Young children mostly keep, and expect others to keep, their promises

Patricia Kanngiesser\textsuperscript{1,2,*}, Bahar Köymen\textsuperscript{2,3}, Michael Tomasello\textsuperscript{2}

\textsuperscript{1}Department of Education and Psychology, Freie Universität Berlin, 14195 Berlin, Germany

\textsuperscript{2}Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, 01403 Leipzig

\textsuperscript{3}School of Health Sciences, University of Manchester, Manchester M13 9PL

Journal of Experimental Child Psychology

*Corresponding author. E-mail address: patricia.kanngiesser@fu-berlin.de (P. Kanngiesser)

Acknowledgments

We would like to thank Carl Bartl, Liane Doerr, Valentina Dondoli, Isabelle Gaillande-Mustoe, Jana Jurkat, Esther Klonowski, Solveig Jurkat, Julia Lang, Eva Siegert, Nadine Schumann, Benjamin Sklarek, Anne Tomm, and Gesa Volland for their help running the study and transcribing and coding the data. And thank you to Magherita Isella for helpful comments on earlier versions of the manuscript.
Abstract

Promises are speech acts that create an obligation to do the promised action. In three studies, we investigated whether 3- and 5-year-olds (N = 278) understand the normative implications of promising in prosocial interactions. In Study 1, children helped a partner who promised to share stickers. When the partner failed to uphold its promise, 3- and 5-year-olds protested and referred to promise norms. In Study 2, when children in this same age range were asked to promise to continue a cleaning task – and they agreed - they persisted longer on the task and mentioned their obligation more frequently than without such a promise. They also persisted longer after a promise than after a cleaning reminder (Study 3). In prosocial interactions, young children thus feel a normative obligation to keep their promises and expect others to keep theirs as well.
Young children mostly keep, and expect others to keep, their promises

Promises express commitments to future actions. Austin was the first to systematically describe promises as speech acts, highlighting that they are more than mere words but a way of acting with words (Austin, 1975). In many societies, saying the words “I promise to do A” or simply “I will do A” under the appropriate circumstances creates an obligation to do A (Searle, 1969). Yet what makes promises binding and people keep their word is debated amongst philosophers (Gilbert, 2011; Prichard, 2002; Rawls, 1955; Scanlon, 1990) as well as behavioural scientists (Charness & Dufwenberg, 2006; Ellingsen & Johannesson, 2004; Vanberg, 2008). Some theorists view promises as social practices that play an important role in sustaining social coordination and cooperation (Bicchieri, 2002; Hume, 1890; Lewis, 1969), and, indeed, experimental work has overwhelmingly shown that rates of cooperation in social dilemmas increase substantially when participants promise each other to cooperate (Orbell, Van de Kragt, Alphons, & Dawes, 1988; Ostrom, Walker, & Gardner, 1992; Sally, 1995).

Studies with children have found that [English-speaking] children begin to talk about commitments (“I will do it”) and produce promises from four to five years of age (Astington, 1988b; Diessel, 2004). Around the same age, children adjust their behaviour following an adult’s request to promise: Han Chinese children were less likely to cheat in a competitive game after having promised to follow the rules (Heyman, Fu, Lin, Qian, & Lee, 2015) and North American children revealed transgressions (i.e., report that they played with a forbidden toy) more frequently after having promising to tell the truth (Lyon & Dorado, 2008; Lyon, Malloy, Quas, & Talwar, 2008; Talwar, Lee, Bala, & Lindsay, 2002). Yet, the effect of promises is more variable and potentially more context specific in children under
five years of age: Heyman et al. (2015) found that promises did not reduce cheating rates in four-year-olds in a competitive game. While Talwar et al. (2002, Exp. 3) found a significant main effect of promises on reducing lying behaviour for 3- to 7-year-olds, closer inspection of the data shows that the lying rates decreased only marginally for 3- to 4-year-olds (3-year-olds: promise vs. discussion - 47% lying vs. 50% lying; 4-year-olds: promises vs. discussion - 67% lying vs. 71% lying).

The above findings suggest that from four to five years of age children produce promises and start to keep their own promises, yet, their abilities to correctly reason about promises have been found to develop later: Studies have shown that younger children do not distinguish between the speech act of promising (or committing) and the performance of the act and often prioritize outcomes when judging others’ behaviour (e.g., Astington, 1988c; Kalish & Cornelius, 2007; Rotenberg, 1980). From seven years of age children reason that speakers are responsible for fulfilling what they promised (Astin
gton, 1988c; Maas & Abbeduto, 1998) and take into account whether someone broke a promise (or a commitment) intentionally or accidentally (Maas & Abbeduto, 2001; Mant & Perner, 1988). But it is not until their teens that [English- and French-speaking] children begin to distinguish promises (“I promise I will play with you”) from other types of speech acts such as predictions (“I promise it will rain”) and assertions (“I promise, it rained”) (Astin
gton, 1988a; Bernicot & Laval, 1996).

The overall developmental picture based on the findings to date is that children younger than five years of age mostly fail to keep their own promises nor do they reason that others ought to keep their promissory obligations. However, two lines of evidence challenge this view: (1) studies on young children’s understanding of joint commitments in collaborative, social activities, and (2) studies on young children’s understanding of normativity.
Firstly, recent work on children’s behavior in collaborative social activities has shown that three-year-olds already behave in a committed manner and expect others’ to be committed to a joint endeavor. Specifically, three-year-olds will reengage an adult when she stops to participate in a joint activity, will acknowledge their own leaving of a joint play activity, and will wait, help, and take over the partner’s role after collaboration (Gräfenhain, Behne, Carpenter, & Tomasello, 2009; Gräfenhain, Carpenter, & Tomasello, 2013). However, these studies both elicited an explicit (verbal or non-verbal) agreement from the child to collaborate and engaged the child in a collaborative activity, and compared this collaboration condition to a condition with individual activity (and no agreement). It is thus unclear whether the increased commitment in the collaboration condition (as compared to the individual condition) resulted from the explicit agreement, the collaborative activity or both. In fact, other work has shown that collaboration by itself (without verbal agreements) makes pre-schoolers more committed to finishing a task – both when the partner is present (Hamann, Warneken, & Tomasello, 2012) and when the partner is absent (Butler & Walton, 2013). In our studies, we wanted to focus exclusively on the effect of verbal agreements (i.e., promises) on young children’s behavior. In each study, we therefore used the same prosocial tasks across all conditions and only varied whether a promise was given or not.

