

Children's intrinsic motivation to provide help themselves after accidentally harming others

Robert Hepach¹, Amrisha Vaish², & Michael Tomasello¹

¹Max Planck Institute for Evolutionary Anthropology

²University of Virginia

Author Note

Robert Hepach, Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology; Amrisha Vaish, Department of Psychology, University of Virginia; Michael Tomasello, Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology.

We thank all the families who participated in the studies. The following people helped with data collection and we are grateful for their assistance: Jana Jurkat, Georg Keller, Lydia Fiedler, Selina Brinch, Sveja Taubert, Christian Kliesch, Diotima Rapp, Katharina Walter, Juliane Buchwald, Lou Haux, Ulrike Kachel, Jana Urban, Anna Schöning, Esther Klonowksi, Kristina Schilke, Freijdis Jurkat, and Ronja Büchner. We thank Freijdis Jurkat, Kristina Schilke, and Erin Boggs for their help with coding, Roger Mundry for statistical advice, Robert Schettler for technical support, Henrik Röthel for building as well as Isabelle de Gaillande-Mustoe, Alina Schönfuß, Cleo Meinicke, and Parissa Chokrai for decorating the house apparatus. The pictures in Figure 1 were taken by Christian Kliesch. We also thank Susanne Hardecker and Maria Plötner for their feedback during the pilot phase of the project.

Correspondence concerning this article should be addressed to Robert Hepach, Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig.

Email: hepach@eva.mpg.de

Abstract

Little is known about the flexibility of children's prosocial motivation. Here, 2- and 3-year-old children's ($n = 128$) internal arousal, as measured via changes in pupil dilation, was increased after they accidentally harmed a victim but were unable to repair the harm. If they were able to repair (or if they themselves did not cause the harm and the help was provided by someone else) their arousal subsided. This suggests that children are especially motivated to help those whom they have harmed, perhaps out of a sense of guilt and a desire to reconcile with them. Young children care not only about the well-being of others but also about the relationship they have with those who depend on their help.

Keywords: children, prosocial behavior, motivation, guilt, reconciliation, helping, pupil dilation, pupillometry

Children's intrinsic motivation to provide help themselves after accidentally harming others

Humans exhibit a remarkable degree of prosocial behavior and the foundations of human prosociality are already evident in young children (see, Dunfield, 2014; Martin & Olson, 2015; Paulus, 2014; Warneken, 2015, for recent reviews). Children as young as 12 to 24 months of age show concern when others are hurt (see Roth-Hanania, Davidov, & Zahn-Waxler, 2011; Zahn-Waxler, Radke-Yarrow, Wagner, & Chapman, 1992). Accordingly, toddlers' concern for others relates to how likely they are to help (Eisenberg & Miller, 1997; Vaish, Carpenter, & Tomasello, 2009). Furthermore, children help others achieve instrumental goals (Warneken & Tomasello, 2006, 2007) and share resources with others (Brownell, Svetlova, & Nichols, 2009).

The spontaneous helping behavior of 2-year-old children seems to be intrinsically motivated. Neither praise nor material rewards facilitate helping behavior (Warneken & Tomasello, 2008) and the presence and encouragement of caregivers does not increase toddlers' helping (Warneken & Tomasello, 2013). Two-year-old children's internal physiological arousal, as measured via changes in pupil dilation, is similarly reduced when they themselves help or a third party helps someone in need. Internal arousal is lower in both cases when children see the needy person receiving no help (Hepach Vaish, & Tomasello, 2012). This finding shows that young children are not primarily motivated to themselves provide help and 'get credit' for it, in which case arousal would only be reduced when they themselves could provide help and would stay high when the third party provided help. Thus, the overall conclusion is that 2-year-old children help others out of a genuine concern for the victim's well-

being and are not concerned with *who* provided the help (see also Hepach, Vaish, & Tomasello, 2013b).

Although this might be true in many cases, there are situations in which it does matter *how* the situation is resolved. Three-year-old children share less with an adult who previously harmed another or who was unwilling to help (Vaish, Carpenter, & Tomasello, 2010; see also Dahl, Schuck, & Campos, 2013; Dunfield & Kuhlmeier, 2010) and show less concern for an adult whose distress is unjustified (Hepach, Vaish, & Tomasello, 2013). At the same age, children's instrumental helping is paternalistic such that children correct others if their requests for help do not fulfill their actual need (Martin & Olsen, 2013). Children also selectively maintain their social relationships with cooperative rather than non-cooperative individuals (Vaish et al., 2010; Vaish, Carpenter, & Tomasello, 2011; see also Vaish & Tomasello, 2014).

Interestingly, a pivotal case where it not only matters *how* the situation is resolved but also *who* resolves the situation is the case of guilt - when one's own actions have accidentally harmed another individual. In such a scenario, not actively repairing and failing to actively reconcile can have negative consequences for the transgressor, who may be perceived as not valuing the relationship and lacking motivation to repair it, and thus jeopardizing the relationship with the harmed individual. Here, the social emotion of guilt serves to motivate reparative or other prosocial behavior towards the one harmed so as to reconcile with the victim and avoid possible exclusion from future collaboration (Cunnigham, Steinberg, Grev 1980; Keltner, 1995). Indeed, 3-year-old children who have caused someone harm (and thus presumably experience guilt) show greater reparative behavior than children who did not cause harm and in turn presumably experienced sympathy rather than

guilt (Vaish, Carpenter, & Tomasello, in press; see also, Kochanska, Gross, Lin, & Nichols, 2002).

While this flexibility in children's prosocial behavior is impressive, Vaish et al. (in press) only found this effect in 3-year-old children. Among 2-year-olds, the authors found increased prosocial behavior whenever harm had been caused, regardless of whether or not children themselves caused the harm. This suggests that although 2-year-old children readily sympathize with and help others, they may be insensitive as to whether or not their own actions have caused others harm and may experience only sympathy and not guilt (Vaish et al., 2015; Zahn-Waxler & Kochanska, 1990). This raises the possibility that children before the age of 3 are truly indifferent as to who resolves helping situations even if the child is responsible for harming the victim. It is also possible, however, that whereas 2-year-old children may not yet know how to go about repairing and restoring the damage they cause, their underlying motivation to reconcile is nonetheless enhanced after they have harmed others. Such underlying physiological changes can be measured via changes in pupil dilation.

Such systematic changes in pupil diameter are linked to activation of the sympathetic branch of the autonomous nervous system (Loewenfeld, 1993). Pupils dilate in response to emotionally-charged images (Bradley, Miccoli, Escrig, & Lang, 2008) and sounds (Partala & Surraka, 2003), and are indicative of experienced internal arousal (Bradley et al., 2008). Furthermore, pupil diameter increases in response to stimuli that hold motivational significance (Nieuwenhuis, De Geus, & Aston-Jones, 2011) and psychologically mediated effects on pupil diameter are evident from as early as infancy (Laeng, Sirois, & Gredebäck, 2012; Sirois & Brisson, 2014). With regards to motivation, 2-year-old children's internal arousal (pupil

dilation) remains increased when they observe a person needing help and decreases if that person is helped (see Hepach et al., 2013; Hepach, Vaish, Grossmann, & Tomasello, in press). On the other hand children's pupil dilation remains increased if no help is provided at all (Hepach et al., 2012) or if the need is not appropriately fulfilled (Hepach et al., in press). Importantly, the greater children's increased arousal, the faster they are to subsequently provide the help themselves (Hepach et al., 2013; in press). Changes in pupil dilation, therefore, reveal children's intrinsic motivation to help and pupils remain more dilated so long as the child's initial motive (to see the person in need be helped) is not fulfilled.

