologically related individuals (Hamilton, 1964). In the context of interactions among non-relatives or strangers, reciprocity can generate stable patterns of support and sharing (Axelrod & Hamilton, 1981; Bowles & Gintis, 2011; Trivers, 1971). Reciprocal theories of cooperation turn on the idea that helping you or deferring to you on one occasion can be in my direct interest if by doing so I increase the likelihood that you will help or defer to me in the future. One way that costly prosocial behaviours such as helping, sharing, and comforting even among non-kin can pay off is if the roles of donor and recipient reliably alternate.

There is ample evidence that behavioural patterns of reciprocity occur not only in humans but in a wide variety of animal taxa. Among the most well-known cases are predator inspection among fish (Milinski, 1987; Milinski, Pfluger, Külling, & Kettler, 1990), particularly sticklebacks (*Gasterosteus aculeatus*) and guppies (*Poecilia reticulata*), blood sharing by vampire bats (*Desmodus rotundus*) (Wilkinson, 1984), and mutual grooming in impala (*Aepyceros melampus*) (Hart & Hart, 1992). As a whole, these studies present strong naturalistic evidence for reciprocal patterns of cooperation in animals. In recent years, there has been ongoing debate as to whether reciprocal cooperation in humans occurs in similar interactions and with the same flexibility and is based on the same psychological foundation as in other species, especially our closest animal relatives, chimpanzees (Brosnan & de Waal, 2002; Tomasello, 2016; Schweinfurth & Call, 2019). The question is: what mechanisms generate reciprocal patterns of cooperation in chimpanzees and humans?

Several mechanisms have been proposed to support reciprocity in primates. Emotional reciprocity is a mechanism that allows individuals to keep track of favours given and received, typically over longer time frames (Brosnan & de Waal, 2002; Jaeggi, De Groot, Stevens, & Van Schaik, 2013; Schino & Aureli, 2009; Tomasello, 2016). The idea is that dyadic social interactions elicit partner-specific emotional responses that in turn guide future behaviour towards that same partner. Recent work suggests that the mammalian bonding hormone oxytocin might play a role in such forms of emotional reciprocity (Crockford et al., 2013; Wittig et al., 2014).

Two other potential mechanisms are attitudinal reciprocity and calculated reciprocity. In attitudinal reciprocity, individuals associate a tag or an attitude with a cooperative partner based on their last encounter. Instead of having to memorize specific cases of past behaviour, a general tag such as “the partner was nice” is sufficient to elicit prosocial responses in the next encounter (Brosnan & de Waal, 2002; Jaeggi, De Groot, Stevens, & Van Schaik, 2013). Calculated reciprocity is based on careful mental scorekeeping of given and received favours. Due to the cognitive demands this entails, calculated reciprocity is thought to be rare in non-human primates (Stevens & Hauser, 2004; Schino & Aureli, 2009; but see by Dufour et al., 2009). Both attitudinal and calculated reciprocity are associated with short-term exchanges (Schweinfurth & Call, 2019), although, in principle, calculated reciprocity can also function over extended time periods.

To answer questions about underlying mechanisms, an important first step is to investigate in what interactions chimpanzee and human reciprocal cooperation occurs. Moreover, studying the developmental trajectories of different types of reciprocal exchanges in young children can provide further insights into the underlying cognitive capacities at play. Some of the most convincing evidence for reciprocity in chimpanzees comes from observational research in which cooperative exchanges are analysed over extended periods of time. This work has shown, for instance, that in the long-term, chimpanzees engage in both in-kind exchanges – e.g. when they exchange grooming for grooming (Gomes & Boesch, 2011), food for food (de Waal, 1989), or support for support (Gomes & Boesch, 2011) – as well as in exchanges across domains, as when they exchange meat for sex (Gomes & Boesch, 2009) or support for sex (Duffy, Wrangham, & Silk, 2007). Likewise, naturalistic observations suggest that human children start cooperating in reciprocal patterns from a young age onwards. Much like in the case of chimpanzees, this involves both in-kind exchanges, e.g. helping for helping, object for object, as well as exchanges across domains, e.g. helping someone in exchange for receiving an object (Fujisawa, Kutsukake, & Hasegawa, 2008).

The evidence for short-term reciprocity in chimpanzees is more mixed: while several studies failed to detect any evidence (Brosnan et al., 2009; Yamamoto & Tanaka, 2009; Jaeggi et al., 2013; Engelmann & Herrmann, 2016; Bueno-Guerra et al., 2019) and others found only weak evidence (Melis, Hare, & Tomasello, 2008, Engelmann, Herrmann & Tomasello, 2015), three recent studies reported positive findings (Benozio et al., 2023; Schmelz et al. 2017; Schmelz et al. 2020).

By contrast, young children consistently demonstrate short-term reciprocity from a preschool age onwards (Olson & Spelke, 2008; Kuhlmeier, Dunfield & O'Neill, 2014; Warneken, 2018; Zhang, Grocke & Tomasello, 2019; Grueneisen & Tomasello, 2017). Warneken and Tomasello (2013) reported that while children's first prosocial acts, at around 2 years of age, emerge spontaneously, children show first signs of contingent reciprocity in the third year of life. Supporting this general developmental picture, Vaish et al. (2018) found that young children selectively act generously toward past benefactors, especially when the benefactor had provided them with benefits intentionally. And House and colleagues (2013) present unequivocal evidence for short-term reciprocity in children at the age of 5.5 years. Around age 5, children also start to engage in future-oriented calculated reciprocity by strategically acting prosocially when others can reciprocate (Engelmann et al., 2013; Grueneisen & Warneken, 2022; Warneken et al., 2019) and this facility has been linked to children's developing skills for prospection and the willingness to delay rewards (Grueneisen et al., 2023).

A conclusion one could draw from these findings is that, while young children clearly show short-term calculated reciprocity from a young age, short-term exchanges in chimpanzees are rare. Instead, chimpanzee reciprocity is largely confined to long-term exchanges, presumably mediated by affective mechanisms (House et al., 2013; Jaeggi et al., 2013). But this conclusion seems premature for at least three reasons. First, studies on chimpanzee reciprocity have often relied on the ”prosocial choice task” in which individuals choose between one option that delivers rewards to themselves and to a partner and another option that delivers rewards only to themselves. However, the methods operationalizing this task often differed markedly between paradigms used with children and chimpanzees. For example, Brosnan et al. (2009) presented chimpanzees with a vertical barpull apparatus which allowed chimpanzees to choose one of two food distributions that could be accessed immediately after each round. House and colleagues (2013) used a set up introduced by Fehr et al. (2008), in which children selected one of two token distributions (presented on cards laying on the ground), which could be exchanged for actual rewards at the end of the experiment. Second, in some studies using the prosocial choice task, no evidence was presented that chimpanzees understood the experimental setup and the validity of the prosocial choice task has been questioned due to its high cognitive demands (e.g., Burkart & Rueth, 2013; Tan et al., 2015). Interestingly, the studies reporting positive evidence for short-term reciprocity in chimpanzees using the prosocial choice task implemented the strictest training procedures and comprehension checks (Benozio et al., 2023; Schmelz et al. 2017; Schmelz et al. 2020). Third, and most importantly, the prosocial choice task has failed to elicit prosociality in chimpanzees at baseline (Jensen et al., 2006; Silk et al., 2005), that is, in the absence of any prior interaction, so it is doubtful whether this is the most promising task for investigating reciprocal prosociality in this species.

By contrast, chimpanzees consistently show prosocial behaviour in helping paradigms. In numerous studies, chimpanzees have been shown to pick up out-of-reach objects (Warneken & Tomasello, 2006), open doors (Warneken et al., 2007), release food (Engelmann et al., 2015; Melis et al., 2011), or hand over tools (Yamamoto et al., 2009). Given the ubiquity of prosociality in these studies, helping paradigms might be particularly suitable for detecting the conditions under which reciprocal motives might enter chimpanzees' cooperative decision-making.

The goals of the current studies were twofold: First, we aimed to study prosociality in children and chimpanzees in a helping context using experimental tasks and study procedures that were as similar as possible and thus allowed for species comparisons. Second, we aimed to investigate whether both samples showed reciprocity, that is, whether subjects' prosocial actions toward a recipient were influenced by whether the recipient had previously behaved prosocially toward the subject.

Thus, in the current set of studies, we presented chimpanzees and children with nearly identical experimental setups using a novel version of an instrumental helping task. We tested 4-5-year-olds because children at this age have been shown to engage in reciprocal cooperation with peers (House et al., 2013; Grueneisen & Tomasello, 2017). Both species participated in the same three conditions. In the *helpful condition*, participants were given the opportunity to help a partner who had previously helped them. In the *unhelpful condition*, they could help a partner who had previously proven *unhelpful*. We had two main research questions. First, we asked whether chimpanzees and children demonstrate helping behaviour by comparing these two social conditions to a *non-social condition* in which the partner was absent. Second, by comparing the *helpful condition* to the *unhelpful condition*, we probed whether chimpanzees and children engage in short-term reciprocity and assist a previously *helpful* partner more than a previously *unhelpful* partner. To overcome methodological shortcomings observed in prior research, we implemented thorough apparatus familiarization phases as well as strict comprehension checks including criteria to assure task understanding before testing.