Secondly, one important characteristic of promises is that they are normatively binding (i.e., they entail an obligation that one should keep one’s word), and research has shown that young children already comprehend the normativity of some social rules and will hold other’s responsible that do not abide to norms (Schmidt & Tomasello, 2012). For example, 3-year-olds will actively correct a puppet-actor that fails to play a game correctly (Rakoczy, Warneken, & Tomasello, 2008) or will intervene if a puppet tries to take someone else’s property (Rossano, Rakoczy, & Tomasello, 2011), often using norm-related language during their interventions (‘You did that wrong’; ‘You must not do this’). Moreover, pre-
schoolers also start to respond to moral transgressions and judge that it is wrong to harm others (Smetana, 2013). To date, no study has investigated whether young children will also comprehend the normative implications of promising (e.g., that it is wrong to break a promise).

Therefore, we conducted a series of studies investigating children’s understanding of promises focusing on their understanding of other’s promissory obligation (Study 1) and their own commitment after having promised to do something (Studies 2 & 3). In all three studies, children were engaged in prosocial interactions (helping) – either with a puppet-partner (Study 1) or adult experimenters (Studies 2 & 3). We conducted the studies with 3-year-olds who have been shown to comprehend both, joint commitments and the normative implications of social rules (Gräfenhain et al., 2009; Schmidt & Tomasello, 2012). We also included 5-year-olds in all studies to explore whether a more nuanced understanding of promises may emerge later during the preschool years (e.g., Heyman et al., 2015).

**Study 1**

In this study, we investigated whether 3- and 5-year-olds expect others to keep their promises. Three- and 5-year-old children helped a puppet partner to retrieve sticker rewards that only the puppet could access. The puppet either promised to share the stickers (promise group) or only asked for help to retrieve them (control group). In both groups, the puppet then decided to keep all the stickers to itself. Since the outcome was held constant in both groups (i.e., the child did not receive any stickers from the puppet), the meaning of the puppet’s action differed due to the preceding social interaction: it either signified breaking of a promise to share (promise group) or an act of stinginess (control group). We coded children’s spontaneous responses to the puppet’s behaviour, focusing in particular on whether they would refer to the puppet’s obligation or promise to share.
Methods

Participants. Forty 3-year-olds (M: 3;7 years, Range: 3;0 – 3;11 years, 20 girls and 20 boys) and 40 5-year-olds (M: 5;6 years, Range: 4;11 – 5;11 years, 20 girls and 20 boys) took part. Ten additional 3-year-olds and nine additional 5-year-olds were excluded because of failure to protest during the warm up phase of the experiment (nine 3-year-olds, eight 5-year-olds), inattentiveness (one 3-year-old), and experimenter error (one 5-year-old). Children were recruited through a database of parents who had agreed to have their children participate in developmental studies. They lived in a medium-sized German city. They were all native speakers of German and came mostly from middle class families.

Procedure. Children interacted with a puppet partner that was animated by a female experimenter (E1). The experiment was divided in two phases: a warm up phase and a test phase.

In the warm up phase, a second female experimenter (E2) introduced children to the puppet, who then played a short warm-up game together. Next, children experienced three warm up protest trials to test children’s willingness to spontaneously correct the puppet. In each warm up trial, the puppet made an instrumental mistake such as drinking from a cup upside down, brushing the floor instead of her teeth, or holding a brush the wrong way around when sweeping the floor. Children had 30 seconds to correct the mistake, while the puppet became increasingly puzzled about failing to achieve its goal. If children did not spontaneously correct the puppet, they were directly asked to do so at the end of each trial. Only children who spontaneously corrected the puppet in at least one of the three trials participated in the test phase. All other children were dropped from the experiment because they were judged to lack the motivation (or to be too inhibited) to correct the puppet. (For the
17 children that were excluded because of failure to protest during the warm up phase, test-trial data is available for 11 children (5 in the promise group and 6 in the control group). For the remaining 6 children, the procedure was stopped after absence of protest in the warm up trials. For the 11 children with test trial data, there was one instance of protest (out of 3 trials) by a three-year-old boy in the control group. The remaining children never protested and only two children spoke to the puppet once (one asked whether the puppet was real, the other asked the puppet to keep pulling on the rope). This means that there was a protest rate of 1/33=3% in the test phase for this subsample of children (i.e., children that did not protest during warmup). We believe that this warrants our decision to exclude these children from our final sample.)

In the test phase, E2 introduced a box from which stickers could be retrieved (see Figure 1). The stickers were placed on trays inside the box and both partners had to pull a rope together to move the trays and take out the stickers. There was one sticker tray on the child’s side and one on the puppet’s side. Children first participated in one practice trial, in which E2 explained how the box worked and told children that they could collect the stickers in their bowl and take them home later (there were two stickers on each side).

Next, children participated in one trust trial, in which the puppet had all four stickers on its side and either promised to share half the stickers with the child and asked the child to pull the rope (promise group) or only asked the child to pull the rope (control group). In both groups, the puppet handed two stickers to the child to establish trust. This trial was followed by three test trials, in which four stickers were always placed on the puppet’s side. At the start of the trial, the puppet made children aware of the stickers. In the promise group, the puppet then promised to share half of the stickers (Ich gebe dir die Hälfte der Aufkleber ab. Das verspreche ich dir. Also versprochen! ‘I will give you half the stickers. I promise.
Promised!’); and in the control group, the puppet did not say anything. In both groups, the puppet then asked the child to pull on the rope.

Crucially, after having retrieved the stickers, the puppet announced that the stickers were very nice and that it would keep them all. In the promise group, this statement signalled breaking of the puppet’s promise to share with the child. In the control group, it simply stated the puppet’s desire to keep all the stickers. The child had 30 seconds to react and possibly protest against the puppet’s behaviour, after which the puppet gave the child two stickers pretending it had forgotten to do so earlier. This last part was included to depict the puppet as forgetful and friendly (and not as mean) and to help children maintain some level of trust in subsequent trials. The procedure was repeated for each of the three test trials. The entire session (including warm up and test phase) lasted approximately 10-15 minutes.

**Data coding and analyses.** All sessions were videotaped. We coded children’s behaviour and transcribed their utterances verbatim from videotape. We focused on children’s protest in the 30 seconds following the puppet’s announcement that she would keep all stickers. We coded children’s protest behaviour according to the following categories:

1. *Promise normative:* children refer to the puppet’s obligation/promise to share stickers using normative language with deontic modal verbs such as ‘must’ or ‘should’ (e.g., *Aber was man versprochen hat, das muss man auch halten.* ‘But you must keep what you promised.’; *Das darfst du aber nicht.* ‘You mustn’t do this.’)