Here we investigated whether having harmed others affects 2- and 3-year-old children's intrinsic motivation (via changes in internal arousal) to help others. Children were presented with two different situations – either they themselves accidentally caused a person harm or they merely witnessed another adult accidentally cause the harm. We then manipulated whether or not children could subsequently help the victim or whether a different individual was quicker to provide the victim with the necessary help. Given prior work on children's sympathy toward others in need (e.g., Vaish et al., 2009; Zahn-Waxler et al., 1992), we predicted that both 2- and 3-year-old children would be motivated to see a person in need be helped if they saw the harm being caused by another individual. In this case, children's internal arousal should similarly decrease if they themselves or a third party provide the necessary help. On the other hand, and given prior work on guilt, we predicted that 3-year-old children would be motivated to actively repair the harm if they had caused it (as opposed to merely seeing the person in need be helped). In such a case, children's motivation to reconcile with the victim would be frustrated if someone else provided the help instead of them, and their internal arousal would thus remain increased. With regard

to the 2-year-old children. we tested two alternative predictions. If children at this age are indeed sensitive as to whether or not they have caused someone harm, then they should, similar to the 3-year-olds, show increased internal arousal if they cannot actively repair and restore the situation. On the other hand, if 2-year-olds respond only with sympathy to having harmed others, then it should not matter to them that *they* actively help. Consequently their internal arousal should similarly decrease when they or someone else restores the situation. As additional measures we coded children's verbal responses and the speed with which they initiated their helping behavior during each test trial but we did not have any prior hypotheses as to how these measures would be affected by the experimental conditions.

General Method

Overview

Two studies were conducted, one with 3- and the other with 2-year-old children. All experimental conditions and crucial manipulations were identical in both age groups but the procedure had to be modified and simplified to meet the attention requirements for 2-year-old children. Each age group is thus presented as a separate study. Prior to conducting the actual studies we ran a pilot study with 5-year-old children to develop the paradigm for our targeted age groups, to determine the time points on which to measure pupil dilation, and to write the analysis algorithms for data processing (see supplementary materials for details). In the following paragraphs, we describe the commonalities between the procedures. Procedural specificities for each age group are described in the respective study methods sections.

The studies were carried out in a mid-sized German city (population approximately 500 000; median household monthly income approximately 1400 €). Participants were recruited and randomly selected from a local database. Families

visited the research institute and all children (predominately white Caucasian) received a small toy after participation. An experimenter explained the study to parents before they gave their informed consent. After the study, children were debriefed, showing that the accident was in fact caused by a technical failure (broken train tracks) and was not anyone's fault. In case children were not given the opportunity to help during they study, they were given a chance after the study to repair and help the victim, who then thanked them for their help.

Materials and Stimuli

Eye movement and pupil diameter were recorded while children sat in front of a wooden apparatus in the shape of a house-front. A computer screen (24-inch for 3-year-old participants; 23-inch for 2-year-olds) was embedded into a window (the outside of the house) together with the actual eye-tracking unit. For 3-year-old children we tracked both eye movements and pupil diameter using a Tobii system (model X120). Stimuli were presented with Tobii Studio (Version 3.2.3). For the 2-year-old children we used an SMI eye-tracker (model RED-m) and stimuli were presented with Experiment Center (Version 3.4). Both eye-trackers recorded eye data with a frequency of 60 Hz. A small webcam (Logitech 1080p) was placed on the inside of the house to provide a live-feed on the computer screen. Children could see what was happening on the other side of the house (the game side) when sitting in front of the window.

During the study we used several toys and objects as part of the game that all participants were familiarized with. Children were made to believe that they operated an electric toy train by hitting a cotton start button with a sponge hammer (3-year-olds) or with their hand (2-year-olds). The train was in fact battery operated and remote-controlled by one of the adults (the observer) present in the study room. In

addition, the train carried a plastic container filled with either green- or red-colored water moving on wooden tracks (see Fig. 1 for details).

To measure pupil dilation, we presented neutral video clips (either blue or purple in color, counterbalanced across test trials) on the computer screen (see Hepach, Vaish, Grossmann, & Tomasello, in press) at the beginning of each test trial (baseline measure) and at the end of the trial (post measure). For 3-year-olds we also presented the clips immediately after the first experimental manipulation and between the pre and post measure to refocus children's attention to the computer screen. The luminance in the study room was kept constant for all participants (approx. 45 Lux from the position where children sat, with the computer screen turned off).

General Procedure and Design

The study was explained to parents and their informed consent was obtained prior to commencement of the study. For the 3-year-olds, parents sat in a separate room from which they could follow the study on a computer monitor. If children were not comfortable, parents stayed inside the study room, sat in a corner out of children's immediate view, and read a magazine. For the 2-year-olds, parents were always present in the study room. Four experimenters ran the study: The first adult (the moderator) introduced the child to the game and remained in the room at all times. The second adult (the victim of harm) and the third adult (the guest) entered the room at specific time points during the study. The victim was the target of later harm, the guest controlled the eye tracking software and, in two of the experimental conditions, caused the accident himself. The fourth adult (the observer) sat to the side of the game. She repaired the harm in two of the experimental conditions. Before children entered the study room, the adults introduced themselves in a separate playroom such that participants were familiar with them. Children entered the study room together

with the moderator and sat in front of the computer screen. Before the actual study began, the eye tracking system was calibrated to participants' eyes. This was done following a standard five-point calibration procedure (Gredebäck, Johnson, & von Hofsten, 2009). The moderator always sat next to the child (position counterbalanced). After the calibration, children saw the live feed on the computer screen and the moderator introduced children to the game. Each study session lasted approximately between 20 and 30 minutes (average test phase duration was 3.5 mins).

Each child was randomly assigned to one of four experimental conditions: child-harms / child-repairs, child-harms / 3rd party (3P) repairs, adult-harms / child-repairs, adult-harms / 3P-repairs. Children participated in two test trials in a between-subject design. Each test trial consisted of a preparation and a test phase. After the first trial, adults engaged children in a brief free play situation before starting with the subsequent test trial. Based on previous work (e.g., Hepach et al., 2012) we included a second test trial to be able to compute an average change in pupil dilation and, in case data could not be collected on the first test trial, to have a second opportunity to measure children's response. Preliminary analyses revealed no effect of trial on the interaction of the factors who harmed and who repaired (see details below).