STUDY 1: CHIMPANZEES

Methods.

Participants

12 chimpanzees (*Pan troglodytes*, 5 females, 7 males; mean age = 12.67 years, age range = 5 to 24 years) participated in this study. Additionally, we trained chimpanzee *stooges* (4 females, mean age = 19.5 years, age range = 10 to 25 years), who acted reliably as either *helpful* or *unhelpful* depending on condition. All participants were unrelated. The chimpanzees lived in two stable groups (group 1 = 24 individuals; group 2 = 16 individuals) at the *Sweetwaters Chimpanzee Sanctuary* inside the *Ol Pejeta Conservancy* in Kenya. All chimpanzees were housed in natural, spacious outdoor and indoor enclosures with regular feedings, water ad-lib and additional daily enrichment provided by the caregivers and researchers. Table 1 depicts a list of chimpanzee subjects' and stooges' age, sex rank and housing group. Within 11 of 24 stooge-subject dyads, the stooge was the higher-ranking individual. 10 chimpanzees (Jojo, Judy, Max, Ajabu, Saidia, Oscar, Safari, Ndaronse, Alley, Victoria) from the housing group 1 did not participate in the experimental phase due to insufficient apparatus understanding (see subject familiarization for more details). For an evaluation of the sample using the STRANGE-framework (Webster & Rutz, 2020) please refer to the Supporting information [SI].

Table 1

Chimpanzee subject and stooge list, N = 12 including age, sex, rank and housing group; average individual rank derived from 8 keepers' questionnaires (3 categories: 1 = high-ranking; 2 = medium-ranking; 3 = low-ranking)

Name	**Age**	**Sex**	**Rank**	**Housing group**
		Subjects		
Mwanzo	16	female	1.9	1
Amahirwe	14	male	1.6	1
Zee	12	male	2.1	1
Julia	12	female	2.9	1
Alikaka	10	male	1.6	1
Edward	10	male	2.4	1
Mary	9	female	2.0	1
George	9	male	2.3	1
Akela	24	female	1.8	2
Uruhara	23	male	1.6	2
Jane	8	female	3.0	2
Roy	5	male	3.0	2
		Stooges		
Cheetah	25	female	2.1	1
Eva	10	female	2.4	1
Amisero	23	female	1.2	2
Tess	20	female	1.6	2

Ethical statement

All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. Chimpanzees voluntarily participated in the study and were never food or water deprived. All research performed was in accordance with the recommendations of the Weatherall report (Weatherall et al., 2006). The ethics committee at the Sweetwaters Chimpanzee Sanctuary (the Sanctuary board members and the veterinarian) approved the experimental study procedure. All research at the sanctuary was non-invasive and strictly adhered to the legal requirements of Kenya (NCST, KWS; Research permit number issued by the National Council for Science and Technology: NCST/RRI/12/1/BS011/220). No medical, toxicological or neurobiological research of any kind is conducted at Sweetwaters Chimpanzee Sanctuary. The chimpanzees were tested in their sleeping/observation rooms.

Materials

We presented chimpanzees with a modified version of a tug-of-war-apparatus from a study on monopolization and turn-taking by Knofe et al. (2019). The apparatuses were positioned in two opposing rooms and consisted of four main items (see Figure 1): Two reward boxes filled with mashed banana, a rope connecting two stick-tools for foraging and a clamp for the rope with an attached release-button. One reward box was placed in the left room. An identical reward box was positioned in the opposite room. Each reward box could be accessed by poking with a stick-tool. Both tools were connected by a rope too short for simultaneous tool-use. Only in the left room, we added a clamp with a release-mechanism for the rope. This mechanism prohibited access to the right stick by clamping a knot tied into the rope. At the same time, the left tool was fully functional. The clamp could be released by pressing the attached release-button. This mechanism allowed the following sequence: Firstly, poking was possible only on the left side, since the rope was clamped. Secondly, pressing the release-button removed the clamp in the left room. Thereupon, the released rope could be pulled towards the right room allowing subsequent tool-use by the right individual. Once the button was pressed and the rope was pulled into the right room, the tool could not be used in the left room any more. Additionally, the functionality of the release-button could be manipulated by inserting a screw, a feature necessary in the two test conditions. Throughout the study, the blocking-mechanism and the release-button remained in the left room. For details on the apparatuses and materials please refer to the SI.

---------------------------------------Figure 1---

Design

After passing a set of criteria in the familiarization phase, all subjects were tested in a within-subject design in two test conditions (*helpful*, *unhelpful*) and one control condition (*non-social control*). In the *helpful* and *unhelpful* test conditions, the subjects interacted with a trained stooge. In the *non-social control*, the subjects were tested alone. Each condition consisted of ten trials. The *helpful* and *unhelpful* conditions were tested in five to six test sessions, each consisting of one to three trials, depending on subjects' motivation. The *non-social control* condition was split in two blocks each consisting of five trials. The first block was conducted before and the second block after the test conditions (ABCA-design). Each block was tested in one to two test sessions consisting of two to five trials (see table S2 in the SI for individual and average number of sessions per condition). The number of trials per session was influenced by subjects' motivation, keeper recommendations and force majeure. The test conditions were counterbalanced across subjects. Embedding the test conditions (B and C) into the *non-social control* (A) was designed to highlight the difference between the two social test conditions and the control and to gain a measure of subjects' baseline tendency to press the release-button (even when this did not result in a benefit for a partner).

The two main questions were whether subjects would help a partner (comparing the social *helpful* and *unhelpful conditions* with the *non-social control*) and if so, whether they engage in short-term reciprocity (comparing the *helpful* with the *unhelpful condition*).

Each trial in the two test conditions consisted of two parts. In the *experience phase*, the stooges either always (*helpful condition*) or never (*unhelpful condition*) helped the subject by pressing the release-button. Consequently, the stooges either ate their own food and then allowed subjects access to their reward box (*helpful condition*) or, in stark contrast, simply ate and left the apparatus without pressing the release-button (*unhelpful condition*) leaving the subject empty-handed. The *experience phase* was immediately followed by the *reaction phase,* the actual test situation, in which the subjects' reaction to the stooges' behaviour was assessed.

Procedure

Subject familiarization

All subjects were individually introduced to the set up. Subjects needed to pass eight criteria to demonstrate their understanding of the apparatuses. Table 2 depicts an overview of all familiarization steps: description of each step including its learning goal, criterion, trial time, average number of trials needed to meet criterion, number of dropped out subjects and provided apparatuses.

To meet the most important apparatus-understanding criterion (familiarization step 7), subjects had to (1) use the tool for foraging in the left room, (2) press the release button, (3) change into the right room and (4) use the previously released tool in the right room in three consecutive trials. Here subjects used all apparatuses provided in the subsequent test conditions (except for the open door between the right and left room). Ten chimpanzees did not meet the basic requirements for the experiment. These individuals lost interest in the apparatuses or lacked apparatus understanding in the familiarization steps and where therefore dropped from the experiment. For a more detailed, formulated description of the familiarization steps, individual performance in each familiarization step including the dropped out individuals please refer to the SI.

Table 2

Familiarization steps of Study 1

Step	Description / learning goals	Criterion	Trial time [min]	Average trials needed	Drop-outs	Provision of apparatus-item		
						Release-button	*Left/ Right reward box*	*Open door connecting rooms with boxes*
1	Poking I: With loose stick-tool for mashed banana, left room / tool use	3 consecutive trials of tool use	3	3	5	no	yes/ no	no
2	Poking II: With stick-tool attached to rope, left room; tool use / rope-habituation	3 consecutive trials of tool use	3	3.08	1	no	yes/ no	no
3	Poking III: With stick-tool attached to rope in the left room, room change, poking in the right room / use of right room and box	3 consecutive trials of tool use in both rooms	6	3.75	1	no	yes/ yes	yes
4	Button intro: Button-pressing for an apple slice / button-habituation, learning to press	6 consecutive, complete button-presses in 1 min	5-20	1.08	1	yes	no/ no	no
5	Release I: Button-pressing by human (left) releases tool in the right room / observing button-pressing consequences	10 observations and consecutive tool uses	1.5	10	-	yes	no/ yes	no
6	Release II: Button-pressing releases tool in right room / understanding button-pressing consequences	3 consecutive trials button pressing; room change; tool use	6	5.75	1	yes	no/ yes	yes
7	Release III: Poking left, button-pressing, room change, poking on the right / understanding setup by combining all steps	3 consecutive trials tool use (left); button pressing; room change; tool use (right)	6	6.17	1	yes	yes/ yes	yes
8	Introduction of distractors against button-play / decrease button-pressing for unwanted reasons	3 consecutive trials tool use; leave tool; avoiding button for 1 min	6	10.67	-	yes	yes/ no	no

Stooge training

The two test conditions were characterized by the contrasting behaviours of the either *helpful* or *unhelpful* stooges. After one stooge was assigned to a subject and condition, they remained the partner for this subject in the same condition throughout the experiment. That is, subjects experienced 10 trials per condition with the same stooge.

