nature human behaviour

Article

https://doi.org/10.1038/s41562-024-01837-4

Trusting young children to help causes them to cheat less

Received: 7 May 2023

Accepted: 23 January 2024

Published online: 20 February 2024

Check for updates

Li Zhao [1,2,6] ✉, Haiying Mao [2,3,6], Paul L. Harris [4] & Kang Lee [5]

Trust and honesty are essential for human interactions. Philosophers since antiquity have long posited that they are causally linked. Evidence shows that honesty elicits trust from others, but little is known about the reverse: does trust lead to honesty? Here we experimentally investigated whether trusting young children to help can cause them to become more honest (total $N = 328$ across five studies; 168 boys; mean age, 5.94 years; s.d., 0.28 years). We observed kindergarten children's cheating behaviour after they had been entrusted by an adult to help her with a task. Children who were trusted cheated less than children who were not trusted. Our study provides clear evidence for the causal effect of trust on honesty and contributes to understanding how social factors influence morality. This finding also points to the potential of using adult trust as an effective method to promote honesty in children.

Humans are the only species with large-scale societies comprising mainly biologically unrelated members[1]. To ensure the smooth functioning of such societies, interpersonal trust is of paramount importance[2–5]. Since antiquity, Eastern and Western philosophers such as Confucius, Plato and Aristotle have recognized that of the many conditions for trust, honesty is essential for its establishment and maintenance[6]. Surprisingly, only in recent decades has scientific evidence begun to emerge to confirm this honesty-to-trust linkage[7]. However, little is known about the reverse relation: does trust lead to honesty? The present research aims to answer this question. More specifically, we examined whether kindergarten children would act more honestly and cheat less after being trusted by an adult.

Scientific research over the past three decades has established that the honesty-to-trust causal linkage is evident among not only adults[8,9] but also children[10–15] and even infants[16,17]. For example, after being told about a story character confessing versus another lying about their transgressions, children as young as age three subsequently trusted the information provided by the former but not the latter, and this selective trust became more pronounced with age[10].

However, one cannot simply assume that because honesty leads to trust, trust will automatically elicit honesty. It is possible that trust may lead to dishonesty. This possibility is based on an important characteristic of interpersonal trust, which by definition refers to confidence in others and a willingness to be vulnerable to them[18–23]. When a person entrusts another, the trustor opens up the possibility of being taken advantage of by the trustee. The trustee may use this opportunity to perform socially questionable acts including acting dishonesty towards the trustor (the vulnerability hypothesis). Alternatively, trust might increase people's honesty. According to social exchange theory[24–27], when people perceive themselves as being trusted by others, they may reciprocate that trust by behaving honestly (the reciprocity hypothesis). Direct experimental evidence is needed to assess the circumstances under which these two different reactions emerge in young children.

To test these two possibilities, we focused on children aged five to six years on the basis of a robust tradition of research on the development of trust and honesty, dating back to the 1920s[28–31]. From a young age, children trust others selectively[32–35]. They are inclined to accept claims made by individuals who have proved knowledgeable[36–40], benevolent[15,41,42], reliable[34,40,43–45] and well intentioned[46–49] and to distrust individuals who are ignorant, malicious, unreliable or ill intentioned. Moreover, by five years of age, children already understand the concepts of trust and honesty and the causal link from honesty to trust; they also act according to such understandings[10,11]. For example,

[1]Zhejiang Philosophy and Social Science Laboratory for the Development and Care of Infants and Young Children, Hangzhou, PR China. [2]Department of Psychology, Hangzhou Normal University, Hangzhou, PR China. [3]Department of Psychology, Heidelberg University, Heidelberg, Germany. [4]Graduate School of Education, Harvard University, Cambridge, MA, USA. [5]Dr Eric Jackman Institute of Child Study, University of Toronto, Toronto, Ontario, Canada. [6]These authors contributed equally: Li Zhao, Haiying Mao. ✉e-mail: zhaoli@hznu.edu.cn