2. *Promise non-normative:* children referred to the promise or its content in a non-normative manner (*Du hast mir das versprochen* ‘You promised it to me’; *Du hast doch gesagt, du gibst mir die Hälfte ab.* ‘You said you would give me half.’; *Du wolltest welche abgeben.* ‘You wanted to share some.’)
(3) *Personal desire:* children expressed their desire for the stickers by complaining about not getting any stickers or directly requested stickers from the puppet (e.g., *Jetzt hab ich wieder keine.* ‘Now, I don’t have any again’; *Kann ich auch mal auf meiner Seite welche haben?* ‘Can I have some on my side?’; *Gib mir meine 2 Aufkleber* ‘Give me my 2 stickers’)

(4) *Simple:* Children said ‘no’ without further (task related) elaboration.

It was only scored whether protest occurred in a trial or not (we did not score the frequency of protest in a trial). We conducted analyses on two types of protest measures:

1. We analysed children’s overall protest in each trial, scoring if they protested in a trial or not (yes = 1, no = 0). We analysed the binary overall protest data with Generalized Linear Mixed Models (GLMM), using the lme4 package in R (Bates, Maechler, Bolker, & Walker, 2013; R Development Core Team, 2014). Protest (yes=1, no=0) was entered as the response variable. We included a two-way interaction of experimental group (promise vs control) x age group (3 vs 5) as predictor as well as the control predictors gender (male vs female) and trial number (z-transformed). Participant ID was entered as a random effect to account for the repeated measures design (i.e., multiple trials per child). We conducted model comparisons with likelihood ratio tests using the ANOVA method in R. Specifically, we compared a full model (all predictors) and a null model (containing the control predictors gender and trial number). In addition, we calculated protest sumscores for each child by adding overall protest across the three trials (range = 0-3). We compared for each age group whether protest sumscores differed between the promise and the control group using non-parametric tests.

2. We analysed the different types of children’s protest. Here we only coded the highest level of protest per trial with promise norm protest being the highest and simple
protest being the lowest category. For the analyses, we calculated a sumscore for promise protest (normative and non-normative promise) and a sumscore for other protest (desire and simple) for each child, adding the number of trials with promise protest (range= 0-3) and the number of trials with other protest (range= 0-3), respectively. Categories were combined for analyses because of the low occurrence of normative promise protest and of simple protest. In addition, we recorded the number of children who protested at least once in the three trials and scored only the highest level of protest they used. Data were analysed for each age group separately using non-parametric statistics.

All data were coded by the first author. A second coder re-coded 25% of the data for reliability purposes ($N = 20$ children in total, equally distributed by age, experimental group, and gender). Agreement between coders on protest categories was very good (equal-weighted $\kappa = .90$). The irr-package in R was used for reliability calculations (Gamer, Lemon, & Singh, Puspendra, 2012).

**Results and Discussion**

First, we investigated children’s overall protest in the three test trials (see Figure 2). We found that 3-year-olds protested in 40% of trials in the promise group and in 42% of trials in the control group, while 5-year-olds protested in 47% of trials in the promise and in 33% of trials in the control group. We analysed children’s protest behaviour with Generalized Linear Mixed Models (GLMM) and found that a full model did not have a significantly better fit to the data than a null model (including only gender and trial), $\chi^2 = 1.95$, $df = 3$, $p = .584$ (see Electronic Supplement Table 1 for details). Thus neither experimental group nor age group significantly predicted children’s overall protest behaviour. Separate analyses for the two age groups using protest sumscores, showed no significant differences between the promise and
the control group for 3-year-olds, \( U = 210, p = .790 \), and for 5-year-olds, \( U = 164, p = .327 \) (Mann-Whitney U tests). Together these analyses indicate that 3- and 5-year-olds in the promise and in the control group protested similarly about not receiving any sticker rewards.

However, we observed qualitative difference in children’s protest behaviour in the two groups. Figure 2 reports the percentages of all four protest categories. Since normative promise protest and simple protest occurred in less than 10% of the trials, respectively, we used only two categories for the subsequent statistical analyses: promise protest (including normative and non-normative promise protest) and other protest (including desire and simple protest). Only children in the promise group used normative and non-normative promise protest like ‘You promised it to me’ or ‘One shouldn’t break a promise’ (promise protest collapsed: 3-year-olds in 17% of trials, 5-year-olds in 27% of trials), while this type of protest was completely absent in the control group. We summed the two types of promise protest across trials for each child and found a significant difference between the promise and the control group for 3-year-olds, \( U = 150, p = .047 \), and for 5-year-olds, \( U = 100, p < .001 \) (Mann-Whitney U tests). Children in the control group used other types of protest: they protested primarily about not getting any stickers or expressed their desire for more stickers using utterances like ‘I have so few and she has so many’ (other protest collapses: 3-year-olds in 42% of trials, 5-year-olds in 33% of trials); these types of protest occurred less often in the promise group (3-year-olds in 23% of trials, 5-year-olds in 20% of trials). There was a trend for a significant difference in other protest sumscores between the promise and the control group for 3-year-olds, \( U = 258, p = .090 \), and no significant difference for 5-year-olds, \( U = 225, p = .437 \) (Mann-Whitney U tests). Finally, we compared how many children protested at least once in the three trials scoring only the highest level of protest (see Table 1): 25% of 3-year-olds in the promise group used promise protest at least once (25% used other protest) and 65% of 3-year-olds in the control group used other protest at least once, \( p = .008 \) (Fisher
exact test). 50% of 5-year-olds in the promise group used promise protest at least once (10% other protest) and 40% of 5-year-olds in the control group used other protest, \( p < .001 \).

These findings show qualitative differences in children’s protest behaviour in the control and the promise group: while children in the control group primarily voiced their desire for stickers, only children in the promise group referred to promising norms and thought the puppet responsible for not sharing. Thus, by three years of age some children in our sample (25%) expect a promisor to perform the promised act and these expectations became more pronounced in five-year-olds (50% of five-year-olds in the sample). It should be noted that 9/50 (18%) of 3-year-olds and 8/49 (16%) of 5-year-olds were excluded from the sample because of failure to protest in the warm-up, which potentially limits the generalizability of our findings.

It is possible that children protested differently in the control and the promise group, because only in the promise group were children told that they would receive half of the stickers (the puppet never mentioned sharing in the control group). Prior work has shown that 3-year-olds share rewards equally after collaborating on a task that is identical to the one used in our study (Hamann, Warneken, Greenberg, & Tomasello, 2011; Warneken, Lohse, Melis, & Tomasello, 2011). This suggests that children in our study likely had an expectation to receive stickers in both conditions. This claim is supported by our finding that children’s overall protest behaviour did not differ between the promise and the control group, indicating that children in both groups were equally dissatisfied with not receiving any stickers. But crucially, only in the promise group did some children mention the puppet’s obligation to share.