Dependent Measures

The primary dependent measure was the baseline-corrected change in children's pupil dilation (see below). We collected two additional variables to provide information on the nature of children's responses. Through coding children's verbal responses we explored whether children expressed guilt-related verbal responses during the test trial. This coding served as a quasi manipulation check for whether our interpretation of a guilt-like situation is justified. The descriptive data are presented in the supplementary materials. The second additional variable we coded was children's

latency to attempt and repair the situation. We did not have any hypotheses as to whether children in the child-harms condition would be faster to help compared to the adult-harms condition. However, we did include children's latency to help as a control variable to rule out that children's change in pupil dilation was influenced by the time they were exposed to the victim (who increased his / her cues for help over time).

Pupil Dilation. The processing and data analysis of pupil dilation followed a standardized procedure. The recorded data were exported to text files that were in turn processed in Matlab (Version 8.2.0.701, R2013b) and R (R Core Team, 2015; Version 3.1.3). First, data for the entire session of each participant were pre-processed. We applied both a quantile cut-off filter (removing the upper tercentile of sample-to-sample differences) as well as a linear interpolation of sample-to-sample gaps not exceeding 70 ms (Hepach et al., 2012, 2013, in press). Next, data for both eyes were averaged, filtered, and gaps were interpolated.

Systematic changes in pupil dilation were analyzed for the neutral clips. Each measurement consisted of a bright image on the computer screen presented twice, and thus eliciting two pupillary light reflexes (Hepach et al., in press). Further processing steps were taken for the data range of each of the two pupillary light reflexes (duration 4.7 s, 284 samples). Samples were discarded when participants did not actually look at the screen and when an additional density filter removed data ranges for which no pupil diameter was recorded between ~ 1 and 2.5 s following stimulus onset. Specifically, we divided the entire range for each pupillary light reflex into 4 equal sized segments. Segments two, three, and four contained the pupillary light reflex (PLR). We only retained data ranges for which at least two adjacent *data samples* were found either in all three or in two adjacent *data segments*. To further

reduce noise in the data, we applied the quantile cut-off filter and again interpolated to each data range of each PLR. Finally, we determined the lowest value of pupil diameter for the first half of each data range thereby identifying the minimum of the PLR. For each presentation of the neutral clips we calculated the average pupillary minimum. As internal arousal increases, the PLR is inhibited such that overall pupil dilation including the pupillary minimum is greater (see also Steinhauer, Condray, & Kasperek, 2000). For statistical analyses, we calculated the change in pupil dilation from the beginning to the end of the test trial, i.e., $(\text{post-measure} - \text{baseline-measure}) / \text{baseline-measure}$. Each participant provided two baseline-corrected values of pupil dilation, i.e., one for each test trial. We further averaged data from both test trials to arrive at one baseline-corrected dependent measure for each participant. In cases where no data could be recorded for one test trial, only the data from the other trial were used.

Latency to help. The latency to help was coded (adult coder blind to the study's hypotheses and conditions) from video tape and measured from the moment children got up from in front of the computer screen after the accident occurred (the moment the train collapsed for the 2-year-old participants) until children reached for the paper towel (see details below). A second adult coder blind to the study's hypotheses and conditions coded the latencies of 25 % of the all trials and agreement between the coders was high for the 3-year-old ($ICC = .925$) and 2-year-old participants ($ICC = .998$).

Study 1

Participants

Participants were 3-year-old children ($n = 64$, 31 girls) between three years and three years six months ($M = 39\text{m } 13\text{d}$; $SD = 47\text{d}$; range: 36m 18d to 41m 25d). A

total of 89 children participated in the study. Data from 64 children were included in the final analyses. Additional children were tested but excluded due to an experimenter error ($n = 2$), equipment failure ($n = 7$), the child not being a native speaker ($n = 1$), parental interference ($n = 1$), because the child did not want to participate ($n = 8$) or because no data could be recorded during either test trial ($n = 6$). In addition, for 15 children one test trial had to be excluded due to equipment failure ($n = 5$), an experimental manipulation error ($n = 1$), parental interference ($n = 1$) or because the child did not want to participate on the second test trial ($n = 8$). Details regarding the exclusion criteria are provided in the supplementary materials.

Procedure

After the calibration, children participated in the familiarization phase in which the moderator explained the game and its components to the child (see Fig. 1). Wooden train tracks could be assembled to direct the train to either the victim's or the child's (and partner's) playing field. Children were made to believe that they could start the train if they sat in front of the window and hit a cloth button with a sponge-hammer (see supplementary materials for details). Following this familiarization phase, the test phase began. Children first participated in a preparation trial and learned that they could use a paper towel to wipe up liquids that accidentally spilled (see supplementary materials for details). At the end of the preparation trial the victim and guest entered the room.

The test trial then began with the victim noticing the water on the moderator's and child's playing board. He asked whether he could get some water for his animals as well. At this point the first experimental manipulation took place. In the two 'child-harms' conditions, the child turned the train around and assembled the train tracks such that the train went toward the victim, and placed the canister with water on the

train. The victim then said: "Ok, [child's name], you start the train so that it goes to my station here (pointing at his playing field) and I will wait with my canister (holding the canister in his hand). Ok?" The moderator asked the child: "So who is going to start the train?" If children did not immediately answer, the moderator said, "You will start the train." In the two 'adult-harms' conditions, it was the guest who prepared the train and who was addressed by the victim. The moderator again made sure that children understood that it was the guest who was going to start the train.

Next, the child, the moderator, and the guest moved in front of window and children sat in front of the computer screen with the two adults on either side (positions counterbalanced). The neutral stimulus was presented and the baseline-measure of pupil dilation was taken. During this time the victim surreptitiously manipulated the train tracks such that the train was bound to fall over once it reached his playing board. Next, the three adults saw the victim through the live feed saying: "Ok, [child's name or guest's name, depending on condition], I'm ready. You can hit the button." The child or the guest then started the train. The victim said: "Thank you [child's or guest's name]. The train is coming my way." Before reaching the victim's playing board, the train stopped moving, prompting him to say: "Oh, you will have to hit the button again [child's or guest's name]." The child or the guest restarted the train and the victim said "Thank you [child's or guest's name]." In all conditions, the guest now excused himself and left the room and only the child, the moderator, and the observer remained in the room. As the train approached the victim's field, it passed the manipulated train tracks and collapsed on the victim's playing board, consequently spilling the water onto the playing field. The victim exclaimed: "Oh no, what is happening? The train did not stop. Oh no! You [child's or guest's name] hit the button too hard." The neutral stimulus was shown on the computer screen.

At this point children moved to the other side of the house and the second experimental manipulation took place. In the child-repairs conditions the child could now reach for the paper towel and hand it to the victim, who thanked the child. In the 3P-repairs conditions, just as the child reached toward the paper towel, the observer grabbed it and handed it to the victim who thanked her in return. If children did not immediately reach for the towel, the victim provided several cues until children attempted to reach for the towel: (1) "Oh no. It spilled." (~ 3 s) (2) The victim looked at the child (~ 2 s). (3) "I need a towel." (~ 3 s) (4) "Can you give me the towel?" With the paper towel in his hand, the victim wiped up the water and turned the train around so that it could go back. The moderator and child moved in front of the computer screen and the neutral stimulus was presented for a third time. Once the victim could be seen again through the live feed, he had wiped up all the water. The moderator encouraged children to hit the button again and the victim stated that he had cleaned everything up and some water was even left for him to play with. Once the train arrived in its initial position, the neutral stimulus was shown for a fourth time and the post measure of pupil dilation was taken.