https://doi.org/10.1038/s41562-024-01837-4

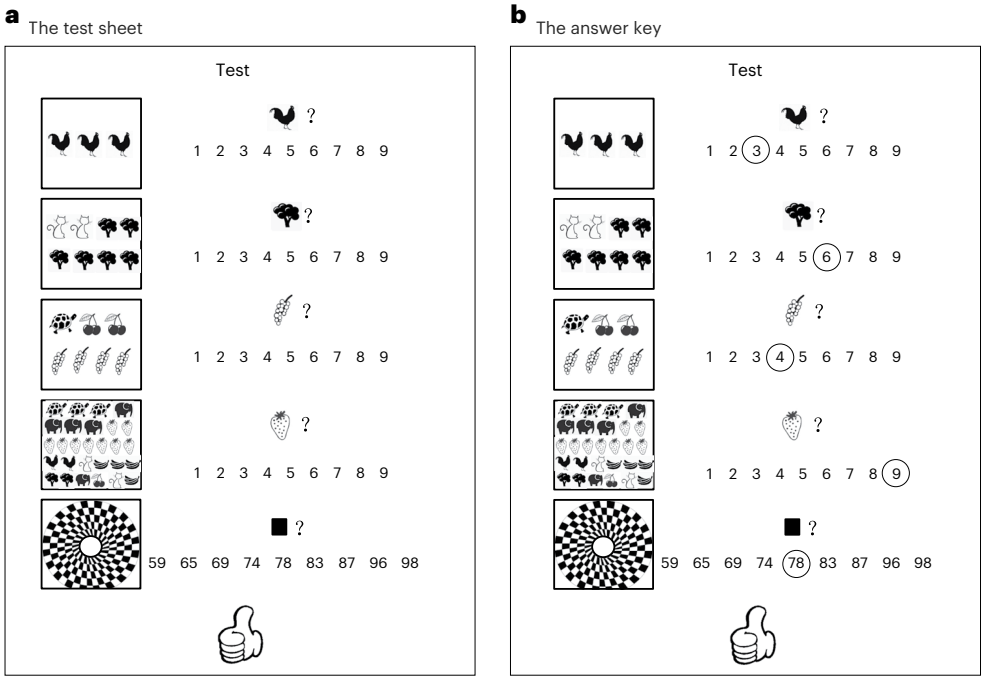

Fig. 1 | Experimental materials. a, The test sheet of the math test that children were asked to complete. **b**, The answer key.

after hearing a story character falsely claiming to be the owner of a toy versus another correctly denying ownership, five-year-olds considered the former to be less trustworthy and did not act on the information provided by this character[11].

Research on honesty has revealed that children are generally honest at age two but begin to cheat and lie shortly thereafter[50,51]. Initially, when enticed by material rewards, nearly all kindergarten children cheat[52–57]. However, when the goal is to attain a better test score, few children under five years cheat, whereas about half of five- to six-year-olds do so[58]. The present research focuses on five- to six-year-olds because their baseline rate of cheating is about 50%. This rate is ideal for testing whether trust enhances cheating (as the vulnerability hypothesis predicts) or reduces it (as the reciprocity hypothesis predicts).

The challenge for testing these two hypotheses experimentally is how best to manipulate trust. Because we aimed to study the trust-to-honesty causal linkage as naturalistically as possible, prior to the behavioural test of honesty, we directly manipulated whether or not children were entrusted to perform a task. Hence, we tapped into an early emerging ability in children: helping. By age two, most children are already highly motivated to help others[59–65]. Accordingly, we reasoned that if we entrusted children by asking them to help, they would do so readily.

To create a situation where helping would appear natural, we capitalized on a common situation in the kindergarten where children are often taken by teachers to be tested elsewhere for various reasons. In the experimental condition, on the way to the testing room, the experimenter asked children to help her by holding an important item (that is, an envelope with the answer key for a test) because her hands were full. Upon arriving at the testing room, the experimenter administered a math test to children, ostensibly to assess their abilities to count but in reality to assess their spontaneous cheating.