To make sure that the two chimpanzee stooges acted in reliable ways, each stooge underwent individualized training. The *helpful* stooges were trained to exhibit the following procedure in the *experience phase*: (1) enter the left experimental room; (2) feed from their reward-box using the stick-tool; (3) press the release-button when the preventing screw is removed and a slice of apple is shown (hidden from the subjects) by the experimenter (4) leave the experimental room to the neighbouring room and receive the apple slice thrown into the room. The *unhelpful* stooges were trained to exhibit the following procedure: (1) enter the left experimental room; (2) feed from the reward-box with the stick-tool; (3) ignore the button, blocked by the screw (4) leave the experimental room to the neighbouring room and get an apple slice thrown into it. Please refer to the SI for the details about the stooge training.

Two stooges were tested in both test conditions (Cheetah and Eva), that is, they first acted *helpfully* or *unhelpfully* with four subjects (randomly assigned), then received retraining and acted in the opposite role with another four individuals. Two additional stooges (Amisero and Tess) acted exclusively *helpfully* or *unhelpfully* with four partners (four additional subjects for whom these stooges were supposed to be retrained for the opposite role did not reach criterion in time to be tested). The four latter subjects for whom Stooge ID was not balanced between conditions did not behave differently than the other subjects, and all reported results hold when we control in the analysis for whether or not Stooge ID was fully balanced (see the SI for details).

Test phase

In the *helpful* and *unhelpful condition*, trials began with an *experience phase*. During this phase, the subject was in the right room without access to the tool. The stooge was in the left room with access to the tool and release-button. First, the subject watched as the stooge fed from the box. In the *helpful condition*, following feeding, the stooge pressed the release-button. Thereupon, the subject could pull the rope with the tool into the right room and feed. The *experience phase* ended once the subject had emptied her box. In the *unhelpful condition*, the stooge fed at her box for approximately one minute. Afterwards, the stooge left the room ignoring the button, which left the subject empty handed. In addition, as an extra safety, the release-button was secretly sabotaged (subjects were not cognizant of the hidden dysfunction) ensuring no accidental helping by button pressing in the *unhelpful condition*. However, *unhelpful* stooges ignored the button reliably.

The *experience phase* was immediately followed by the *reaction phase*. Subject and stooge exchanged their positions (with the subject now in the left room provided with the release-button). The time between the two phases was kept to a minimum and never exceeded five minutes. The procedure of the *reaction phase* resembled the *experience phase* with the release-button being always functional, though. A trial ended one minute after the subjects left their tool unattended for longer than ten seconds (irrespective of whether they came to feed or pressed the release-button).

In the *non-social control* condition, the subjects were tested without a stooge partner. There was no need of an *experience phase*, and subjects participated only in the *reaction-phase*.

Coding and analyses

All trials were video-taped. The main binomial dependent variable *button pressing* was coded from video. Trials were assigned a “1” if subjects pressed the release button in the *reaction phase* and a “0” if they did not. The coding time frame was set to 1 minute: 10 seconds after the subjects left their tool and box unattended and untouched, the subjects had 50 additional seconds to potentially press the button. If subjects left the tool for less than 10 seconds and came back just holding the tool or feeding again, the timeframe was reset until the subject left the apparatus for more than 10 seconds. 20% of the data were coded independently by a research assistant, who was unaware of the study design and hypotheses. Inter-rater reliability for button-pressing (Cohen’s kappa, unweighted) was very good ($K = 0.952$, $N = 72$). For the complete data see the supporting information.

In a preliminary analysis, we first fit a Generalized Linear Mixed Model (GLMM, Baayen, 2008) to inspect if condition order (the order in which subjects completed the three conditions) and session number affected button presses. For our main analysis, we fit a GLMM with binomial error structure. The dependent variable was whether or not subjects pressed the button on a given trial. The test predictors were condition (*helpful*, *unhelpful*, *non-social control*), trial number (1-10) within condition, and their interaction. We included the random effect of subject ID to account for the fact that subjects contributed multiple data points. We also included the random slopes components of condition and trial number nested within subject ID.

Analyses were conducted in R (R Core Team, 2022) using the function 'glmer' of the R-package lme4 (Bates et al., 2014). We first compared the full model described above with a null model not including the test predictors but retaining the random effect and random slopes using a likelihood ratio test. We ran hypotheses-driven tests of individual predictors using likelihood ratio tests (*drop1* command of *lme4* package) only after this full-null model comparison revealed a significant effect of the test predictors combined (this approach has been shown to reduce Type 1 error rates by preventing multiple testing issues, Forstmeier & Schielzeth, 2011).

Results

The preliminary analysis revealed that condition order and session number did not have a significant effect ($\chi^2(1) = 0.67$, $p = .412$ and $\chi^2(1) = 0.79$, $p = .373$, respectively). These factors were thus dropped from the analysis. The main analysis indicated that the predictors condition, trial number, and their interaction combined significantly affected subjects' button presses (full-null-model comparison: $\chi^2(5) = 17.42$, $p = .004$). Further analyses revealed a significant interaction between condition and trial number ($\chi^2(2) = 8.63$, $p = .013$). We followed up the interaction by investigating the effect of trial number in the three conditions separately. This revealed that, in the *helpful* condition, subjects were less likely to press the button over trials ($\chi^2(1) = 9.70$, $p = .002$): subjects started out pressing the button at relatively high levels (50% in trials 1-5) but became gradually less likely to do so (23% in trials 6-10). Trial number did not have a significant effect in the *unhelpful* condition ($\chi^2(1) = 0.00$, $p = .986$, overall mean = 30%), or in the *non-social* condition ($\chi^2(1) = 0.08$, $p = .783$, overall mean = 12.5%; see figure 2).

We also tested the effect of condition in the first five trials and in the last five trials. This revealed that condition significantly affected button presses in first five trials ($\chi^2(2) = 27.96$, $p < .001$). Specifically, subjects were more likely to press the button in the *helpful* than in the *non-social* condition ($p < .001$) and in the *unhelpful* condition than in the *non-social* condition ($p = .009$) and in the *helpful* than in the *unhelpful* condition ($p = .009$). By contrast, condition did not significantly affect button presses in the last five trials ($\chi^2(2) = 2.30$, $p = .316$).

--Figure 2--

When disregarding the interaction, there was a significant main effect of condition ($\chi^2(2) = 6.32$, $p = .042$). Follow-up analyses revealed that, pooled across all trials, there was no significant difference between the *helpful* and *unhelpful* condition ($p = .311$), but button presses were significantly more common in the *helpful* condition than in the *non-social* condition, ($p = .003$) and in the *unhelpful* condition than *non-social* condition ($p = .068$). See the SI for details.

STUDY 2: CHILDREN

Methods.

Participants

36 pre-school children (N = 36; mean age = 5.05 years; age range = 4.74 to 5.26 years; 18 girls, 18 boys) were tested in day-care centres in a medium sized German city. In the test conditions subjects were tested with a same-gender peer stooge (total 21 stooges; mean age = 5.64 years; age range = 4.84 to 5.97 years; 12 girls, 9 boys). These stooges were recruited from the same day-care centres as the participants but from different groups. The interacting partners were not related and mostly did not know each other (4 dyads played once before). 18 trained stooges interacted with just one subject and 3 very motivated stooges were tested with a second participant. 24 stooges broke the defined rules while interacting with 32 subjects and had to be excluded from the study. We recruited new stooges and tested new subjects. For details on the dropout criteria and our ethical statement see the SI.

Materials

The children were presented with an apparatus similar to the one used with the chimpanzees in Study 1. The tasks were fitted for the specific preferences of the two species. Instead of poking for mashed bananas with a stick-tool as the chimpanzees in Study 1, children dipped a stamp into an ink-filled tube for stamping on a sheet (see Figure 3). Each child on each side had such a stamp which were connected by a short rope. The stamps were needed for applying marks into five printed circles on provided sheets. Before the stamps could be used, they had to be dipped into a tube containing an ink pad. As in Study 1, the rope could be blocked by clamping a knot tied into it. Blocking the rope prohibited stamping on the right side since the rope was too short to utilize the stamp. Thus, only the left individual could use the stamp. Pressing the release-button attached to the clamp on the left side, released the rope now allowing stamping on the right side. Therefore, the individual on the left side could start to apply marks and allow stamping on the right side by pressing the button. The children sat on pillows on each side of the apparatus and were divided by a transparent barrier. For protection against unintended stamps, the participants wore coats. Walkie-talkies for experimenter- stooge communication allowed the experimenter to remind the stooges live (child with cap, see Figure 3) how to act during trials. In-ear headphones prevented subjects from listening in (subjects were told stooges used a hearing aid). For details on the materials see the SI.