At the end of the first test trial, the moderator and child moved to the game side to play. The moderator put a new paper towel on the bench and joined the victim and the child to play with the toy animals on the playing boards for a brief period of time. At the end of this phase, the victim left the room saying he wanted to check on the guest. This marked the beginning of the second test phase, which was identical to the first, with one preparation trial and one test trial.

Data Analyses

For the main analyses regarding the baseline-corrected change in children's pupil dilation at the end of the test trial (post measure; averaged across both trials), we

ran an analysis of variance (ANOVA) with children's change in pupil dilation as the dependent measure and the interaction of who caused the harm and who repaired the harm as the independent factors. In addition, we included participants' gender and the identity of the observer (five different adults for the entire study) as control variables. Preliminary analyses revealed no interaction of trial number with the hypothesized interaction of who caused the harm and who provided the help and preliminary visual inspection indicated that both the dependent measures as well as the residuals were normally distributed and that there was no relation between the residuals and the fitted values (see supplementary materials for details). Posthoc pairwise comparisons were carried out using Mann-Whitney-U tests.

Results and Discussion

Pupil dilation. Three-year-old children's pupil dilation after the resolution of the situation varied as a function of both who harmed the victim and who got to subsequently help him, $F(1, 55) = 4.36, p = .04, 95\% \text{ CI } [-0.11 -0.002], r^2_{adj} = .12$ (see Fig. 3). Children who caused the harm had a greater increase in pupil dilation when they did not themselves help ($M = 0.07; SD = 0.06$) compared to when they did help ($M = 0.009; SD = 0.05$), $U = 199, \text{ CI } [0.02 0.09], p = .006$. Children who did not cause the harm showed similar levels in pupil dilation when they did not themselves help ($M = 0.03; SD = 0.05$) compared to when they did help ($M = 0.02; SD = 0.05$), $U = 149, \text{ CI } [-0.03 0.05], p = .45$. In addition, children who could not themselves help after causing the harm showed greater pupil dilation compared to when they could not help after not having caused the harm, though this was a statistical trend, $U = 178, \text{ CI } [-0.005 0.07], p = .06$. There was no effect of gender, $F(1, 55) = 2.22, p = .14$, or identity of the observer, $F(1, 55) = 0.57, p = .69$ (supplementary materials for more details on the analyses).

Latency. On the first test trial, all children in both of the child-repairs conditions helped by handing the victim the paper towel but the latency to help did not differ between the two conditions: Children who caused the harm helped on average after 13.5 s ($SD = 6.6$ s) and children who did not cause the harm helped after 11.1 s ($SD = 3.7$ s), $U(\text{Mann-Whitney}) = 603, p = .22$. Similarly on the second test trial, except for one child, all children in the child-repairs conditions helped by handing the victim the paper towel. Again, the latency to help did not differ between the two conditions: Children who caused the harm (both 3P-repairs and child-repairs) helped on average after 7.2 s ($SD = 5.4$ s) and children who did not cause the harm (both 3P-repairs and child-repairs) helped after 10.3 s ($SD = 11.5$ s), $U = 340, p = .21$.

Pupil Dilation & Latency to help. An additional analysis was carried out to investigate whether the statistical interaction effect of who caused the harm and who repaired it was influenced by the time (in s) it took children to reach for the paper towel, i.e., their latency to help. This is because the recipient increased his / her cues over time, and children's motivation to provide the help themselves may have been influenced by the time they were exposed to the situation. We transformed the variable of children's latency to achieve a symmetrical distribution of values. The effect of who harmed the recipient and who resolved the situation on children's change on pupil dilation was not influenced by the time children took to reach for the paper towel, GLMM: $\chi^2(1) = .32, p = .57$ (see supplementary analyses for more details).

These main results regarding the change in children's pupil dilation suggests that children's initial motive to help the victim was different between the child-harms and adult-harms conditions. When children witnessed the victim being harmed by another person, they were indifferent as to who resolved the situation and the levels of

their internal arousal were similar if they or a third party helped the victim. On the other hand, when children caused the harm, their motivation changed such that now children wanted to actively provide the help themselves. Children who could not do so showed greater levels of internal arousal, presumably because their initial motive to reconcile with the victim was frustrated. In line with previous work, 3-year-old children who cause others harm are motivated to act prosocially not just due to sympathetic concern but also because they experience feelings of guilt that motivate reparative and reconciliation behavior (Kochanska et al., 2002; Vaish et al., 2015). One resulting question is how this motivation to reconcile develops and whether younger children show similar flexibility in their motivation to help others.

Study 2

Participants

Participants were 2-year-old children ($n = 64$, 32 girls) between 2 years and 2 years 6 months ($M = 27\text{m } 29\text{d}$; $SD = 53\text{d}$; range: 24m 2d to 30m). A total of 109 children participated in the study. Data from 64 children were included in the final analyses. Additional children were tested but excluded due to an experimenter error ($n = 1$), equipment failure ($n = 5$), parental interference ($n = 2$), because they did not want to participate in the procedure ($n = 24$), because of an experimental manipulation error ($n = 4$), or because no data could be recorded during either test trial ($n = 9$). In addition, for 19 children, one test trial had to be excluded due to equipment error ($n = 6$), an experimental manipulation error ($n = 5$), or because the child did not want to participate on the second test trial ($n = 8$). Details regarding the exclusion criteria are provided in the supplementary materials.

Procedure

The procedure was adapted from Study 1 but several crucial changes had to be made to make the study more engaging and comprehensible for the younger age group. The moderator sat behind children and always closed her eyes during the measurements of pupil dilation. After the calibration, the victim (also present from the beginning of the study) picked up a toy animal and brought it to the child (who was still sitting in front of the computer screen). The moderator and child both got up and placed the animal back on one of the playing fields. They took a seat again in front of the computer screen such that the moderator could point out to the child that the toy was still in the same position where they had left it. Next, both the child and moderator got up again and the moderator showed the train to the child. The victim excused herself and left the study room. There was no separate familiarization trial and the moderator proceeded with the preparation trial for the first test phase. The main differences in the procedure compared to the 3-year-olds were as follows: Children sat in front of the computer screen only at the very beginning (baseline measure) and at the very end (post measure) of each preparation and test trial. Children moved to the game side immediately after the baseline measure and only moved back in front of the computer screen before the post measure. The start button was positioned next to the train tracks and children hit it with their hand instead of the hammer. This made the mechanism more intuitive for children. During the first preparation trial the child had multiple opportunities to re-start the train. The train stopped five times on its way to the moderators' and child's playing field. Each time, the child had an opportunity to restart it. The moderator unloaded the water containers and turned around the train. On its way back the child had the opportunity to restart the train on three additional instances. After hitting the train for the last time and

while it was still moving, the moderator and child moved in front of the computer screen again. Once seated, the child saw the train move to its final location; at this point the neutral clip was presented and the post-measure of pupil dilation was taken.