Earlier work had shown that it is not until age seven (possibly somewhat earlier) that children reason that speakers are responsible for fulfilling promises (“I promise I will play with you”) but not for fulfilling predictions (“I promise it will rain”; Astington, 1988c; Maas
& Abbeduto, 1998). We show that pre-schoolers already assume that promises obligate a speaker to perform the promised act – though our study design does not allow us to draw any conclusions as to whether pre-schoolers differentiate promises and predictions. Yet, a series of related studies has shown that pre-schoolers direct their protest at actors for failing to follow a directive (e.g., ‘Put the red thing there!’) and at speakers for failing to make accurate predictions or assertions (e.g., ‘The red thing is/will be there’) (Lohse, Gräfenhain, Behne, & Rakoczy, 2014; Rakoczy & Tomasello, 2009). Taken together this indicates that young children may have a more nuanced understanding of speech acts and speaker’s responsibility than previously assumed. One of the questions left open is whether the different developmental onsets arise from methodological differences between studies (action based methodologies vs. interviews about third party stories) or are indicative of children’s conceptual developmental (implicit, action-based understanding vs. explicit knowledge about the correct use of promises). Future research could help to clarify this question by combining first party action-based and third party judgement-based methodologies in studying young children’s understanding of different speech acts.

**Study 2a**

While the findings from Study 1 indicate that from three years of age children expect others to keep their promises in a prosocial interaction, evidence to date suggests that children younger than 5 years of age mostly fail to keep their own promises (Heyman et al., 2015; Talwar et al., 2002). However, in these studies children were asked to promise to refrain from engaging in forbidden behaviours such as cheating or lying. It is an open question whether the findings indicate a general inability to keep promises in younger children (e.g., due to limited inhibitory control skills) or whether they indicate a context-specific inability (e.g., limited to forbidden behaviours). In fact, another line of evidence has shown that three-year-
olds behave committedly after a joint commitment to a collaborative task (Gräfenhain et al., 2009; Gräfenhain, Carpenter, & Tomasello, 2013) suggesting that contexts could play a role in scaffolding children’s promise keeping.

We therefore wanted to investigate whether young children will keep their own promissory obligations using a collaborative helping task. To get a graded measure of children’s commitment to the helping task, we used increasingly persistent cues to distract children from the task. Specifically, we engaged 3- and 5-year-olds in a cleaning task with an adult experimenter. When the experimenter had to leave the task, she either asked children a series of prompt questions to elicit a spontaneous promise to continue cleaning (promise group) or immediately left without asking any prompt questions (control group). A second experimenter then tried to entice children away from the cleaning task with an exciting toy. We expected children in the promise group to resist temptation longer than children in the control group.

Methods

Participants. Thirty-nine 3-year-olds ($M = 3;9$ years, $Range = 3;6 – 3;11$ years, 19 girls and 20 boys) and 40 5-year-olds ($M = 5;9$ years, $Range = 5;5 – 5;11$, 20 girls and 20 boys) took part in the study. Two additional 3-year-olds were excluded: one due to a technical error (video camera not switched on) and one who never committed to the task in the promise group. Children were recruited in the same manner as in study 1.

Procedure. Children were tested individually in a quiet room of their nursery. Half of the children were assigned to the promise group and the other half to the control group. Children in both groups first played a familiarisation warm-up puzzle with the two female experimenters (E1 and E2). Next, E1 showed the child a marble track in one corner of the
room, while E2 left the room (pretending to look for a kindergarten head teacher). The marble track was colourfully decorated and made jingle sounds when marbles were thrown in (see Figure 3). While the child was distracted playing with the marble track, E1 walked backwards and pretended to spill a bucket of shredded paper on a carpet. After the accident, E1 drew the child’s attention to her mishap, explained that she would need a clean carpet later on and asked the child whether she could help with the clean-up. It should be noted that this means that children in both groups (control and promise group) were initially asked to commit to helping E1 with the cleaning task. E1 and child spent about 15 seconds cleaning the carpet, when E2 entered the room and told E1 that the nursery school teacher wanted to talk to her about something important. This was introduced to give E1 a credible excuse to leave the room.

Before she left the room, E1 asked children in the promise group a series of prompt questions to elicit a spontaneous commitment to continue the cleaning task. The following questions were asked in a fixed sequence: (1) What should we do? Do you have any idea? (*Was sollen wir machen? Hast du eine Idee?*), (2) Who will keep cleaning? (*Wer räumt hier den weiter auf?*), (3) Could you maybe do it? (*Könntest du das vielleicht machen?*), (4) Do you promise to keep cleaning? Promised, ok? (*Versprichst du mir, dass du weiter aufräumst? Also versprochen?*). If children spontaneously committed after one of the earlier prompt questions, all remaining prompt questions except the promise question (question 4) were omitted. That is, all children in the promise group were eventually asked to promise. E1 then announced that she would be back soon. In the control group, E1 omitted all prompt questions and simply stated that she would be back soon.

Next, children in both groups took part in the distraction phase of the experiment. While E1 left the room, E2 sat down in front of the marble track with her back towards the child. She then spent 60 seconds trying to entice the child away from the cleaning task using
increasingly explicit distraction cues (one cue every 15 seconds). (1) After 15 seconds, E2 simply commented that the track was great (*Ah, toll.*), (2) after 30 seconds, she threw a marble in the track (which produced a jingle sound) commenting that this is how the marble track worked (*Ah, so funktioniert das.*), (3) after 45 seconds, E2 threw another marble in the track, turned to look at the child and said that the marble track was great (*Oh schau mal, das ist eine tolle Murmelbahn.*), and (4) after 60 seconds, E2 turned again towards the child and asked whether the child wanted to play (*Spielst du mit?*). If the child refused to play, E2 asked her why not (*Warum denn nicht?*). At this point E1 returned, thanked the child for cleaning and allowed her or him to play with the marble track. The entire test session lasted approx. 10 minutes.

**Data Coding and Analyses.** The study was videotaped and children’s behaviour was later coded from videotape. For children in the promise group, we coded how many prompt questions were asked before children committed to the cleaning task using a score from 0 to 4 (with higher scores indicating increasing levels of prompting). For example, children received a score of 0 if they volunteered to help before E1 had asked the first prompt question and they received a score of 4 if they only committed after being asked to promise (question 4: Do you promise to keep cleaning?).

For children in both groups, we coded how committed they were to continue the cleaning task. We calculated a commitment score from 0 to 5 based on the number cues E2 needed to use in order to distract the child from her task, with higher scores indicating greater levels of commitment. Distraction was scored if the child started playing with the marble track (e.g., by inserting a marble) or if she stopped cleaning and left the carpet for more than 15 seconds (we scored the distraction cue at the start of the 15 seconds). For example, children received a score of 0 if they left the carpet immediately after E1 had left the room, a
score of 1 if they left after the first distraction cue, and a score of 5 if they kept cleaning even after the final distraction cue (i.e., after E2 asked whether they wanted to play and why not).