--Figure 3--

Design

As in Study 1, two test conditions (*helpful*, *unhelpful*) and one control condition (*non-social control*) were conducted. Each participant was tested in one condition consisting of five trials (one session). The total number of 5 trials was unknown to the subjects. Per condition, 12 children were tested (N = 36). In the test conditions, the subjects were tested with stooge partners who acted like subjects too (but were actually secretly trained collaborators). The stooges' cooperation with the experimenters was revealed to the subject at the end of the experiment. The two test conditions included two phases. In the *experience phase*, the stooge acted either *helpfully* or *unhelpfully* after stamping, allowing the subject to either access the stamp or not. Prior to the next phase, stooge and subject changed sides. In the *reaction phase*, the actual test, the subject could stamp first and potentially use the release-button to help the stooge. In the *non-social control* the subjects were tested alone (and did not undergo an *experience phase*) and remained on the left side of the apparatus.

Procedure

Subject familiarization

Before the test phase, subjects were individually presented with a single familiarization trial for understanding the set-up and task. At the beginning, the subject was introduced to the previously trained stooge outside the experimental room. Subjects were told that they would play together later and that the other child already received the introduction. The Walkie-talkie of the stooge was explained as a hearing aid. Subjects were then individually presented with a single familiarization trial before the test phase. The setup was the same as in the experimental phase but without a partner. Thus, each side of the apparatus contained a stamping sheet, a tube with an ink pad and a stamp interconnected by a rope. The release button for the clamped rope was on the left side. Subjects were taught how to apply marks. The following two rules were established (which the participants were reminded of before every trial): Use fresh ink for every new mark; apply just one mark to each circle printed on the sheet. Then, subjects were shown the blocked rope preventing access to the stamp on the right side. The release-button on the left side was explained and demonstrated. After the child pressed it herself, she was allowed to apply marks on the right side.

Quickly, all tested children learnt applying marks since using stamps was part of the kindergartens' play routines. We always asked nursery school teachers in advance if children were familiar with rubber stamps. This was always the case.

Children understood the apparatus intuitively. Button-pressing seemed an easy task for the children, too. No additional, individual training was necessary.

Stooge training

All stooges were trained individually to guarantee reliable behaviour in the *experience* and *reaction phase* of the *helpful* and *unhelpful condition*. For details on the procedure and sequence of the stooges' act see the paragraph below and the SI.

Test phase

During the test phase, at the beginning of all trials, the rope was blocked (stamping in the right room was therefore not possible), and sheets were provided. In the *experience phase*, the participants sat down, stooges on the left side, subjects on the right side. The experimenter repeated the rules (use fresh ink and just one mark per circle) and left the room; the *experience phase* started. The subject watched the stooge apply marks. In the *helpful condition*, the stooge pressed the release-button after finishing her sheet. Thereupon, the subject could pull the rope to the right side and use the stamp. In the *unhelpful condition*, the stooge applied the marks, put the stamp away and ignored the button for 30 seconds. When subjects requested button-pressing, stooges were instructed to stare to the ground, shake their heads or state "No.". In the *helpful condition* the *experience phase* ended once the subject finished her sheet and in the *unhelpful condition* after the subject had remained empty handed for 30 seconds. Before the *reaction phase,* the experimenter changed the sheets. Subjects and stooges exchanged their position (with the subject now on the left side provided with the release-button). The time between the two phases was kept to a minimum and never exceeded five minutes. The release-button was always functional in the *reaction phase*. The procedure of the *reaction phase* was identical to the *experience phase*, with one change: a trial ended one minute after the subjects left the stamp unattended for longer than ten seconds (irrespective of whether they used the stamp or pressed the release-button). In the *non-social control* condition, subjects started in the *reaction-phase* (there was no *experience phase* as there was no partner).

Coding and analyses

The coding procedure was analogous to Study 1. 20% of the child data was coded independently by a research assistant who was unaware of the study design and hypothesis. Inter-coder reliability for button-pressing (Cohen's kappa, unweighted) was very good ($K = 0.943$, $N = 36$).

We aimed to follow the same general analytic approach as in Study 1. However, due to complete separation issues – all children in the *helpful* condition pressed the button on all trials – conducting GLMMs in our main analysis became unfeasible. We therefore pooled subjects' button presses over the five trials and conducted Kruskal Wallis *H*-tests and Mann-Whitney *U*-tests to investigate condition effects. We used GLMMs to separately test for trial effects in the *unhelpful* and *non-social* condition.

Results

The Kruskal-Wallis H-test revealed a difference in button-pressing between conditions ($H = 23.52$, $df = 2$, $p < 0.001$), see Figure 4. Children pressed the release-button significantly more often (and thus helped more) in the two social test conditions than in the *non-social control* condition (Mann-Whitney *U*-tests, *helpful–non-social*: $U = 0$, $n_1 = n_2 = 12$, $p < 0.001$; *unhelpful–non-social*: $U = 25$, $n_1 = n_2 = 12$,

p = 0.003). Children also pressed the release-button more in the *helpful* condition than in the *unhelpful* condition (*helpful–unhelpful*: U = 36, $n_1 = n_2$ = 12, p = 0.014). In fact, all 12 children in the *helpful* condition pressed the release-button on all 5 trials. In the *unhelpful* condition, children pressed the button in 62% of the trials and children's button-presses decreased slightly over trials ($\chi^2(1) = 3.59$, p = .058). In the *non-social* condition, the button was presses at consistently low levels (6% on average) with no significant changes of over trials ($\chi^2(1) = 0.26$, p = .609). See the SI for details.

--Figure 4--

DISCUSSION

We investigated chimpanzees' and children's immediate reactions towards a conspecific partner who had either previously helped them or not. We asked two main questions. One, do chimpanzees and children show prosociality in a helping task, i.e. are they more likely to press a button if by doing so they can benefit a partner? Two, does the level of prosociality vary as a function of the partner's previous behaviour? That is, are chimpanzees and children more likely to help a partner who has previously acted *helpfully* compared to a partner who has previously acted *unhelpfully*? We found evidence for prosociality in both species: chimpanzees and children engaged in a low-cost action to help a partner obtain a tool for accessing a reward.

We also found clear evidence for short-term reciprocity in children: participants calibrated their level of prosociality to their partner's previous behaviour and acted more prosocially towards a *helpful* partner and less prosocially towards an *unhelpful* partner. Chimpanzees showed a similar pattern, but the picture was less straightforward: Over the course of the first five trials, chimpanzees acted more prosocially when interacting with *helpful* partners; over the course of the last five trials, they did not differentiate between *helpful* and *unhelpful* partners.

The current results replicate and extend previous work on chimpanzees' and children's prosocial tendencies (Melis, 2017; Warneken & Tomasello, 2009). The novel release-button prosociality task used in the current study detected converging evidence for chimpanzees' and children's prosocial motivations. Both species pressed the release-button more frequently in the social test conditions (*helpful* and *unhelpful*) compared to the *non-social control* condition. Note that we presented chimpanzee subjects with a particularly demanding test of prosociality. Our design required subjects to switch twice between a social and a non-social setup (first from a non-social to a social situation, and then back to a non-social setting; ABCA design). Such switches between conditions can be hard for chimpanzees and often result in carry-over effects. Yet, the act of button-pressing was observed rarely in the control trials before and after the two social test conditions, making the difference between control and test conditions even more remarkable.

In our experiment, children were clearly influenced by their partner's previous behaviour. In fact, when children were given a chance to help a previously *helpful* partner, their probability of doing so was 1. In other words, the twelve children that participated in this condition acted prosocially towards the *helpful* partner in a 100% of trials. The current results cannot answer the question whether – compared to a helping baseline – children's helping increased in the *helpful condition*, decreased in the *unhelpful condition*, or showed both patterns. Thus, future work on children's reciprocity should include a baseline helping condition (without a previous interaction) to investigate whether, in a helping context, children show a negativity bias (Vaish, Grossmann, & Woodward, 2008) and are more likely to express negative reciprocity (responding unhelpfully to an antisocial partner) than positive reciprocity (responding helpfully to a prosocial partner). Note that in 38% of the trials in the *unhelpful* condition children ignored the request for pressing the release button for over one minute. This behaviour might be indicative of negative reciprocity such that children withheld cooperation from previously uncooperative actors.

We also found some evidence for short-term reciprocity in chimpanzees. Specifically, in the *helpful* condition, subjects started out choosing prosocially at relatively high levels but then their prosociality decreased over time. By contrast, prosocial choices remained constant at medium and low levels in the *unhelpful* and *non-social* conditions, respectively.