During the test trial, the victim gave additional instructions to make the procedure more comprehensible for participants. After the train was prepared and ready to move toward her playing field, he said: "So, you (looking at either the child or guest) hit this (pointing at the start button) so that the train goes right here to my side. And I will wait with my canister. And [child's name or guest's name] watches." Next either the child or the guest (depending on condition) started the train by hitting the button on the game side. The rest of the procedure was identical to that of the 3-year-olds.

Data Analysis

Similar to the 3-year-olds we measured the change in children's pupil dilation from the beginning of the test trial to the end. The dependent measure for the analysis incorporating both test trials was square-root transformed to achieve normal distribution of residuals. Preliminary analyses revealed no interaction of trial number with the hypothesized interaction of who caused the harm and who provided the help and preliminary visual inspection indicated that both the dependent measures as well as the residuals were normally distributed and that there was no relation between the residuals and the fitted values (see supplementary materials for details).

Results

Pupil dilation. 2-year-old children's pupil dilation after the resolution of the situation varied at the level of a statistical trend as a function of both who harmed the victim and who got to subsequently help, $F(1, 58) = 3.74, p = .058, 95\% \text{ CI } [-0.19, 0.003], r^2_{adj} = .09$ (see Fig. 3). Children who caused the harm had a greater increase

in pupil dilation when they did not themselves help ($M = 0.07$; $SD = 0.04$) compared to when they did help ($M = 0.02$; $SD = 0.07$), $U = 202$, $CI [0.02\ 0.09]$, $p = .004$.

Children who did not cause the harm showed similar levels in pupil dilation when they did not themselves help ($M = 0.03$; $SD = 0.06$) compared to when they did help ($M = 0.03$; $SD = 0.05$), $U = 120$, $CI [-0.05\ 0.03]$, $p = .78$. In addition, children who could not themselves help after causing the harm showed greater pupil dilation compared to when they could not help after not having caused the harm, $U = 202$, $CI [0.01\ 0.08]$, $p = .004$. There was no effect of gender, $F(1, 58) = 1.54$, $p = .22$, or identity of the observer, $F(1, 58) = 1.59$, $p = .21$.

Latency. On the first test trial, with the exception of one child who used the paper towel to actively wipe off the water himself/herself, all children in the child-repairs conditions (both child-harms and adult-harms) helped by handing the victim the paper towel. Children who caused the harm helped on average after 18.35 s ($SD = 9.65$ s) and children who did not cause the harm helped after 16.32 s ($SD = 10.13$ s), $U = 565$, $p = .24$. On the second test trial, 4 children helped by putting the paper towel on the spilled water and 25 helped by handing the victim the paper towel. Children who caused the harm helped on average after 9.97 s ($SD = 8.25$ s) and children who did not cause the harm helped after 11.33 s ($SD = 10.3$ s), $U = 414$, $p = .75$.

Pupil Dilation & Latency to help. We transformed the variable of children's latency to achieve a symmetrical distribution of values. The effect of who harmed the recipient and who resolved the situation on children's change on pupil dilation was not influenced by the time children took to reach for the paper towel, GLMM: $\chi^2(1) = .89$, $p = .35$ (see supplementary analyses for more details).

Pupil Dilation including both age groups. We carried out an additional analysis to investigate whether the interaction of who harmed and who repaired the

situation differed statistically between the two studies and hence between the two age groups. The data from both studies were combined and the dependent measure was the baseline-corrected change in pupil dilation. We included a 3-way-interaction of who harmed, who repaired, and age as well as the control factors gender and experimenter identity into the model. This analysis did not reveal a statistically significant 3-way-interaction, $F(1, 114) = 0.09, p = .77$, and no main effect of age, $F(1, 114) = 0.27, p = .61$ (see supplementary materials for details). At the same time the interaction of who harmed and who repaired remained statistically significant, $F(1, 114) = 6.89, p = .01, 95\% \text{ CI } [-0.1 \text{ } 0.009], r^2_{adj} = .1$. In addition, there was a statistically significant of gender, $F(1, 114) = 4.27, p = .04, 95\% \text{ CI } [0.0004 \text{ } 0.04]$. Boys showed greater overall increase in pupil dilation ($M = 0.04, SD = 0.06$) than girls ($M = 0.02, SD = 0.05$).

General Discussion

The two studies presented here demonstrate that children are intrinsically motivated to provide the necessary help themselves if their actions have caused others harm. Both 2- and 3-year-old children's internal arousal, as measured via changes in pupil dilation, increased if they could not remedy the harm they had caused and merely watched another adult provide the help. Children's internal arousal did however decrease if they could help the victim. The pattern of children's internal arousal was different in situations in which they were merely bystanders and saw another party causing the harm. In this case children at both ages were not primarily motivated to repair the harm themselves. In line with prior work, we found that children were motivated to see others helped (Hepach et al., 2012), but this motivation was specific to cases in which they are not responsible for the victim needing help.

Prior work had established that children demonstrate more reparative behavior after they have accidentally harmed others (Vaish et al., in press). The present results add importantly to those findings by showing that guilt-inducing situations affect children's underlying motivation such that they want to actively reconcile with the victim. This allows children to be recognized for their helpful intention (see also Keltner, 1995; Kochanska et al., 2002; Vaish et al., in press). Children, therefore, are not only selective with regards to whom they help (Dahl et al., 2013; Dunfield & Kuhlmeier, 2010; Vaish et al., 2010) but also whether they want to provide the necessary help themselves.

It is important to note that in the present studies the situation was always remedied for the victim and children consequently saw a completed action in all experimental conditions before the post measure of pupil dilation was taken. Therefore, it is unlikely that the condition effects were a result of children witnessing incomplete actions or unresolved physical situations. The only difference between conditions was who had caused the harm and who resolved the situation. Furthermore, we did not find a condition effect of latency, i.e., the time children took to help the victim, suggesting that children were equally engaged and motivated to help the adult across the conditions. What differed was the underlying motivation for *why* children attempted to help the adult.

The fact that the interaction of who harmed and who repaired resulted in a statistically significant difference for the 3-year-old children but remained at a statistical trend for the 2-year-old children could suggest that younger children's sensitivity to having harmed others is not yet robust and fully developed by the age of 2 (see also Vaish et al., in press). While we did not find an effect of age in the present studies, this is merely suggestive, given that the procedures of both present studies

were modified to meet the attention requirements of each age group and were not identical. Nevertheless, it is noteworthy that while 2-year-old children's reparative behavior does not specifically increase after harming others (Vaish et al., in press), our results are at least suggestive that their internal arousal may be affected. This hints that children's physiological responses in these complex social situations may develop before their actual behavior. In turn, this speaks to the importance of assessing children's physiological responsiveness in social settings, and provides an interesting avenue for future research on prosocial motivations across social contexts, including studying how different cognitive abilities (e.g. Miller & Marcovitch, 2015) influence prosocial motivations.