The math test was developed by Zhao et al.[58] and derived from the classic cheating tasks invented by Voelker[31] and Hartshorne and May[28]. It consists of five test items, with each item depicting a set of animals, fruits or geometric shapes (Fig. 1a). The children's task is to count the number of exemplars in the set and then circle the correct answer. Of the five items, the first four were easy, as pilot testing showed that they could be answered successfully by children of this age with a passing rate of 100% without cheating. However, the final item was exceptionally difficult, and it was impossible for children of this age to answer it correctly without peeking at the answer key. The answer key was identical to the test sheet except that the correct answer to each test question had been circled (Fig. 1b). Before the children finished the test, the experimenter stepped out of the room, supposedly to deal with an important matter in the next office. She left the answer key on a table near where the children were seated. Unbeknownst to the children, their behaviour during the experimenter's absence was recorded by a hidden video camera. We assessed whether or not the children refrained from peeking at the answers.

We conducted five preregistered studies that systematically tested the two contrasting hypotheses of entrustment (see Table 1 for an overview). In all studies, children were tested individually in a quiet room at their kindergarten. The room contained two identical tables (Fig. 2). Children took the math test at the table on the right which had a digital countdown timer on it. The second table was placed 0.6 m away to the left of children. We chose 0.6 m as the distance between the two tables for two reasons. First, according to prior studies[58,66], when children were sitting down, it was easy for them to see the answers on the table next to them at a distance of 0.6 m, without having to stand up or step towards the answer key to peek. Second, according to Zhao et al.[66], the baseline cheating rate at the inter-table distance of 0.6 m was 54%, which optimally allowed children's cheating rate to increase or decrease depending on our experimental manipulations.

Study 1 investigated whether trusting children to help influenced their cheating behaviour. We randomly assigned children to either an experimental condition or a control condition. Because we were uncertain whether any causal link between trust and honesty existed, we manipulated trust as strongly as possible. Specifically, in the experimental condition (that is, the answer key experimental condition), the experimenter asked the children to help hold the envelope that contained the answer key to the test that they were about to take. To ensure that the children were aware that they were being trusted to do so, the experimenter explicitly mentioned that she trusted them and therefore asked them for help (that is, a message implying that the children were trusted to be helpful). Upon arriving at the testing room,

https://doi.org/10.1038/s41562-024-01837-4

Table 1 | Overview of the experimental design in each condition of all studies in the present research

Study	Condition	Helping behaviour	Trust-related messages	Trustor	Tester
Study 1	Answer key experimental	Help (holding the answer key to a test)	Trusted to help; trusted not to cheat	E1	E1
	Control	No help	No trust-related messages	E1	E1
Study 2	House key experimental	Help (holding house keys)	Trusted to help; trusted not to cheat	E1	E1
	House key control	Help (holding house keys)	No trust-related messages	E1	E1
Study 3	Modified house key experimental (Version 1)	Help (holding house keys)	Trusted to help only	E1	E1
	House key control	Help (holding house keys)	No trust-related messages	E1	E1
Study 4	Modified house key experimental (Version 2)	Help (holding house keys)	Trusted not to cheat only	E1	E1
Study 5	Modified house key experimental (Version 3)	Help (holding house keys)	Trusted to help only	E1	E2

'No help' indicates that the experimenter did not ask the children for help. 'Help (holding the answer key to a test)' indicates that the experimenter asked the children to help hold the answer key to the test that they were about to take. 'Help (holding house keys)' indicates that the experimenter asked the children to help hold her house keys. 'No trust-related messages' indicates that the children were not given trust-related messages. 'Trusted to help' indicates that the experimenter stated she trusted the children to help her before they rendered help, and after they rendered help, she thanked the children for the help and told them that they were being trusted to help in the future. 'Trusted not to cheat' indicates that the experimenter told the children that they were trusted not to cheat in the following math test. 'Trustor' indicates the experimenter who asked the children for help or gave the children trust-related messages. 'Tester' indicates the experimenter who gave the children the math test.

the experimenter thanked the children for taking good care of the envelope and stated that she would similarly trust them in the future (that is, another message implying that the children were trusted to be helpful). Furthermore, she told them that she trusted them not to cheat during her absence (that is, another message implying that she trusted them to act honestly). The children thus received two types of trust messages, one about the help they provided and the other about not cheating. In the control condition, the experimenter did not ask the children for help, and they were not given trust-related information. If the reciprocity hypothesis holds, it is expected that children may cheat less in the experimental condition than in the control condition. Alternatively, if the vulnerability hypothesis is favoured, children may cheat more in the experimental condition than in the control condition.