These results are in line with recent studies reporting reciprocal cooperation in chimpanzees using the prosocial choice task (Benozio et al., 2023; Schmelz et al., 2017; 2020) and suggest that, rather than being detectable only over long time horizons, chimpanzees also show reciprocity in the short-term, at least in some situations. A potential mechanism underlying these results is attitudinal reciprocity (Brosnan & de Waal, 2002). That is, subjects may have associated a positive attitude or tag to helpful partners (e.g., "this individual was nice"), which in turn elicited helping behaviour in subsequent interactions. It is unclear, however, why attitudinal reciprocity should lead to a decrease in prosociality over time as observed in the *helpful* condition.

The fact that chimpanzees helped more in the *helpful* than in the *unhelpful* condition only in the first five trials is also surprising from an emotional bookkeeping perspective (Schino & Aureli, 2009), since the partner's continued prosociality should have strengthened subjects' partner-specific emotional responses. Subject's prosocial behaviour towards helpful partners thus should have increased rather than decreased over time. A potential explanation is that, at first, chimpanzees in the *helpful* condition returned the partner's favour in an attempt to secure the partner's prosociality in subsequent rounds. Over time, they may have realized that their partner proved *helpful* irrespective of their own behaviour, thus resulting in a drop in prosocial choices in later trials. What speaks against this explanation, however, is that this form of calculated prospective reciprocity requires sophisticated planning abilities and is not shown by human children until age 5 to 7 in comparable experimental settings (Grueneisen et al., 2023; Warneken et al., 2019). Hence, while evidence for short-term reciprocity has been accumulating, the exact psychological mechanisms underlying this phenomenon are not yet well-understood and deserve further investigation. It should be noted, however, that when pooled over the whole experiment, chimpanzees did not help more in the *helpful* than in the *unhelpful* condition and we did not have strong a priori predictions regarding a trial effect. The current findings thus need to be interpreted with caution.

On a methodological level, we contend that the helping context is highly suitable to study reciprocal prosociality in chimpanzees. Chimpanzees show helping behaviour in several experimental tasks, including the current one. Compared to the prosocial choice task, which is cognitively demanding and requires subjects to infer conspecifics' preferences from the payoffs alone, chimpanzees flexibly decode others' intentions in helping paradigms and provide targeted prosociality accordingly (Engelmann et al., 2015; Melis et al., 2011; Warneken et al., 2007; Yamamoto et al., 2009). To discern the extent and flexibility of chimpanzees' reciprocal cooperation, it would be useful to implement similar reciprocity manipulations using different helping tasks (e.g., opening doors, handing out-of-reach objects, delivering food).

In the past, the lack of rigorous comprehension checks has often compromised the interpretability of negative findings. Indeed, experimental studies which found evidence for reciprocal cooperation in chimpanzees (Benozio et al., 2023; Schmelz et al., 2017; 2020) implemented extensive training procedures and only included subjects who clearly demonstrated task understanding. The current study further emphasizes the importance of ensuring task understanding prior to conducting systematic tests of reciprocity and prosociality. Subjects received an extensive familiarization in which 8 criteria had to be met to ensure complete understanding. In children, breaking down the familiarization into several steps with separate criteria was not necessary since we could use language to explain the set-up. Within a single trial all children proved proficient at (1) using the stamp on the left side (by finishing their stamping pattern), (2) pressing the button autonomously which released the tool on the right side (3) changing sides and (4) using the stamp on the right side. Thus, before entering the experimental phase both species were prepared and sufficiently familiarized.

The current set-up, which only allows one individual to use the tool at a time, might have been interpreted as a limited resource problem. As recent studies on complex resource dilemmas demonstrated, chimpanzees choose monopolization over prosocial strategies (Koomen and Herrmann, 2018b; Knofe et al., 2019, Koomen and Herrmann, 2018a). However, the in-depth familiarization with the set-up and the finding that chimpanzees pressed the button more in the social than in the non-social conditions makes this alternative interpretation unlikely.

There were several other minor methodological differences between the chimpanzee and children study regarding the rewards (banana versus stamps), the number of trials per condition (10 versus 5), and the number of trials per session (1-2 versus 5). We also used a within-subjects design in the chimpanzee study and a between-subjects design in the children study and stooges were trained slightly differently. The rewards were chosen to keep motivations high in both study populations throughout the experiment and to correspond to subjects’ everyday experiences. Chimpanzees are highly food-motivated and poke for various food items as part of their natural behavioural repertoire. The children in the current study all used stamps as toys in their playing routine (we asked kindergarten teachers of all groups). Using non-food rewards for children further avoided issues around hygiene and dietary allergies.

We chose to administer 10 trials in study 1 since, in a previous study (Knofe et al., 2019), chimpanzees were motivated to poke for food for 10 trials when tested over several sessions (1-2 trials per session). When piloting study 2, we found that children's interest in stamping gradually decreased after 5 trials. Testing multiple sessions was not practicable due to the kindergartens' changing schedules and the increased risk of either the subject or the stooge being absent the following day. We therefore decided to reduce the total amount of trials to 5 in one session. Interestingly, both species discriminated between partners' *helpful* and *unhelpful* behaviour already in the first half. Certainly, it would be interesting to know if children would gradually reduce reciprocating in the *helpful* condition like chimpanzees in study 1 if we would have increased the total number of trials. But this scenario is rather unlikely since all children reciprocated *helpful* behaviour with the probability of 1 in all 5 trials. To maximize the number of subjects in study 1, all chimpanzees who passed the 8 familiarization steps were tested using a within-subject design in all conditions. In study 2, as explained above, testing subjects on consecutive days was impracticable and we therefore chose a between-subject design.

The differences in the stooge-subject relationship reflect the actual living conditions of chimpanzees and human children. In study 1, we always chose subjects and stooges from the same chimpanzee groups since there was no alternative and prosocial acts mostly occur between individuals from stable groups in the wild (e.g. de Waal, 1989). Thus, all chimpanzees knew each other. Similarly, in study 2, human subjects and stooges knew each other, too. Even though subjects and partners were not from the same kindergarten group, they were from the same kindergarten and thus familiar with one another. However, it is worth noting that only 4 of the 24 child dyads explicitly stated that they had played together previously.

We cannot rule out that the stooge behavior felt less consequential to chimpanzees given their long-lasting history with the stooge partners. However, the fact that subjects differentiated between helpful and unhelpful partners, at least in the first five trials, indicates that chimpanzees also considered their partners' actions in the experiment. Moreover, subjects could not repay the stooges' behaviors in different currencies such as grooming, food, support, or aggression since they never interacted directly with the stooge between the *experience* and *reaction phases*. Providing or withholding help was thus the only way to reciprocate the partner's action.

In study 1, chimpanzee stooges were trained to exhibit the same procedure in every trial. They learned they would receive an apple slice after pressing or ignoring the button in the experience phase and after the reaction phase. The experimenter was in close proximity to the stooge (but hidden from the subject by a thick curtain) and with an apple in his hand to focus the stooge on the high-value reward rather than on communicating with the subject. Using this procedure, begging behaviors could not be ruled out, but were not detected (a systematic analysis was not possible as stooge behavior was not videotaped during the reaction phase). In study 2, stooges' communication was tightly controlled and under strict inspection. This measure proved to be imperative especially in the *unhelpful* condition. Though stooges received detailed guidelines beforehand and clear instructions via walkie talkie (inaudible for the subjects) during the experiment to ensure they do not give away their role in the experiment, 32 dyads had to be excluded from the experiment because 24 stooges revealed their play act and told the subjects about their instructions not to press the button (for details on the drop out criteria see the SI). Without these strict instructions which limited interaction and communication towards the subjects, drop-out rates would have been even higher. This is an interesting fact in itself, since many stooges decided to disregard the experimenter's instructions rather than leaving the *helpful* behaviour of the subjects unreciprocated.

The chimpanzee subjects included in the study live in a naturalistic environment with a rich social life, in many ways comparable to wild populations, increasing our confidence in the generalizability of the current findings (yet, some effects of rearing history or life in captivity cannot be ruled out; see Webster & Rutz, 2020). Although, the chimpanzees had prior testing-experience, the current set up was completely new to them. Since the chimpanzees were familiarized with the new setup, carry over effects from former studies are rather unlikely.

An interesting question is whether dyad composition affects reciprocity in chimpanzees. For instance, dyads high in tolerance (e.g. Melis, Hare & Tomasello, 2006) or friendships and trust (e.g. Engelmann & Herrmann, 2016) might show higher levels of direct reciprocity. On the other hand, intra-dyadic tolerance could just as well allow tolerance of short-term imbalances, which are typically compensated over a longer time period, thus reducing short-term reciprocity. Testing for short-term reciprocal interactions in composed dyads respective their social hierarchy, tolerance and friendship could be an interesting topic for future research.