We also coded whether children referred to their promise or the cleaning task at any point during the distraction phase using the following coding scheme:

(a) Promise: Children explicitly refer to their promise to clean. For example, *Hab ich versprochen* (‘I promised’).

(b) Normative cleaning: Children use normative language when referring to their commitment to do the cleaning task. For example, *Aber ich muss das aufräumen.* (‘But I have to clean this’), *Ich muss das hier alles fertig machen.* (‘I have to finish all of this’).

(c) Simple (non-normative cleaning): Children refer to their cleaning without using normative language. For example, *Ich räume hier auf.* (‘I am cleaning’), *Ich mach die Schnipsel.* (‘I am doing the paper shreds’).

If children used different types of protest, we only coded the highest level (promise being the highest and simple being the lowest category).

Data were analysed in R using the exactRankTests package for all non-parametric tests, the ordinal package for all ordinal linear models, the Kendall package for correlations and the irr package for reliability calculations (Christensen, 2015; Gamer, Lemon, & Singh, Ian Fellows Puspendra, 2012; McLeod, 2011). We used the Kendall package (Kendall-function) and calculated Kendall rank correlations because our data was ordered and contained ties. The Kendall-function (1) computes a tau-b coefficient which is suited for tied data and (2) approximates a p-value for tied data that is an adequate approximation of an exact p-value. In the ordinal linear models, we entered commitment scores as response variable (ordered factor from 0 to 5); age group (3 vs. 5), experimental group (promise vs. control) and their two-way interaction as predictors; and gender (female vs. male) as a control
predictor. We conducted model comparisons with likelihood ratio tests using the ANOVA method in R. Specifically, we compared a full model (all predictors) and a null model (control predictor gender) as well as a full and a reduced model (excluding certain predictors).

A second coder scored 25% of the data for reliability purposes (equally distributed by age group, experimental group, and gender). Agreement between observers was good for children’s responses to the prompt question in the promise group, i.e. at which question children agreed to do the cleaning task (weighted $\kappa = .71$; $N_{\text{reliability}} = 10$ children). All other reliability analyses were calculated together with data from study 2b as the same reliability coder coded the three groups ($N = 30$ children in total, equally distributed by age group, experimental group, and gender). Agreement between observers was good for children’s commitment scores in the distraction phase (equal-weighted $\kappa = .80$) and for children’s spontaneous mentioning of their commitment (equal-weighted $\kappa = .74$).

**Results and Discussion**

First, we analysed how many prompt questions children in the promise group needed before committing to the cleaning task. 90% of the 3-year-olds and 80% of the 5-year-olds, respectively, only committed after they were asked the third prompt question (‘Could you maybe do it?’; see Table 2 for details). The number of prompt questions did not differ significantly between 3-year-olds ($M = 3.0$, $SD = 0.33$) and 5-year-olds ($M = 2.75$, $SD = 0.85$), $U = 209$, $p = .401$ (Mann-Whitney U test). There was no significant correlation between children’s willingness to commit to the task (i.e., the number of prompt questions asked) and how committed they were to completing the task in the distraction phase (3-year-olds: $\tau_B = -0.23$, $p = .303$; 5-year-olds: $\tau_B = 0.08$, $p = .737$; Kendall correlation). No child in the promise group spontaneously used the word *versprechen* (‘promise’), but all children in this
group were eventually asked by E1 to promise to complete the task. Children in the control group were never asked any prompt questions.

Next, we analysed children’s behaviour in the distraction phase, finding that children in the promise group were more committed to continue the cleaning task than children in the control group (see Figure 4). We analysed children’s commitment scores using ordinal linear regression models finding that a full model had a significantly better fit to the data than a null model (including only the control predictor gender), $\chi^2 = 16.28$, $df = 3$, $p < .001$ (see Electronic Supplement Table 2 for details). Next, we dropped the non-significant interaction term between age group and experimental group from the full model, which did not change the model’s fit significantly, $\chi^2 = 0.01$, $df = 1$, $p = .917$. The reduced model revealed that children in the promise group had significantly higher commitment scores than children in the control group, $Z = 3.39$, $p < .001$, and that 5-year-olds had significantly higher commitment scores than 3-year-olds, $Z = 2.03$, $p = .042$. In addition, there was a trend for girls to have higher commitment scores than boys, $Z = 1.82$, $p = .069$. Separate comparisons for each age group revealed a non-significant difference between the control and promise group for 3-year-olds, $U = 126$, $p = .055$, and a significant difference for 5-year-olds, $U = 100$, $p = .005$ (Mann-Whitney U tests). Taken together, these findings indicate that promises resulted in higher levels of commitment to complete the task – with 5-year-olds being clearly committed and 3-year-olds showing some signs in this direction.

In the distraction phase, some children spontaneously mentioned their commitment to the cleaning task. There were no significant age differences, $p > .999$ (Fisher exact test). In the promise group, 14 children (36%) mentioned their commitment to the task (see Table 3 for further details). The majority of those children used normative language (e.g., *Aber ich muss das aufräumen*, ‘But I have to clean here’; $N_{normative} = 11$) and two of those children explicitly referred to their promise. In contrast, significantly fewer children ($N_{total} = 6$ [15%]
with $N_{normative} = 4$) mentioned their commitment in the control group, $p = .041$ (Fisher exact test). Overall, children’s justifications revealed that more children in the promise group than in the control group viewed the cleaning task as a normative obligation.

To summarize, we found that 3- and 5-year-olds who were asked to promise to continue a tedious cleaning task stayed committed for longer and mentioned their commitment more frequently than children who were not asked to promise. Even though we were unable to elicit spontaneous promises from children, the adult experimenter’s request of a promise (“Do you promise to clean?”) induced a promissory commitment. Children’s ability to keep elicited promises thus emerges earlier than previously found (e.g., Heyman et al., 2015). There are a number of possible explanations for why our findings differ from the findings of previous studies: (1) It is conceivable that our cooperative helping task reduced incentives for young children to defect on their promise as compared to the competitive games or transgression situations of previous studies (Heyman et al., 2015; Lyon & Dorado, 2008). (2) Irrespective of the task-context, our series of distraction cues may have provided us with a more fine-grained measure of children’s behaviour than the dichotomous measure (e.g., cheat yes/no) used in other studies. (3) Alternatively, our study may have differed from previous studies in the extent to which young children had to exercise self-control in the face of competing desires. Future studies could, for example, directly compare promise keeping abilities in competitive vs cooperative contexts to disentangle these different accounts.