To assess the influence of helping behaviour in and of itself on children's cheating behaviour, in Study 2, we randomly assigned children to either an experimental condition or a control condition. In both conditions, the experimenter asked the children to help her. However, children in the experimental condition, similar to those in the experimental condition of Study 1, received trust messages, whereas children in the control condition, though being asked to help the experimenter, did not receive any trust messages. If explicit affirmations of trust are crucial, then children should cheat significantly less (as predicted by the reciprocity hypothesis) or more (as predicted by the vulnerability hypothesis) in the experimental condition than in the control condition. However, if the mere act of helping the experimenter is sufficient to affect cheating, then children's cheating should be similar in both conditions.

Fig. 2 | Experimental setup. A schematic depiction of the experimental setup. The answer key was the same size as the child's test sheet, and the inter-table distance was 0.6 m.

We also made another modification in Study 2. In Study 1, the children were asked to help hold the envelope that contained the answer key to the test that they were about to take. They were then asked to refrain from peeking at the very answer key that they had provided help with. This link between the target of help and the target of cheating might be responsible for any possible effects between the conditions of Study 1. To rule out this possibility, all children in Study 2 were asked to help hold the experimenter's house keys.

In the experimental conditions of Studies 1 and 2, children received two types of trust messages: in addition to telling the children that she trusted them to help her, the experimenter stated that she would trust them not to cheat. To examine whether simply telling children that they were trusted to help would be sufficient to engender the trust effect, we conducted Study 3, in which we randomly assigned children to either an experimental condition or a control condition. As in Study 2, in both conditions of Study 3, the experimenter asked the children to help hold her house keys. In the modified house key experimental condition (Version 1), the experimenter told the children that they were trusted to help but did not mention that they were trusted not to cheat. The house key control condition was the same as that in Study 2. The children were asked to help but received no trust-related messages. If telling children that they were trusted to help was sufficient to make them act more honestly (as predicted by the reciprocity hypothesis) or more dishonestly (as predicted by the vulnerability hypothesis), the cheating rate in the experimental condition should be significantly different from that in the control condition, in which no trust messages were given.

We then conducted Study 4 to assess whether a message to children about being trusted not to cheat alone would be sufficient to cause them to cheat differently. In this study, there was a single modified house key experimental condition (Version 2) in which the experimenter asked the children to hold her house keys but did not mention their being trusted to help. Instead, she stated that she trusted them not to cheat. If this statement was sufficient to affect children's cheating, children should cheat significantly differently than those in the control conditions of Studies 2 and 3 (that is, the house key control conditions).

Study 5 also included a single modified house key experimental condition (Version 3) that was identical to the experimental condition