Reciprocity provides a pathway for cooperation to evolve among pairs of nonrelatives. The exchange of goods and services in reciprocal interactions can take place over longer time frames or over short time frames. While the current findings provide some evidence that chimpanzees might engage in short-term reciprocity, several experimental and observational studies suggest that chimpanzees' reciprocal tendencies are expressed primarily in long-term cooperative relationships (Jaeggi et al., 2013; Gomes & Boesch, 2009; Gomes, Mundry, & Boesch, 2009). Young children, on the other hand, form not only long-term reciprocal relationships, but have also consistently been shown to match a partner's prosocial or antisocial behaviour in the short term. The current study confirms that, compared to chimpanzees, children show short-term reciprocal cooperation at substantially higher levels. A key to understanding this species' difference might lie in the species' respective social ecology: chimpanzees live in small-scale and stable social groups that are characterized by repeated encounters within fission fusion dynamics (Couzin, 2006), making long-term balancing mechanisms ideal. Young children, in addition to forming such long-term relationships, are also exposed to more fleeting encounters with relative strangers, requiring more short-term balancing mechanisms (Seabright, 2004).

In summary, the novel helping task used in the current study revealed clear evidence for prosocial motivation in both chimpanzees and human children. Children further showed a strong tendency to engage in short-term reciprocity by helping a helpful partner and withholding help from an unhelpful partner. In chimpanzees, the evidence was more mixed with subjects showing a similar pattern as children only in the first half of the experiment after which they treated helpful and unhelpful partners equally. These findings confirm the notion that short-term reciprocity is more common in humans. However, they also add to a growing body of evidence showing that chimpanzee reciprocity is not confined to long-term exchanges and occurs in some short-term interactions too.

DATA AVAILABILITY STATEMENT

The data associated with this manuscript was uploaded as part of the Supplementary Information.

819

FIGURE LEGENDS

Figure 1. Set up study 1; Tools connected by a short rope were needed to access the identical reward-boxes filled with mashed banana. Firstly, the rope could be accessed in the left room only; simultaneous tool use was impossible. The rope was restricted by a clamp preventing the individual in the right room from pulling and using the tool. The individual in the left room could help by pressing the attached release-button, which removed the clamp and thereby released the rope allowing the other participant to use the tool. The tool was inaccessible for the left individual once pulled to the right room.

Figure 2. Probability of button presses over the 10 trials of the experiment in the helpful, unhelpful, and non-social condition. Scattered dots represent individual data points. Solid lines are fitted regression lines (binomial logistic regressions), shaded regions indicate 95% confidence intervals.

Figure 3. Set up study 2; Stamps connected by a short rope could be used to mark sheets after being dipped into tubes containing ink pads; simultaneous use of the stamps was impossible. The rope was restricted by a clamp and kept the individual on the right side from using the tool. The individual on the left side could help by pressing the attached release-button allowing the other participant to use the stamp. Experimenters could secretly communicate with the stooge (left child with cap) using a Walkie-talkie. The scene depicts the experience phase of the helpful condition at that time the stooge just finished stamping and is pressing the release-button for the subject (right child).

Figure 4. Probability of button presses over the 5 trials of the experiment in the helpful, unhelpful, and non-social condition. Scattered dots represent individual data points. Solid lines are fitted regression lines (binomial logistic regressions). Due to complete separation issues calculating confidence interval was not feasible.

REFERENCES

Amici, F., Aureli, F., Mundry, R., Amaro, A. S., Barroso, A. M., Ferretti, J., & Call, J. (2014). Calculated reciprocity? A comparative test with six primate species. *Primates, 55*(3), 447-457. doi:10.1007/s10329-014-0424-4

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science, 211*(4489), 1390-1396.

Bates, D., Maechler, M., Bolker, B. & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-7, http://CRAN.R-project.org/package=lme4.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.

Benozio, A., House, B. R., & Tomasello (2023). Apes reciprocate food positively and negatively. *Proceedings of the Royal Society B: Biological Sciences*, 290, 20222541.

Bowles, S., & Gintis, H. (2011). *A Cooperative Species: Human Reciprocity and Its Evolution*: Princeton University Press.

Brosnan, S. F., & de Waal, F. B. M. (2002). A proximate perspective on reciprocal altruism. *Hum Nat, 13*(1), 129-152. doi:10.1007/s12110-002-1017-2

Brosnan, S. F., Silk, J. B., Henrich, J., Mareno, M., Lambeth, S., & Schapiro, S. (2009). Chimpanzees (*Pan troglodytes*) Do not Develop Contingent Reciprocity in an Experimental Task. *Animal Cognition, 12*(4), 587-597. doi:10.1007/s10071-009-0218-z

Bueno-Guerra, N., Völter, C. J., de las Heras, Á., Colell, M., & Call, J. (2019). Bargaining in chimpanzees (Pan troglodytes): The effect of cost, amount of gift, reciprocity, and communication. *Journal of Comparative Psychology, 133*(4), 542–550. doi.org/10.1037/com0000189

Burkart, J. M., & Rueth, K. (2013). Preschool children fail pri- mate prosocial game because of attentional task demands. *PLoS ONE*, *8*(7), e68440. doi:10.1371/journal.pone.0068440

Couzin, I. D. (2006). Behavioural Ecology: Social Organization in Fission–Fusion Societies. *Current Biology, 16*(5), R169-R171. doi:10.1016/j.cub.2006.02.042

Crockford, C., Wittig, R. M., Langergraber, K., Ziegler, T. E., Zuberbuhler, K., & Deschner, T. (2013). Urinary oxytocin and social bonding in related and unrelated wild chimpanzees. *Proc Biol Sci, 280*(1755), 20122765. doi:10.1098/rspb.2012.2765

de Waal, F. B. M. (1989). Food Sharing and Reciprocal Obligations Among Chimpanzees. *Journal of Human Evolution, 18*(5), 433-460.

Dufour, V., Pelé, M., Neumann, M., Thierry, B., & Call, J. (2009). Calculated reciprocity after all: Computation behind token transfers in orang-utans. Biology Letters, 5, 172-175.

Duffy, K. G., Wrangham, R. W., & Silk, J. B. (2007). Male Chimpanzees Exchange Political Support for Mating Opportunities. *Current Biology, 17*(15), R586-R587. doi:10.1016/j.cub.2007.06.001

Engelmann, J. M. & Herrmann E. (2016). Chimpanzees Trust Their Friends. *Current Biology, 26*(2), 252-256. doi: 10.1016/j.cub.2015.11.037

Engelmann, J. M., Herrmann, E., & Tomasello, M. (2015). Chimpanzees trust conspecifics to engage in low-cost reciprocity. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1801), 20142803. doi: 10.1016/j.cub.2015.11.037

Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in Young Children. *Nature, 454*(7208), 1079-1083. doi:10.1038/nature07155

Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology, 65*, 47–55. https://doi.org/10.1007/s00265-010-1038-5

Fujisawa, K. K., Kutsukake, N., & Hasegawa, T. (2008). The Stabilizing Role of Aggressive Children in Affiliative Social Networks Among Preschoolers. *Behaviour, 145*(11), 1577-1600. doi:10.1163/156853908786131289

Gomes, C. M., & Boesch, C. (2009). Wild Chimpanzees Exchange Meat for Sex on a Long-Term Basis. *PLOS ONE, 4*(4), e5116. doi:10.1371%2Fjournal.pone.0005116

Gomes, C. M., & Boesch, C. (2011). Reciprocity and Trades in Wild West African Chimpanzees. *Behavioural Ecology and Sociobiology, 65*(11), 2183-2196. doi:10.1007/s00265-011-1227-x

Gomes, C. M., Mundry, R., & Boesch, C. (2009). Long-term reciprocation of grooming in wild West African chimpanzees. *Proc Biol Sci, 276*(1657), 699-706. doi:10.1098/rspb.2008.1324

Grueneisen, S., Leimgruber, R. L., Vogt, R., & Warneken, F. (2023). Prospection and delay of gratification support the development of calculated reciprocity. *Cognition,* 105369.

Grueneisen, S., & Tomasello, M. (2017). Children coordinate in a social dilemma by taking turns and along dominance asymmetries. *Developmental Psychology, 53*, 265–273. https://doi.org/10.1037/dev0000236

Grueneisen S. & Warneken, F. (2022). The development of prosocial behavior – from sympathy to strategy. *Current Opinion in Psychology,* (43), 323-328. doi: 10.1016/j.copsyc.2021.08.005.

Hamilton, W. D. (1964). The Genetical Evolution of Social Behaviour. I & II. *Journal of Theoretical Biology, 7*(1), 1-52. doi:10.1016/0022-5193(64)90038-4

Hart, B. L., & Hart, L. A. (1992). Reciprocal allogrooming in impala, Aepyceros melampus. *Animal Behaviour, 44*(6), 1073-1083. doi:10.1016/S0003-3472(05)80319-7

Henrich, N., & Henrich, J. (2007). *Why Humans Cooperate: A Cultural and Evolutionary Explanation* (1st ed.). New York: Oxford University Press.