One question that we were not able to address (because we failed to elicit spontaneous promises) is whether children distinguish between spontaneous and elicited promises. Research with adults has shown that promises that are uttered spontaneously in free flowing communication are more effective in enhancing trust and cooperation than “bare promises” (i.e. sending a pre-written message with the words “I promise”; Charness & Dufwenberg, 2006, 2010). Moreover, developmental work on children’s helping behaviour has revealed
that five-year-olds (but not 3-year-olds) helped longer with a task when they were given a choice to help as compared to being asked to help (Rapp et al., submitted). Together, these findings tentatively suggest that spontaneous promises may create even stronger obligation for older preschoolers than elicited promises. This remains an interesting avenue for future research to investigate.

One developmental pattern emerged in our study: irrespective of condition 5-year-olds in our study were more committed to the task than 3-year-olds. Similarly, Gräfenhain et al. (2009) found that 4-year-olds took longer to leave a (joint or parallel) game than 3-year-olds. In addition, studies on the development of children’s self-control have shown that with increasing age children are better able to regulate their actions (e.g., stop repeating a previously successful action that is no longer effective) and attention (e.g., focus despite a distracting noise; Herrmann, Misch, Hernandez-Lloreda, & Tomasello, 2014). Thus, older children are better at focusing on their current task/game and are less easily distracted to explore a more attractive activity, which may explain the developmental findings in our study. The possibility remains that these observed age differences are specific to the distractor task we used (i.e., playing with an attractive toy) and that a different task – for example, a commitment to overcome an obstacle or persist on a difficult puzzle - may reveal no age related changes in overall commitment.

So far, we had always used the word versprechen (‘promising’) when investigating pre-schoolers’ understanding of promissory commitments. While saying “I promise.” or “Do you promise?” is usually the most unambiguous way of signalling a commitment, utterances such as “I will do it.” or “Will you do this?” can fulfil the same function (Searle, 1969). In fact, developmental work has shown that 3- to 10-year-old French children view future-directed speech acts (“I will do it”) and assertions with a predictive content (“It will be done”) as promises (Bernicot & Laval, 1996). However, under some circumstances such as
when statements containing “promise” and “will” provide conflicting evidence (e.g., regarding the location of a hidden toy), older but not younger children trust promise-statements more than will-statements (Lyon & Evans, 2014). To further investigate whether the word “promise” is necessary to elicit task commitment in our study, we conducted a follow-up study with 3- and 5-year-olds where we elicited commitments without asking children to promise.

**Study 2b**

In this follow-up to Study 2a, we explored pre-schoolers’ commitment to the cleaning task when commitments are elicited without the word “promise”. We repeated the experiment from Study 2a, but excluded the last prompt question and only asked children “Could you maybe do it?”. We chose this question because the majority of children had committed to the cleaning task after this prompt in Study 2a.

**Methods**

**Participants.** Nineteen 3-year-olds (\(M = 3;9\) years, \(Range = 3;6 – 3;11\) years, 10 girls and 9 boys) and 20 5-year-olds (\(M = 5;9\) years, \(Range = 5;4 – 5;11\) years, 11 girls and 9 boys) took part in the study. Two additional 3-year-olds were excluded because they were too shy and did not fully participate in the study. Six additional 5-year-olds were excluded because they spontaneously committed to the cleaning task before the last prompting question was asked. Children were recruited in the same manner as in the previous studies and tested in a quiet room in their local kindergarten.

**Procedure.** The procedure was identical to Study 2a with the exception that we did not ask children to promise to continue cleaning (commitment-without-promise group). Specifically,
E1 asked all the prompting questions up to “Could you maybe do it?” in quick succession (prompt question 3). Children who spontaneously committed to the task before the final prompt question were excluded from the study (6 5-year-olds). The study was conducted by the same two female experimenters as Study 2a.

**Data Coding and Analyses.** Data coding and analyses and reliability coding were identical to Study 2a. We included the data from Study 2a in all statistical analyses. Results of the reliability analyses are reported in Study 2a (section Data Coding and Analyses).

**Results and Discussion**

An overview of children’s commitment scores can be found in Figure 4. We analysed children’s commitment scores using ordinal linear regression models: a full model (with age group, experimental group and their two-way interaction) had a significantly better fit to the data than a null model (including only the control predictor gender), $\chi^2 = 19.39, df = 5, p = .002$ (see Supplementary Table 3). Removing the non-significant two-way interaction between experimental group and age group did not significantly change the reduced model’s fit to the data, $\chi^2 = 0.33, df = 2, p = .850$. The reduced model revealed that children in the current study had significantly higher commitment scores than children in the control group (Study 2a), $Z = 2.60, p = .009$; but commitment scores did not differ significantly from children in the promise group (Study 2a), $Z = 1.20, p = .232$. We also found that 5-year-olds had significantly higher commitment scores than 3-year-olds, $Z = 2.19, p = .029$, and girls than boys, $Z = 2.57, p = .010$. Eliciting a commitment from children without using the word “promise” thus had a similar effect on young children’s behaviour as eliciting a promise.

Moreover, we coded whether children spontaneously mentioned their commitment and found that only six children (15%) did so in the current study (see Table 3). These were fewer children than in the promise group (Study 2a), $p = .068$, and similar numbers of
children compared to the control group (Study 2a), \( p > .999 \). It is possible that the omission of the word promise made the normative dimension of the commitment less salient for children.

Our findings suggested that children are committed to the task even when the word “promise” is omitted. This is in line with previous work on French-speaking children’s understanding of promises and future-directed commissive speech acts (“I will do it”; Bernicot & Laval, 1996) and supports theoretical work arguing that using the word “promise” is not a necessary condition for an utterance to count as a verbal commitment (Searle, 1969).

In studies 2a and b, children were always asked a series of prompt questions to elicit a commitment. We had introduced the prompt questions to induce spontaneous promises from children, but the possibility remains that the questions simply reminded children of the cleaning task and thus resulted in higher commitment compared to the control group. To control for this alternative explanation of our findings, we conducted a third study with 3- and 5-year-olds, in which we contrasted a promise to clean (without repeated prompts) with a cleaning reminder (see Heyman et al., 2015, for a similar control).

**Study 3**

In this study, we investigated 3- and 5-year-olds’ promissory commitment to help in a collaborative cleaning task, contrasting a promise to clean (promise group) with a cleaning reminder (control group). We expected that children in the promise group would be more committed to the task than children who only received a cleaning reminder.

**Methods**

**Participants.** Forty 3-year-olds \((M = 3;8 \text{ years}, \ Range = 3;5 - 3;10 \text{ years, 21 girls and 19 boys})\) and forty 5-year-olds \((M = 5;8 \text{ years}, \ Range = 5;4 – 5;11 \text{ years, 20 girls and 20 boys})\)
took part in the study. Five additional 3-year-olds were excluded because they did not want to participate (3 children) or to promise (2 children). One additional 5-year-old was excluded because s/he needed to use the bathroom during the study. Children were recruited in the same manner as in the previous studies and tested in a quiet room in their local kindergarten.