https://doi.org/10.1038/s41562-024-01837-4

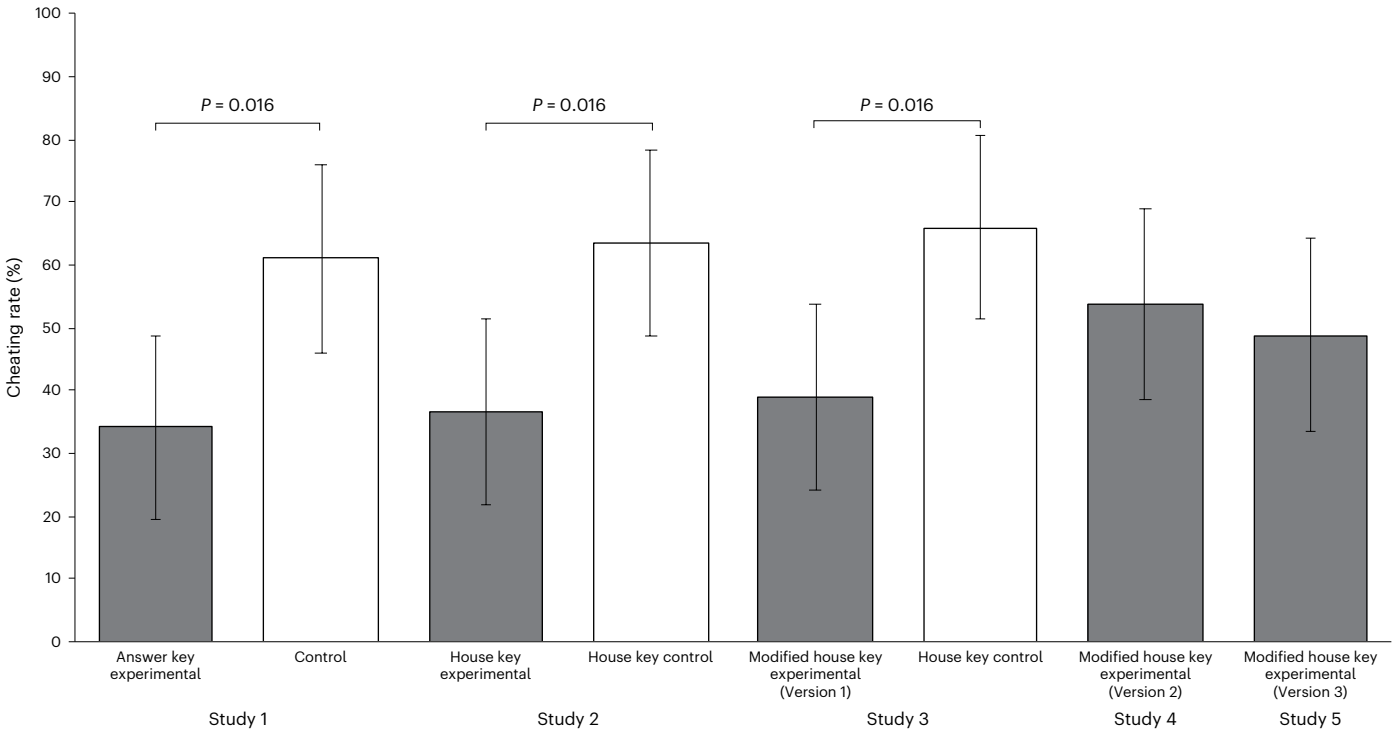

Fig. 3 | Cheating rates across conditions and studies. Percentage of children who cheated in each condition across the five studies (total $N = 328$, $n = 41$ for each condition of each study). The data are presented as the number of children who cheated divided by the total number ($n = 41$) of children in each condition, with error bars representing 95% CIs. Binary logistic regression analyses showed that in Study 1, the cheating rate in the answer key experimental condition was significantly lower than that in the control condition (Wald = 5.77; d.f. = 1; $P = 0.016$; OR = 0.33; 95% CI, 0.13 to 0.80; Cohen's $d = -0.61$). In Study 2, the cheating rate in the house key experimental condition was significantly lower than that in the house key control condition (Wald = 5.76; d.f. = 1; $P = 0.016$; OR = 0.33; 95% CI, 0.13 to 0.81; Cohen's $d = -0.61$). In Study 3, the cheating rate in the modified house key experimental condition (Version 1) was significantly lower than that in the house key control condition (Wald = 5.77; d.f. = 1; $P = 0.016$; OR = 0.33; 95% CI, 0.13 to 0.80; Cohen's $d = -0.61$). In Study 4, the cheating rate in the modified house key experimental condition (Version 2) did not differ significantly from the control conditions in Studies 1, 2 and 3 (Wald = 1.43, d.f. = 3, $P = 0.699$). In Study 5, the cheating rate in the modified house key experimental condition (Version 3) did not differ significantly from the control conditions in Studies 1, 2 and 3 (Wald = 2.90, d.f. = 3, $P = 0.407$). All statistical tests were two-tailed.

of Study 3 with one exception. This study involved two experimenters, with Experimenter 1 (E1, the trustor) asking children for help and Experimenter 2 (E2, the tester) giving children the math test. This study was to test whether the reciprocity hypothesis, if true, would hold even when the trustor was absent. One possibility is that the trust effect is specific to the trustor who has expressed trust: children would behave more honestly to reciprocate the trust expressed by the trustor on the basis of the principle of mutual benefit. Alternatively, after hearing trust expressed by one adult, children may have a generalized idea that any other adult would trust them. They may reciprocate that assumed trust with increased honesty regardless of whether the adult in question has expressed trust in them.