The present studies were designed to address a specific research question regarding the flexibility of children's prosocial motivation using a novel research method. Yet this is only a first step and more research is needed to further improve the method and collect additional data, including children's live gaze behavior, so as to further explore the underlying motivation of children's prosocial behavior. A further limitation of the present studies is that our procedure did not elicit verbal responses from enough children to perform statistical analyses on the types of responses children provide to seeing the accident. In the present case there was no obvious pattern in children's comments between the child-harms and the adult-harms conditions. Future studies could adopt elicited response paradigms such as asking children specific questions, which has also been applied in studying children's responses to guilt (Vaish et al., 2011, in press). Finally, note that the present studies were carried out with 64 children at each age, and in one cultural context. It is worthwhile to speculate whether children in other cultures would show different arousal responses, especially given that the emotion of guilt plays a more prominent

role in Western culture (Fessler, 2007). No study, thus far, has assessed children's internal arousal to others needing help in cross-cultural comparisons.

Another important point to address is the question of whether children in the two child-harms conditions were merely motivated to restore the physical situation and wipe up the spilled water as opposed to reconciling with the victim. That is, it is possible that the increased arousal in the child-harms / 3P-repairs condition resulted from children wanting to themselves remedy the accident they had caused regardless of whether their actions had harmed someone else. For the 3-year-old children this is an unlikely explanation given that Vaish et al. (2015) explicitly tested this alternative and found that children were less motivated to repair if those actions did not cause harm to anyone. Moreover, prior work shows that already at 24 months of age, children show less internal arousal when they witness objects drop from a table in the absence of a person who needs the objects compared to when a person is present and actually needs the objects to complete a task (Hepach et al., in press). Furthermore, children in Hepach et al.'s (in press) study were indifferent as to how the non-social situation was resolved but in the social situation, children responded with greater pupil dilation if the person did not get the relevant object for his task. This suggests that 2-year-old children are less motivated to see non-social accidents repaired. With regard to the present findings, it is further important to note that in the 3P-repairs conditions, the observer adult always waited until children moved toward the paper towel before she herself reached for it. In that sense children's action was always thwarted. Nevertheless, only in cases where children themselves had caused the harm did the results show greater internal arousal, which suggests that children's increased pupil dilation in the child-harms / 3P-repairs condition is not simply a general frustration about failure to follow through with an action. In addition, in the

conditions where children could indeed help the victim, the majority of children did so by handing over the paper towel instead of taking the paper towel and wiping up the spilled water themselves. Likewise, there was no difference in the latencies with which children attempted to help the adult between the child-harms and adult-harms conditions.

Reconciling with others is an important aspect of maintaining social relationships. When children accidentally harm others, they risk being perceived as not valuing the relationship if they fail to repair the harm. One part of reconciling is showing appeasing, guilt-like displays to elicit sympathy from the victim (Keltner, 1995). The way in which children show that they regret having harmed another individual can vary between age groups. What the present results suggest is that while children, by the age of two, may not show guilt in the form of increasing their helping (Vaish et al., 2015), they nevertheless care about who provides the help and want to provide assistance themselves if they have caused the harm. The ontogenetic roots of guilt may run as deep as the second year of life. One possible future avenue is to vary the relationship children have with the experimenter. Children may not be motivated to repair the harm they have caused if the adult had previously been unwilling to cooperate and engage with the child (see Behne, Carpenter, Call, & Tomasello, 2005; Dunfield & Kuhlmeier, 2010).

In summary, 2- and 3-year-old children are motivated to help a victim themselves in a guilt-provoking situation, so as to reconcile with the victim. On the proximate psychological level, while children's prosocial behavior is not driven by general concerns to get credit for helping, children do care about providing the help themselves specifically in situations in which not acting could jeopardize their relationships with others. From early development and as interdependent group

members, humans care not only about the wellbeing of others but also about the integrity of their relationship with others.

References

- Bates, D., Maechler, M., & Bolker, B. (2011). *lme4: Linear mixed-effects models using Eigen and Eigenfaces*. R package version 0.999375-42.
- Behne, T., Carpenter, M., Call, J., & Tomasello, M. (2005). Unwilling versus unable: infants' understanding of intentional action. *Developmental Psychology*, 41(2), 328.
- Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, 45(4), 602-607.
- Brownell, C., Svetlova, M. & Nichols, S. (2009). To share or not to share: When do toddlers respond to another's needs? *Infancy*, 14(1), 117-130.
- Dahl, A., Schuck, R. K., & Campos, J. J. (2013). Do young toddlers act on their social preferences?. *Developmental Psychology*, 49(10), 1964.
- Cunningham, M. R., Steinberg, J., & Grev, R. (1980). Wanting to and having to help: Separate motivations for positive mood and guilt-induced helping. *Journal of Personality and Social Psychology*, 38(2), 181.
- Dunfield, K. A., & Kuhlmeier, V. A. (2010). Intention-mediated selective helping in infancy. *Psychological Science*, 21, 523-527.
- Dunfield, K., Kuhlmeier, V. A., O'Connell, L., & Kelley, E. (2011). Examining the diversity of prosocial behavior: Helping, sharing, and comforting in infancy. *Infancy*, 16(3), 227-247.
- Dunfield, K. A. (2014). A construct divided: prosocial behavior as helping, sharing, and comforting subtypes. *Frontiers in Psychology*, 5.
- Fessler, D. (2007). From appeasement to conformity. *Self-conscious emotions: Theory and Research*, 174-193.
- Gredebäck, G., Johnson, S., & von Hofsten, C. (2009). Eye tracking in infancy research. *Developmental Neuropsychology*, 35(1), 1-19.

- Hamlin, J. K., Wynn, K., Bloom, P., & Mahajan, N. (2011). How infants and toddlers react to antisocial others. *Proceedings of the national academy of sciences*, 108(50), 19931-19936.
- Hepach, R., Vaish, A., & Tomasello, M. (2012). Young children are intrinsically motivated to see others helped. *Psychological Science*, 23(9), 967-972.
- Hepach, R., Vaish, A., & Tomasello, M. (2013). Young children sympathize less in response to unjustified emotional distress. *Developmental Psychology*, 49(6), 1132-1138.
- Hepach, R., Vaish, A., & Tomasello, M. (2013b). A new look at children's prosocial motivation. *Infancy*, 18(1), 67-90.
- Hepach, R., Vaish, R., Grossmann, T., & Tomasello, T. (in press). Young children want to see others get the help they need. *Child Development*.
- Hoffman, M. L. (1982). Development of prosocial motivation: Empathy and guilt. *The development of prosocial behavior*, 281, 313.
- Keltner, D. (1995). Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement, and shame. *Journal of Personality and Social Psychology*, 68, 441-454. doi:10.1037/0022-3514.68.3.441
- Kochanska, G., Gross, J. N., Lin, M-H., & Nichols, K. E. (2002). Guilt in young children: Development, determinants, and relations with a broader system of standards. *Child Development*, 73, 461-482.
- Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry A Window to the Preconscious? *Perspectives on Psychological Science*, 7(1), 18-27.
- Loewenfeld, I. E. (1993). *The Pupil: Anatomy, Physiology, and Clinical Applications*. Detroit, MI: Wayne State University Press.
- Martin, A., & Olson, K. R. (2013). When kids know better: paternalistic helping in 3-