**Procedure.** The procedure was identical to Study 2 with the following exceptions:
(1) In the promise group, children were directly asked to promise to clean without using the series of prompt questions (*Versprichst du mir, dass du hier aufräumst?* ‘Do you promise to clean here?’). In the control group, they were reminded that there was some cleaning left to do (*Es gibt hier noch was zum Aufräumen* ‘There is something to clean here’).
(2) We used a different marble track (a wooden one; see Electronic Supplementary Figure 1) and a new team of experimenters (E1 was male and E2 was female).

**Data Coding and Analyses.** Data coding and analyses and reliability coding were identical to Studies 2a and b. Some children spontaneously made generic normative statements during the distraction phase (e.g., *Der Teppich muss ganz sauber sein.* ‘The carpet needs to be completely clean.’). We did not score these utterances in the normative cleaning category as they did not refer to the child’s own cleaning obligation.

A second coder scored 25% of the data (*N* = 20 children) for reliability purposes (equally distributed by age group, experimental group, and gender). There was good agreement between observers for children’s commitment scores in the distraction phase (equal-weighted κ = .86) and for children’s spontaneous mentioning of their commitment (equal-weighted κ = .95).
Results and Discussion

When we compared a direct promise with a cleaning reminder, we found that children in the promise group had on higher commitment scores than children in the control group (see Figure 5). Ordinal linear regression models revealed that a full model had a significantly better fit to the data than a null model (including only the control predictor gender), $\chi^2 = 14.47, df = 3, p = .002$ (see Electronic Supplement Table 4 for details). Next, we dropped the non-significant age group and experimental group interaction from the full model, which did not change the model’s fit significantly, $\chi^2 = 0.06, df = 1, p = .800$. The reduced model revealed that children in the promise group had significantly higher commitment scores than children in the control group, $Z = 3.39, p < .001$. There was a trend for 5-year-olds to have higher commitment scores than 3-year-olds, $Z = 1.66, p = .097$. Separate analyses for the two age groups showed a significant difference between the promise group and the control group for both, three-year-olds, $U = 130, p = .048$, and five-year-olds, $U = 107, p = .007$ (Mann-Whitney tests). These findings confirm that from three years of age children’s commitment to continue the cleaning task increases after an elicited promise and that older children persist overall longer on the task than younger ones.

Children spontaneously mentioned their commitment to the task during the distraction phase. There was no significant age difference, $p > .999$ (Fisher exact test). In the promise group, 21 children (53%) mentioned their commitment to the task (see Table 3 for further details). The majority of those children used normative language ($N_{\text{normative}} = 14$) and one child explicitly referred to his/her promise. In contrast, significantly fewer children ($N_{\text{total}} = 6 [15\%]$ of which $N_{\text{normative}} = 5$) mentioned their commitment in the control group, $p < .001$ (Fisher exact test).

These results support and extend our previous findings: from three years of age children were more committed to a collaborative cleaning task after an elicited promise than
after a cleaning reminder. This supports our claim that children’s higher commitment scores in the promise and commitment group in Studies 2a & b were a result of their verbal commitments and not simply due to the prompt questions reminding children of the cleaning task. Moreover, our results extend the findings of Heyman and colleagues (2015) to a younger age group, who showed that 5-year-olds refrained from cheating on a card game more after an elicited promise than after a rule reminder. Commitment scores in this study were overall higher than in Study 2, which may have a due to variations in procedural details (wording, materials, experimenter team) – importantly though significant differences between commitment in the control and the promise groups were replicated.

It is possible that the control condition in this study – reminding children that there was something to clean – as well as the control condition in Study 2a created some uncertainty on the part of the child about what behaviour was expected of them. For example, children may have assumed that the adult experimenter expected them continue with the cleaning task. Indeed, previous work has shown that collaborative or joint activities induce commitments in young children (Butler & Walton, 2013; Hamann et al., 2012). It is thus likely that children in the control conditions in Study 2a and 3 felt some commitment to the joint cleaning tasks. Importantly, however, our findings in both studies clearly show that promises or explicit verbal commitments create normative obligations that are stronger than the (implicit) commitments elicited by collaborative activities.

**General Discussion**

We investigated German pre-schoolers’ understanding of other’s promissory obligations and their ability to keep their own promises and demonstrated across three studies that pre-schoolers understand that promises entail a normative obligation to fulfil the promised act. Specifically, they protested and referred to promise norms when a puppet-partner broke a
sharing promise (Study 1) and helped an adult longer in a collaborative cleaning task after an elicited promise or verbal commitment (Studies 2 & 3). Our findings support the view that children begin to comprehend promissory obligations in social interactions around three years of age, an earlier age than what is documented in the literature (e.g., Astington, 1988a; Heyman et al., 2015). Furthermore, they suggest that promises in prosocial or collaborative situations may help young children internalize a sense of obligation to helping others, and that these promises have important normative implications.

We propose that children’s early understanding of promises develops through their participation in joint social activities. Their understanding of promises is preceded by their understanding of joint intentions, actions and goals that emerges during the first 18 months of life (Bruner, 1975; Carpenter, 2009; Tomasello et al., 2005). By three years of age, preschoolers who are engaged in joint activities have been found to acknowledge their leaving of a task to their partner, to help their partner achieve her goal, and to persist longer in a challenging task, indicating a sense of commitment to the joint endeavour (Butler & Walton, 2013; Gräfenhain et al., 2009; Gräfenhain et al., 2013; Hamann et al., 2012; Warneken, Gräfenhain, & Tomasello, 2012). In addition, our findings show that when promises are uttered during joint activities, young children are committed to their own promises and expect that others ought to keep theirs.

Interestingly, pre-schoolers in our studies participated in joint activities both in the control condition and the promise condition: they jointly retrieved rewards with a puppet or cleaned a carpet with an adult. Based on the findings cited above, one would expect that in the control conditions children may have developed a sense of commitment to the joint endeavour. However, when promises were involved in addition to joint activities, children referred more frequently to promise norms or normative obligations, indicating an understanding of the normative implications of promising (beyond a mere sense of
commitment). While joint activities may thus lead to expectations about one’s own and other’s behaviour such as expecting someone to play their part in a joint activity and to contribute to the joint goal (Michael, Sebanz, & Knoblich, 2015), promises possess a strong *normative* force (i.e., that one *ought to* keep them). That is, promises or explicit verbal commitments create normative obligations that are stronger than the implicit commitments elicited by joint activities.