Results

Figure 3 shows the cheating rate for each condition across the five studies.

Study 1

Figure 3 shows that the cheating rate was only 34.1% in the answer key experimental condition. In the control condition, it was 61%.

We conducted a planned binary logistic regression analysis with cheating behaviour (0, no cheating; 1, cheating) as the predicted variable and condition (0, control condition; 1, answer key experimental condition) as the predictor. All statistical analyses reported below are two-tailed.

The model was significant ($\chi^2_{1,N=82} = 5.99$, $P = 0.014$, Nagelkerke $R^2 = 0.09$). The condition effect was significant, showing that the cheating rate in the answer key experimental condition was significantly lower than that in the control condition ($\beta = -1.10$; s.e. $\beta = 0.46$; Wald = 5.77; d.f. = 1; $P = 0.016$; odds ratio (OR) = 0.33; 95% confidence interval (CI), 0.13 to 0.80; Cohen's $d = -0.61$). This result demonstrated that trusting children to perform an act of helping significantly reduced their tendency to cheat.

However, these findings should not necessarily be taken as evidence to support the reciprocity hypothesis but not the vulnerability hypothesis, because the condition effect might not be due to the fact that the children heard the experimenter's expressed trust in them. Rather, they may have become less likely to cheat because they helped the experimenter. It is possible that after performing one kind of prosocial behaviour (helping), the children became more inclined to perform another prosocial behaviour[67,68]. To test whether the findings of Study 1 were specifically due to the expressed trust by an adult or to a generalization of prosocial behaviour, we conducted Study 2.

Study 2

In Study 2, both the house key experimental and house key control conditions involved children helping the experimenter, but the children received trust-related messages from the experimenter only in the house key experimental condition. Figure 3 shows that the cheating rate was higher in the house key control condition (63.4%) than in the house key experimental condition (36.6%).

We conducted a planned binary logistic regression analysis on cheating behaviour (0, no cheating; 1, cheating), with condition (0, house key control condition; 1, house key experimental condition) as the predictor. The model was significant ($\chi^2_{1,N=82} = 5.98$, $P = 0.015$, Nagelkerke $R^2 = 0.09$). The condition effect was significant, showing that the

cheating rate in the house key experimental condition was significantly lower than that in the house key control condition ($\beta = -1.10$; s.e. $\beta = 0.46$; Wald = 5.76; d.f. = 1; $P = 0.016$; OR = 0.33; 95% CI, 0.13 to 0.81; Cohen's $d = -0.61$). Thus, the experimenter's expressed trust significantly reduced the children's tendency to cheat.

In Study 2, we made two important changes: (1) we changed the target of helping from the answer key to the house keys, and (2) the experimenter asked children for help in both the experimental and control conditions. Despite these changes, we replicated the trust effect observed in Study 1, supporting the reciprocity hypothesis. Furthermore, the results suggest that the target of helping (that is, the house keys) and the target of cheating (that is, the answer key) do not need to be the same for the trust effect to be observed. More importantly, these results indicate that the experimenter's invitation to help—which could potentially be interpreted by children as an act of trust on the part of the experimenter—did not influence children's cheating significantly. Rather, explicit affirmations of trust—not tacit implications of trust—reduced cheating.

In Studies 1 and 2, the experimenter expressed trust in children to help and not to cheat simultaneously, leaving the question of whether expressing trust in children to help alone was sufficient to drive the trust effect. Study 3 addressed this question by removing the statement about children being trusted not to cheat, as used in the house key experimental condition of Study 2.

Study 3

Study 3 had two conditions. In the modified house key experimental condition (Version 1), children were first asked to help the experimenter to hold her house keys. After the children rendered help, the experimenter told them that they were trusted to help. However, she did not mention that they were trusted not to cheat. In the house key control condition (identical to the house key control condition of Study 2), after the children rendered help, the experimenter did not tell them that they were trusted to help or that they were trusted not to cheat.

As shown in Fig. 3, 39% of children cheated in the modified house key experimental condition (Version 1), whereas 65.9% cheated in the house key control condition.