House, B., Henrich, J., Sarnecka, B., & Silk, J. B. (2013). The development of contingent reciprocity in children. *Evolution and Human Behaviour, 34*(2), 86-93. doi:10.1016/j.evolhumbehav.2012.10.001

Jaeggi, A. V., De Groot, E., Stevens, J. M. G., & Van Schaik, C. P. (2013). Mechanisms of reciprocity in primates: testing for short-term contingency of grooming and food sharing in bonobos and chimpanzees. *Evolution and Human Behaviour, 34*(2), 69-77. doi:10.1016/j.evolhumbehav.2012.09.005

Knofe, H., Engelmann, J., Tomasello, M., & Herrmann, E. (2019). Chimpanzees monopolize and children take turns in a limited resource problem. *Scientific reports*, *9*(1), 1-7. doi:10.1038/s41598-019-44096-4

Koomen, R., & Herrmann, E. (2018a). The effects of social context and food abundance on chimpanzee feeding competition. *American journal of primatology*, *80*(2), e22734.doi:10.1002/ajp.22734

Koomen, R., & Herrmann, E. (2018b). Chimpanzees overcome the tragedy of the commons with dominance. *Scientific reports*, *8*(1), 1-12.doi:10.1038/s41598-018-28416-8

Kuhlmeier, V. A., Dunfield, K. A., & O'Neill, A. C. (2014). Selectivity in early prosocial behavior. *Frontiers in Psychology*, *5*, 836.

Melis, A. P. (2017). The evolutionary roots of prosociality: the case of instrumental helping. *Curr Opin Psychol, 20*, 82-86. doi:10.1016/j.copsyc.2017.08.019

Melis, A. P., Engelmann, J. M., & Warneken, F. (2018). Correspondence: Chimpanzee helping is real, not a byproduct. *Nature communications*, *9*(1), 1-3.

Melis, A. P., Hare, B., & Tomasello, M. (2006). Engineering cooperation in chimpanzees: tolerance constraints on cooperation. *Animal Behaviour*, *72*(2), 275-286. doi:10.1016/j.anbehav.2005.09.018

Melis, A. P., Hare, B., & Tomasello, M. (2008). Do Chimpanzees Reciprocate Received Favours? *Animal Behaviour, 76*(3), 951-962. doi:10.1016/j.anbehav.2008.05.014

Melis, A. P., Warneken, F., Jensen, K., Schneider, A. C., Call, J., & Tomasello, M. (2011). Chimpanzees Help Conspecifics Obtain Food and Non-Food Items. *Proc Biol Sci, 278*(1710), 1405-1413. doi:10.1098/rspb.2010.1735

Milinski, M. (1987). TIT FOR TAT in sticklebacks and the evolution of cooperation. *Nature, 325*(6103), 433-435. doi:10.1038/325433a0

Milinski, M., Pfluger, D., Külling, D., & Kettler, R. (1990). Do Sticklebacks Cooperate Repeatedly in Reciprocal Pairs? *Behavioural Ecology and Sociobiology, 27*(1), 17-21. doi:10.1007/BF00183308

Mundry, R., & Fischer, J. (1998). Use of Statistical Programs for Nonparametric Tests of Small Samples Often Leads to Incorrect P Values: Examples From Animal Behaviour. *Animal Behaviour, 56*(1), 256-259. doi:10.1006/anbe.1998.0756

Olson, K. R., & Spelke, E. S. (2008). Foundations of cooperation in young children. *Cognition*, *108*(1), 222-231. doi:10.1016/j.cognition.2007.12.003

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-projectorg/.

Raihani, N. (2021). *The social instinct: How cooperation shaped the world*. Random House.

Schino, G., & Aureli, F. (2009). Reciprocal Altruism in Primates: Partner Choice, Cognition, and Emotions. In H. J. Brockmann, C. T. Snowdon, T. J. Roper, M. Naguib, & K. E. Wynne-Edwards (Eds.), *Advances in the Study of Behaviour, 39* (Vol. 39, pp. 45-69): Academic Press.

Schmelz, M., Grueneisen, S., Kabalak, A. Tomasello, M. (2017). Chimpanzees return favors at a personal cost. *PNAS, 114*(28), 7462-7467. doi: 10.1073/pnas.1700351114

Schmelz, M., Grueneisen, S., & Tomasello, M. (2020). The psychological mechanisms underlying reciprocal prosociality in chimpanzees (Pan troglodytes). Journal of Comparative Psychology, 134(2), 149–157. doi:10.1037/com0000200

Seabright, P. (2004). *The Company of Strangers: A Natural History of Economic Life*: Princeton University Press.

Siegel, S., & Castellan, N. J. (1988). *Nonparametric Statistics for the Behavioural Sciences* (2nd ed.). New York: McGraw-Hill.

Silk, J. B., Brosnan, S. F., Vonk, J., Henrich, J., Povinelli, D. J., Richardson, A. S., . . . Schapiro, S. J. (2005). Chimpanzees Are Indifferent to the Welfare of Unrelated Group Members. *Nature, 437*(7063), 1357-1359. doi:10.1038/nature04243

Stevens, J. R., & Hauser, M. D. (2004). Why be nice? Psy- chological constraints on the evolution of cooperation. *Trends in Cognitive Sciences, 8*, 60–65.

Schweinfurth, M.K. & Call, J., (2019). Revisiting the possibility of reciprocal help in non-human primates. *Neuroscience & Biobehavioral Reviews,* 104, 73-86. doi:10.1016/j.neubiorev.2019.06.026

Tan, J., Kwetuenda, S., & Hare, B. (2015). Preference or paradigm? Bonobos show no evidence of other-regard in the standard prosocial choice task. *Behaviour, 152,* 521-544.

Tennie, C., Jensen, K., & Call, J. (2016). The nature of prosociality in chimpanzees. *Nat Commun, 7*, 13915. doi:10.1038/ncomms13915

Tomasello, M. (2016). *A Natural History of Human Morality*: Harvard University Press.

Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *Quarterly Review of Biology, 46*(1), 35-57.

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not All Emotions Are Created Equal: The Negativity Bias in Social-Emotional Development. *Psychological Bulletin, 134*(3), 383-403. doi:10.1037/0033-2909.134.3.383

Vaish A., Hepach R., Tomasello M. (2018). The specificity of reciprocity: Young children reciprocate more generously to those who intentionally benefit them.

Journal of Experimental Child Psychology,16, 336-353. doi:10.1016/j.jecp.2017.11.005

Warneken, F. (2018) How Children Solve the Two Challenges of Cooperation. *Annual Review of Psychology, 69*, 205–29.

Warneken, F., Hare, B., Melis, A. P., Hanus, D., & Tomasello, M. (2007). Spontaneous Altruism by Chimpanzees and Young Children. *PLoS Biology, 5*(7), e184. doi:10.1371/journal.pbio.0050184

Warneken, F., Sebastián-Enesco, C., Benjamin, N. E., & Pieloch, K.A. (2019). Pay to play: Children's emerging ability to use acts of generosity for selfish ends. *Journal of Experimental Child Psychology, 188*, 104675.

Warneken, F., & Tomasello, M. (2006). Altruistic Helping in Human Infants and Young Chimpanzees. *Science, 311*(5765), 1301-1303. doi:10.1126/science.1121448

Warneken, F., & Tomasello, M. (2009). Varieties of Altruism in Children and Chimpanzees. *Trends in Cognitive Sciences, 13*(9), 397-402. doi:10.1016/j.tics.2009.06.008

Warneken, F., & Tomasello, M. (2013). The emergence of contingent reciprocity in young children. *Journal of Experimental Child Psychology*, 116, 338–50.

Weatherall, Goodfellow, Harris, Hinde, Johnson, Morris, . . . Tickell. (2006). The use of non-human primates in research. *Academy of Medical Sciences*.

Webster, M. M. & Rutz, C. (2020) How STRANGE are your study animals? *Nature* 582, 337–340

Wilkinson, G. S. (1984). Reciprocal Food Sharing in the Vampire Bat. *Nature, 308*(5955), 181-184. doi:DOI 10.1038/308181a0

Wittig, R. M., Crockford, C., Deschner, T., Langergraber, K. E., Ziegler, T. E., & Zuberbuhler, K. (2014). Food sharing is linked to urinary oxytocin levels and bonding in related and unrelated wild chimpanzees. *Proc Biol Sci, 281*(1778), 20133096. doi:10.1098/rspb.2013.3096

Yamamoto, S., Humle, T., Tanaka, M., 2012. Chimpanzees' flexible targeted helping based on an understanding of conspecifics' goals. Proc. Natl. Acad. Sci. 109, 3588–3592.