- year-old children. *Developmental Psychology*, 49(11), 2071.
- Martin, A., & Olson, K. R. (2015). Beyond Good and Evil What Motivations Underlie Children's Prosocial Behavior?. *Perspectives on Psychological Science*, 10(2), 159-175.
- Miller, S. E., & Marcovitch, S. (2015). Examining executive function in the second year of life: Coherence, stability, and relations to joint attention and language. *Developmental Psychology*, 51(1), 101.
- Nieuwenhuis, S., De Geus, E. J., & Aston-Jones, G. (2011). The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology*, 48(2), 162-175.
- Nowak, M. A., & Sigmund, K. (1998). Evolution of indirect reciprocity. *Nature*, 393, 573-577.
- Paulus, M. (2014). The emergence of prosocial behavior: why do infants and toddlers help, comfort, and share? *Child Development Perspectives*, 8(2), 77-81.
- R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Roberts, G. (2005). Cooperation through interdependence. *Animal Behaviour*, 70(4), 901-908.
- Roth-Hanania, R., Davidov, M., & Zahn-Waxler, C. (2011). Empathy development from 8 to 16 months: Early signs of concern for others. *Infant Behavior and Development*, 34(3), 447-458.
- Schielzeth, H., & Forstmeier, W. (2009). Conclusions beyond support: Overconfident estimates in mixed models. *Behavioral Ecology*, 20(2), 416-420.

- Sirois, S., & Brisson, J. (2014). Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(6), 679-692.
- Trivers, R. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, 46, 35-57.
- Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., & Herrmann, E. (2012). Two key steps in the evolution of human cooperation. *Current Anthropology*, 53(6), 673-692.
- Vaish, A., Carpenter, M., & Tomasello, M. (2010). Young children selectively avoid helping people with harmful intentions. *Child Development*, 81(6), 1661-1669.
- Vaish, A., Carpenter, M., & Tomasello, M. (2011). Young children's responses to guilt displays. *Developmental Psychology*, 47(5), 1248-1262.
- Vaish, A., & Tomasello, M. (2014). The early ontogeny of human cooperation and morality. *Handbook of Moral Development*, 279-298.
- Vaish, A., Carpenter, M., & Tomasello, M. (in press). The early emergence of guilt-motivated prosocial behavior. *Child Development*.
- Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, 311, 1301-1303.
- Warneken, F., & Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy*, 11(3), 271-294.
- Warneken, F., & Tomasello, M. (2008). Extrinsic rewards undermine altruistic tendencies in 20-month-olds. *Developmental Psychology*, 44(6), 1785 - 1788.
- Warneken, F., & Tomasello, M. (2009). Varieties of altruism in children and chimpanzees. *Trends in Cognitive Sciences*, 13(9), 397-402.
- Warneken, F., & Tomasello, M. (2013). Parental presence and encouragement do not influence helping in young children. *Infancy*, 18(3), 345-368.

- Warneken, F. (2015). Precocious Prosociality: Why Do Young Children Help?. *Child Development Perspectives*, 9(1), 1-6.
- Wickham, H. (2009). *ggplot2: elegant graphics for data analysis*. Springer: New York.
- Zahn-Waxler C, Kochanska G. (1990). The origins of guilt. In *The Nebraska Symposium on Motivation 1988: Socioemotional Development*, ed. RA Thompson, 36:183–258. Lincoln: Univ. Nebraska Press
- Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology*, 28(1), 126.

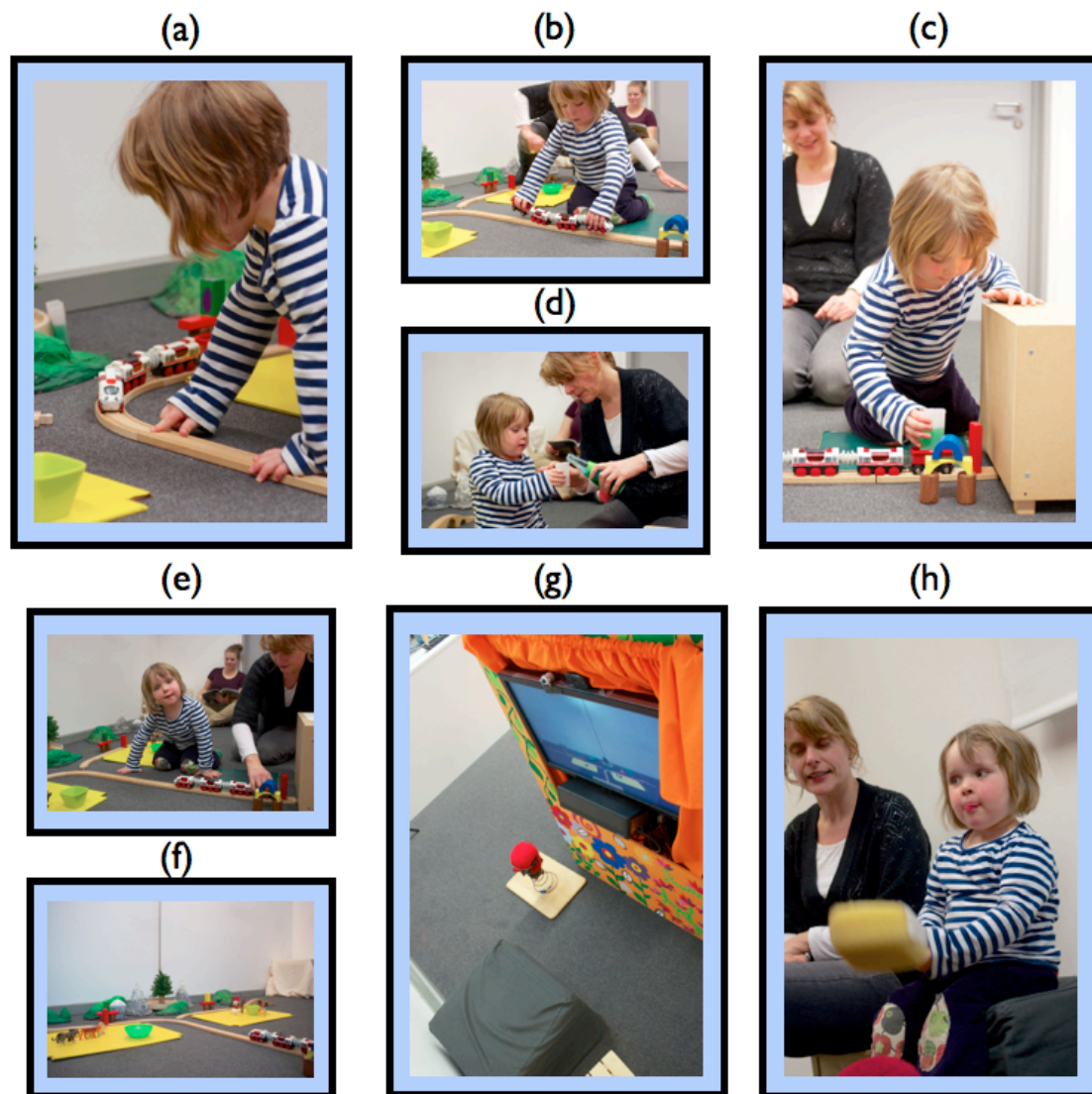


Figure 1. Images illustrating materials and key events of the procedure in study 1.