We conducted a planned binary logistic regression analysis on cheating behaviour (0, no cheating; 1, cheating), with condition (0, house key control condition; 1, modified house key experimental condition (Version 1)) as the only predictor. The model was significant ($\chi^2_{1,N=82} = 5.99$, $P = 0.014$, Nagelkerke $R^2 = 0.09$). Inspection of the model revealed that the cheating rate in the modified house key experimental condition (Version 1) was significantly lower than that in the house key control condition ($\beta = -1.10$; s.e. $\beta = 0.46$; Wald = 5.77; d.f. = 1; $P = 0.016$; OR = 0.33; 95% CI, 0.13 to 0.80; Cohen's $d = -0.61$).

This result replicated the trust effect found in Studies 1 and 2 and demonstrated that this effect was due to the fact that children were given the message about being trusted to help but did not depend on the additional message about being trusted not to cheat.

Study 4

Study 4 tested whether telling children that they were trusted not to cheat alone was sufficient to reduce cheating. In the modified house key experimental condition (Version 2), the statement about the experimenter trusting the children to help was removed, but the statement about her trusting them not to cheat was reintroduced.

The cheating rate in this new condition was 53.7% (Fig. 3). We conducted a planned binary logistic regression analysis to compare the cheating rate in the current modified house key experimental condition (Version 2) to those in the control conditions in Studies 1, 2 and 3. This model was not significant ($\chi^2_{3,N=164} = 1.43$; $P = 0.699$; Nagelkerke $R^2 = 0.01$; Bayes factor, 2.52). Inspection of the model revealed that the condition effect was not significant (Wald = 1.43, d.f. = 3, $P = 0.699$). The cheating rate in this condition was not significantly different from those in the control conditions of Studies 1, 2 and 3 (61%, 63.4% and 65.9%, respectively); neither did the rates in the control conditions significantly differ from each other ($\beta = 0.10$, 0.21 and 0.11; s.e. $\beta = 0.46$, 0.46 and 0.46; Wald = 0.05, 0.21 and 0.05; d.f. = 1, 1 and 1; $P = 0.820$, 0.647 and 0.817; OR = 1.11, 1.23 and 1.11; 95% CI, 0.45 to 2.73, 0.50 to 3.06 and 0.45 to 2.77; for the control condition of Study 1 versus the control condition of Study 2, the control condition of Study 1 versus the control condition of Study 3 and the control condition of Study 2 versus the control condition of Study 3, respectively). Telling the children that they were trusted not to cheat thus did not significantly reduce their cheating. However, this conclusion should be interpreted with caution: although a Bayes factor of 2.52 suggests evidence in favour of the null hypothesis (no differences) because it is greater than 1, the fact that it is less than 3 suggests that the support for this null conclusion is not strong.

Study 5

Study 5 tested whether the effect of trust messaging on cheating was specific to a trustor who directly trusted the children or transferable to a non-trustor. We conducted a modified house key experimental condition (Version 3) that was identical to the experimental condition of Study 3 with one exception. In this condition, one experimenter (E1) asked the children for help, and another (E2) gave them the test.

The cheating rate in this experimental condition was 48.8% (Fig. 3). We conducted a similar planned binary logistic regression analysis as that in Study 4 to compare the cheating rate in the current modified house key experimental condition (Version 3) to those in the three control conditions in Studies 1, 2 and 3. This model was not significant either ($\chi^2_{3,N=164} = 2.91$; $P = 0.405$; Nagelkerke $R^2 = 0.02$; Bayes factor, 1.27). Inspection of the model showed the condition effect to be not significant (Wald = 2.90, d.f. = 3, $P = 0.407$). It thus appeared that there was no significant decrease in the cheating rate relative to that in the control conditions when children interacted with an adult who either did not ask them to help or asked them for help but did not provide explicit trust-related messages about helping. In other words, trust messaging reduces cheating only when children are facing the trustor who has trusted them, presumably to reciprocate her trust by acting honestly. However, this conclusion should be interpreted with caution: although a Bayes factor of 1.27 suggests evidence in favour of the null hypothesis (no differences), the fact that it is less than 3 suggests that the support for this null conclusion is not strong.