Arguably, we only investigated children’s understanding of promises in interpersonal contexts and did not study whether pre-schoolers will refer to or enforce promise norms as uninvolved third parties. The later behaviour is one indication that children understand that norms do not only apply to themselves but also to other members of their group. There is a growing body of developmental work showing that from three years of age, Western children actively enforce social and moral norms towards third parties in a variety of contexts such as game rules or ownership norms (Rakoczy et al., 2008; Rakoczy & Schmidt, 2013; Rossano et al., 2011; Tomasello & Vaish, 2013). This suggests that pre-schoolers understand in principle that norms apply to all individuals within similar contexts. Furthermore, we take the fact that pre-schoolers in the promise conditions in our studies used normative language (‘I have to clean here’, ‘You shouldn’t do this’) and not just statements of personal liking or disliking (‘I want to clean’, ‘I want more stickers’) as evidence of a normative understanding of promises.

That promises and verbal commitments possess normative force is remarkable - or as Prichard (2002) stated: “Once call some act a promise and all question of whether there is an obligation to do it seems to have vanished.” There are two mechanisms that may explain why people keep their word: external sanctions or internalized norms. The external sanctions account is dismissed by scholars on the grounds that many studies have shown that adults keep their promises in the absence of third party sanctions (Orbell et al., 1988; Ostrom et al., 1992); instead, different psychological mechanisms have been suggested to explain this
behaviour such as an aversion to let others down (Charness & Dufwenberg, 2006) or a preference for keeping one’s word (Ellingsen & Johannesson, 2004). Our findings that young children already reference norms and obligations (and not adult authority or fear of punishment) support the psychological view and suggest that promise norms begin to be internalized early in development – although our data does not allow us to differentiate between different mechanistic accounts (e.g., disappointing others, preference for word keeping, fear of others’ disapproval).

The finding that young children do already understand some normative implications of promising, does not imply that there are no developmental changes in promise-related behaviors or that children always keep their promises. For one, we found that across conditions older children were more committed to our cleaning task (Studies 2 and 3), which is most likely related to developmental changes in children’s abilities for self-control (Herrmann et al., 2014). This explanation has been championed by other authors (Heyman et al., 2015) to explain why younger children in their study were not affected by promises. We suggest that in addition to self-control, contextual factors (e.g., cooperative vs. competitive situations, task demands) influence whether young children avoid the temptation to renge on their promises. Furthermore, with age children’s reasoning about promises becomes more sophisticated and they increasingly consider intentions when judging whether it is OK for others to break a promise or not (Maas & Abbeduto, 2001; Mant & Perner, 1988). One may also expect that older children’s decisions to break their own promises could be influenced by whether they believe that they have a good reason to do so (e.g., it is OK to break a promise to help someone in immediate need but not OK to break a promise to play a fun game). Even in the absence of good reasons for breaking a promise, children (and adults) may sometimes break promises to enhance their self-interest or the interests of those close to them.
One final caveat is that most of the work on promises and norms mentioned in this paper (including our own) focuses on children and adults in large-scale Western societies. One cross-cultural study on promises has found that expectations about when it is obligatory to comply with promises and preference for how to communicate promises vary across German, Chinese and Tongan participants (Beller, Bender, & Song, 2009). Moreover, past work has documented cross-cultural variation in whether children prioritize general normative principles such as justice (US children) or personal obligations (Indian) when judging moral dilemmas (Miller & Bersoff, 1992). It is thus plausible that similar cultural variation may exist with regards to whether promises are understood as societal norms or (inter-)personal obligations. Future work is needed to better understand promise-related practices and beliefs in different societies and how these shape children’s developing understanding of promises.
References


doi:10.1111/desc.12272


doi:10.1016/j.jecp.2015.04.013


doi:10.1371/journal.pone.0086958


doi:10.1177/0956797610395392
Tables

Table 1 *Number of children who protested at least once in the three test trials in Study 1 (reporting only the highest level of protest that occurred)*

<table>
<thead>
<tr>
<th>Age</th>
<th>Condition</th>
<th>Promise</th>
<th>Norm</th>
<th>Personal Desire</th>
<th>No Protest</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Promise</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>0</td>
<td>13</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Promise</td>
<td>10</td>
<td>2</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>0</td>
<td>8</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>

Table 2 *Number of children in the promise group in Study 2a who spontaneously committed to continue with the cleaning task by the different prompt questions. All children were eventually asked to promise. (Questions are abbreviated)*

<table>
<thead>
<tr>
<th>Age</th>
<th>(0) No question</th>
<th>(1) What should we do?</th>
<th>(2) Who will clean?</th>
<th>(3) Could you do it?</th>
<th>(4) Do you promise?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>17</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>16</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3 *Number of children who spontaneously mentioned their commitment to the cleaning task in the distraction phase of studies 2 and 3*

<table>
<thead>
<tr>
<th>Study</th>
<th>Age</th>
<th>Condition</th>
<th>Promise</th>
<th>Normative</th>
<th>Simple</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Study 2</td>
<td>3</td>
<td>Promise</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Control</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Commit w/out promise</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Promise</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Control</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Commit w/out promise</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>Study 3</td>
<td>3</td>
<td>Promise</td>
<td>0</td>
<td>8</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Control</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Promise</td>
<td>1</td>
<td>6</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Control</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>20</td>
</tr>
</tbody>
</table>
Figures

Figure 1. The set-up used in study 1. The child sat on the orange pillow. The box was covered with Plexiglas and the child and the puppet had to pull the rope together to move the board with the stickers towards them. The stickers could then be retrieved through the hole(s) in front of the puppet and the child, respectively.
Figure 2. Percent of trials in which 3- and 5-year-olds in the promise and in the control groups referred to the promise norm (light gray bar) or used other types of protest (personal desire or simple protest combined in white bar).
Figure 3. Set-up used in Study 2. E1 accidentally spilled paper shreds on the blue carpet and asked the child to help clean up. E2 later sat in front of the marble track on the right and tried to distract the child from the cleaning task.
Figure 4. 3- and 5-year-olds’ commitment scores in the control group and the promise group in study 2a and in the commitment group in study 2b. Higher scores indicate greater commitment to the cleaning task. The number of children that reached the respective commitment score is indicated by the area of the bubbles (Range = 0 – 13). The mode is indicated by the white numerals, the median by the solid gray line and the second and third quartile by the dotted gray boxes.
Figure 5. 3- and 5-year-olds’ commitment scores in the control group (cleaning reminder) and the promise group in study 3. Higher scores indicate greater commitment to the cleaning task. The number of children that reached the respective commitment score is indicated by the area of the bubbles (Range = 0 – 14). The mode is indicated by the white numerals, the median by the solid gray line and the second and third quartile by the dotted gray boxes.