Yamamoto, S., & Tanaka, M. (2009). Do Chimpanzees (*Pan troglodytes*) Spontaneously Take Turns in a Reciprocal Cooperation Task? *Journal of Comparative Psychology, 123*(3), 242-249. doi:10.1037/a0015838

Zhang, Z., Grocke, P., & Tomasello, M. (2019). The influence of intention and outcome on young children's reciprocal sharing. *Journal of experimental child psychology*, *187*, 104645. doi: 10.1016/j.jecp.2019.05.012

FIGURES

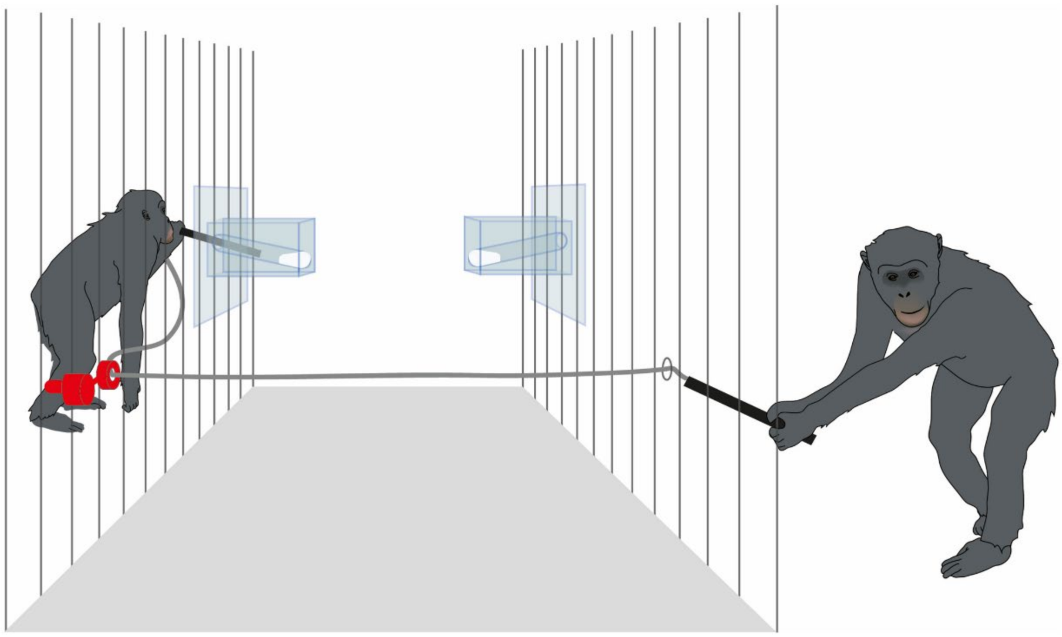

Figure 1. Set up study 1; Tools connected by a short rope were needed to access the identical reward-boxes filled with mashed banana. Firstly, the rope could be accessed in the left room only; simultaneous tool use was impossible. The rope was restricted by a clamp preventing the individual in the right room from pulling and using the tool. The individual in the left room could help by pressing the attached release-button, which removed the clamp and thereby released the rope allowing the other participant to use the tool. The tool was inaccessible for the left individual once pulled to the right room.

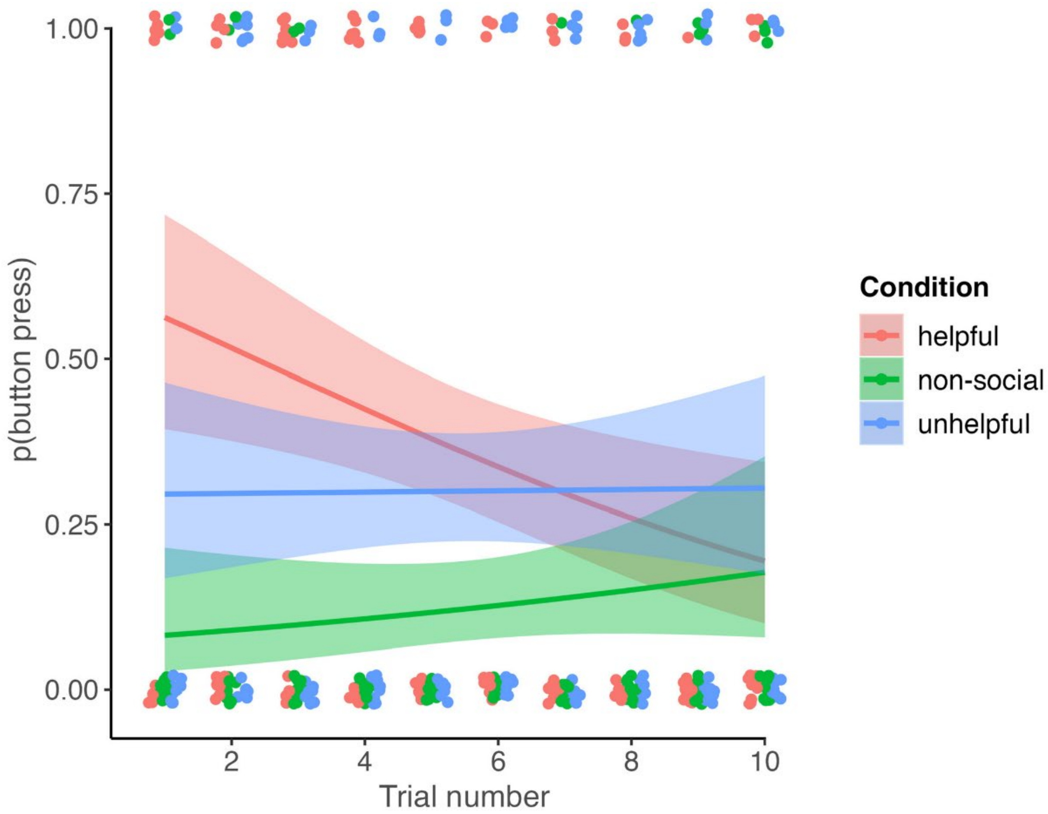

Figure 2. Probability of button presses over the 10 trials of the experiment in the helpful, unhelpful, and non-social condition. Scattered dots represent individual data points. Solid lines are fitted regression lines (binomial logistic regressions), shaded regions indicate 95% confidence intervals.

Figure 3. Set up study 2; Stamps connected by a short rope could be used to mark sheets after being dipped into tubes containing ink pads; simultaneous use of the stamps was impossible. The rope was restricted by a clamp and kept the individual on the right side from using the tool. The individual on the left side could help by pressing the attached release-button allowing the other participant to use the stamp. Experimenters could secretly communicate with the stooge (left child with cap) using a Walkie-talkie. The scene depicts the experience phase of the helpful condition at that time the stooge just finished stamping and is pressing the release-button for the subject (right child).

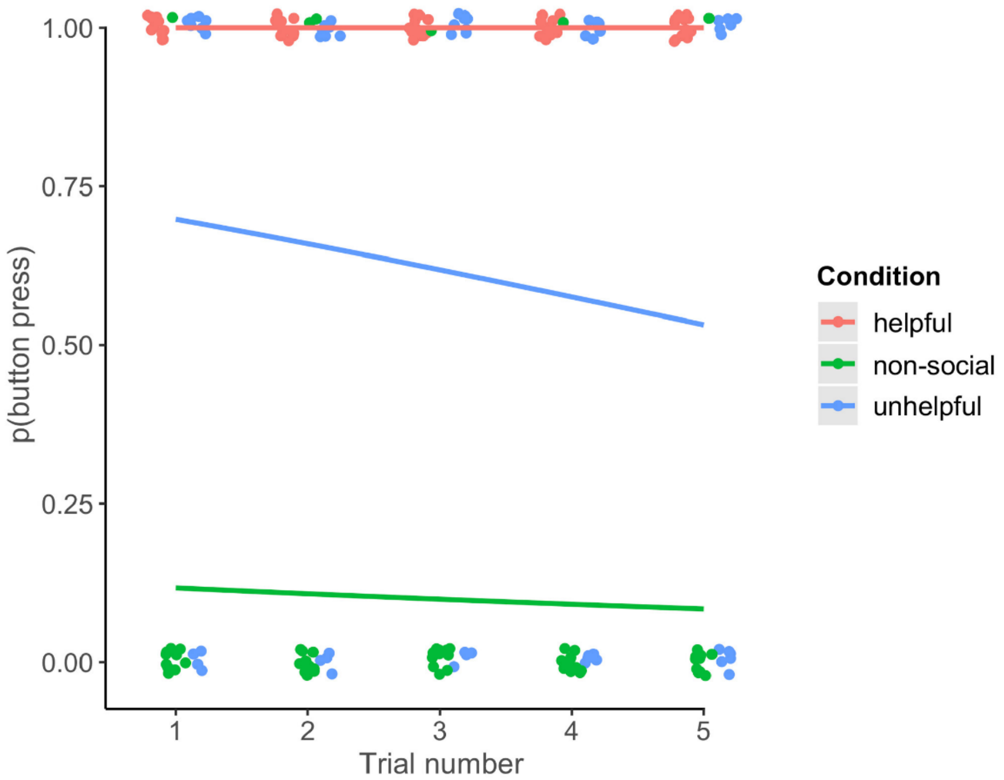

Figure 4. Probability of button presses over the 5 trials of the experiment in the helpful, unhelpful, and non-social condition. Scattered dots represent individual data points. Solid lines are fitted regression lines (binomial logistic regressions). Due to complete separation issues calculating confidence interval was not feasible.