The child prepares the train tracks to make the train go either to her and the moderator's playing field or to the victim's playing field (the yellow boards on both sides of the track) (a). In the child-harms condition, the child places the train in its initial position and loads the train with water at the beginning of the test trial (b-d). The situation at the beginning of a test trial. The train is prepared. The tracks are set toward the victim's playing field (left side). The green canisters on the playing fields hold the water. At this point the victim's canister is empty. The observer sits off to

the side reading a magazine (e-f). The live feed on the window side when children sat in front of the computer screen. The red cloth button had to be hit with the sponge hammer to start the train (g-h).

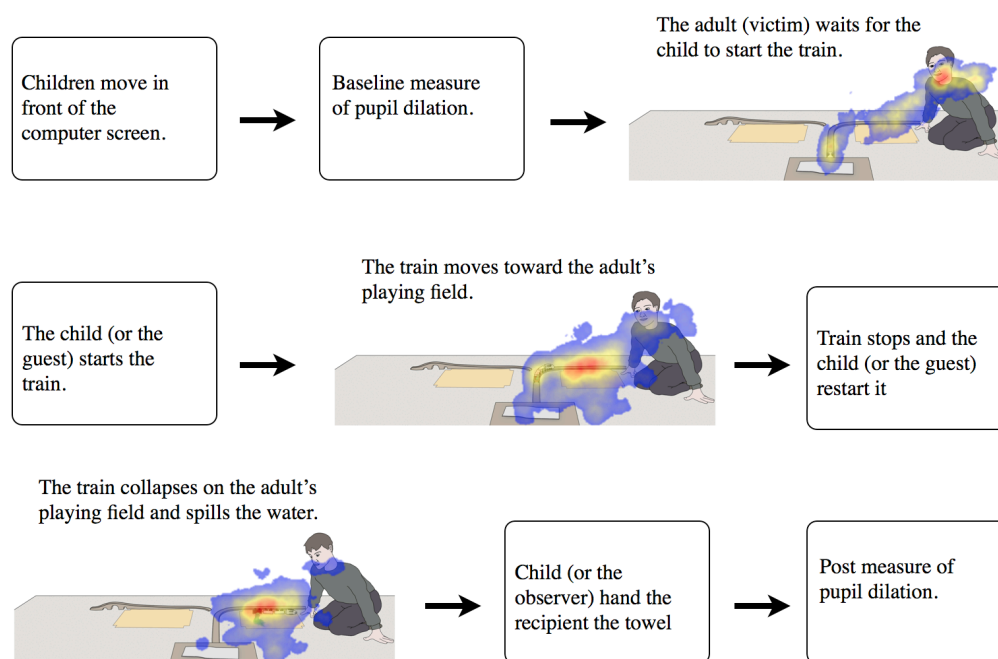


Figure 2. The sequence of events during the test trial for the 3-year-old participants. The illustrations represent key events during the study. Children saw an actual live image of the victim on the computer screen. Across participants the actual live position of the victim may have varied such that the illustrations here represent the prototypical configurations of each element of the scene. In addition to analyzing changes in pupil dilation, we plotted children's viewing patterns while they watched the live feed during the test trials because these data were readily available from the eye tracker (see supplementary materials for details). The purpose was to illustrate the distribution of children's attention during the test trials (see Fig. 2). The resulting gaze-overlay is illustrated by an attention heatmap representing areas with greatest overall number of mapped gaze points (red) and areas with fewest number of gaze points (blue). These gaze data were not analyzed.

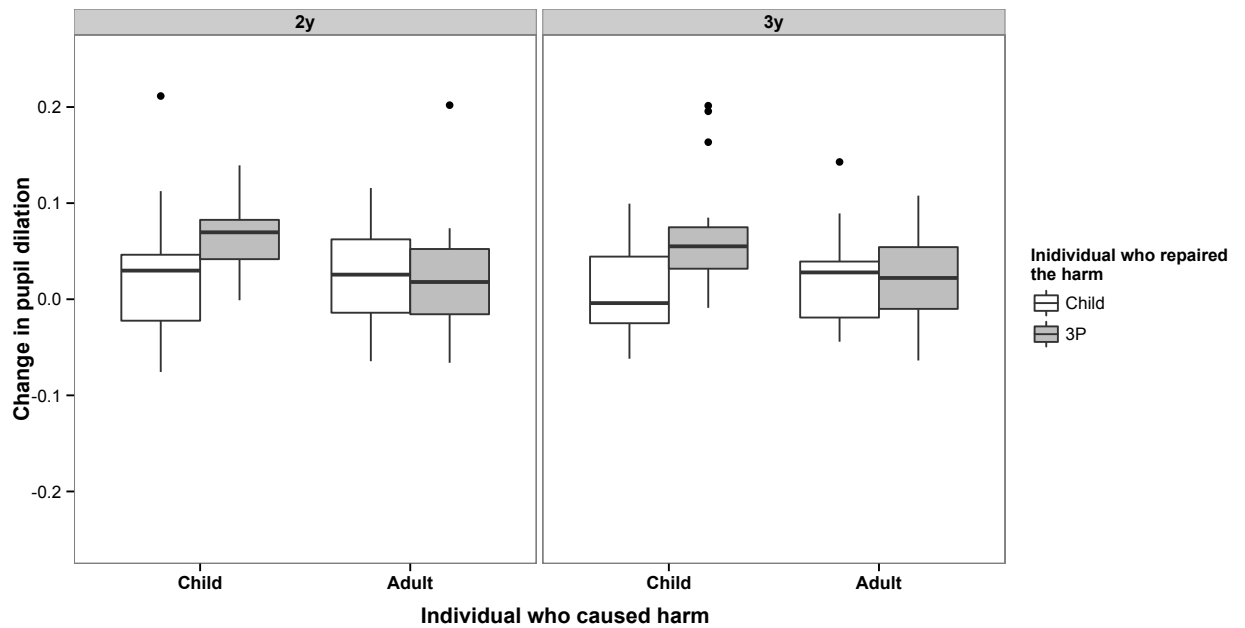


Figure 3. Box-and-whisker plots of the mean relative change in children's pupil dilation as a function of the factors who caused the victim harm (child or adult) and who subsequently got to repair the harm (child or 3P). The upper and lower edge of the boxes mark the first and fourth quartile, respectively. The group medians are represented by the solid lines inside the boxes. The lines above and below each box capture extreme values and the dots represent extreme values which exceed the interquartile distance by more than times 1.58 divided by the square-root of the sample size in each condition (16). Graphics created using the *ggplot2* package (Wickham, 2009) in *R*. * $p < .05$, ** $p < .01$