We also performed non-preregistered analyses on the temporal measures of children's behaviour during the experiment and on the gender and age (in months) effects for all five studies. We found these measures and the gender and age (in months) effects to be generally not significant across conditions (Supplementary Tables 1–3 and Supplementary Figs. 1–3).

Discussion

Empirical evidence in recent decades has supported the honesty-to-trust linkage. The present research addressed the reverse causal link: does trust lead to honesty? With five preregistered studies using an experimental design, we systematically investigated whether trusting children would cause them to be more or less honest. We specifically tested two contrasting hypotheses. The reciprocity hypothesis predicts that trust promotes honesty, whereas the vulnerability hypothesis predicts that trust leads to increased dishonesty.

In Study 1, we randomly assigned five- to six-year-olds to either an experimental condition in which they were trusted by an experimenter to help her or a control condition in which they were not. The children were then given a math test in which they could cheat to answer all questions correctly. Children cheated significantly less in the experimental condition than in the control condition. Study 2 showed that this trust effect was not caused by the mere act of children's helping the experimenter or being asked to help. Rather, it was caused by the experimenter's expression of trust, which included her telling the

children that she trusted them to help her and that she would trust them not to cheat. This study also revealed that children do not treat a request for help as an implicit and non-verbal affirmation of trust by the adult; instead, explicit affirmations of trust are needed to engender the trust effect. Study 3 removed the message about trusting children not to cheat and found that as long as children received the trust-related messages about helping, the trust effect remained. Study 4 further showed that the message about trusting children not to cheat was ineffective in nudging children away from cheating. Study 5 revealed that the trust effect was specific to the experimenter who trusted the children and did not generalize to a non-trustor.

Taken together, these findings support the reciprocity hypothesis. They show that children around five to six years already understand and practise a strategy of reciprocity, one of the fundamental social exchange strategies for maintaining positive interpersonal relationships[69]. More specifically, when children are trusted to help and receive explicit affirmations of trust from another person, they reciprocate by acting honestly towards that person. Our findings do not support the contrasting vulnerability hypothesis. Although the experimenter displayed vulnerability by trusting the children, they did not take advantage of her vulnerability by cheating on the test. Trust thus appears to encourage children to be more honest, not less.

This positive effect of trust towards children is important because it implies that children's honesty can be enhanced by adults' expressed trust rather than by threats or punishment. Previous research has shown that negative social messages such as censure, punishment and intimidating threats can have negative effects on children's prosocial behaviour[70–74]. Our results extend this finding by showing that positive messages—notably, expressed trust—can have positive effects on children's honesty.

These findings also fail to support the moral licensing hypothesis[75,76], which predicts that trust leads to dishonesty, just like the vulnerability hypothesis. The moral licensing hypothesis posits that after having performed certain desirable deeds, people will become more inclined to engage in morally questionable acts because doing good deeds has satisfied their self-image of being a moral person and provides a 'licence' that allows for morally questionable behaviour later on. According to this idea, after people have been trusted by another person to perform a task, they may increase their propensity to engage in dishonest behaviours because doing so will not damage their moral self-image.

The present research was not designed specifically to test this hypothesis, because previous studies have shown that young children do not show a moral self-licensing effect due to their immaturity in moral identity development[77,78]. Nevertheless, by virtue of supporting the reciprocity effect, we demonstrated that the moral licensing hypothesis may not apply to five- to six-year-olds. In other words, five- to six-year-olds may not be as cynical as older children and adults, who appear to be more likely to take advantage of a more positive social situation as an opportunity for doing something immoral[79].

Another notable finding of the current research is that the trust effect observed in our studies was selectively confined to the trustor who provided trust messaging, and it did not generalize to the non-trustor. This finding about the boundary condition of the trust effect is consistent with the extensive literature regarding young children's selective trust. It is well established that children before age six do not blindly trust any adults[80–85]. They are already highly capable of deciding whom to trust depending on an individual's ability, reliability, intentions, motivations and past history of social interactions[32,44,86]. It is thus highly adaptive for children in Study 5 not to transfer what they had learned from the trustor to the non-trustor, who was a complete stranger to them with no prior history of social interaction. However, had this non-trustor been a familiar teacher with whom the children had had positive social interactions, the trust effect might have been transferred to them.

It is also worth noting that the current paradigm makes the act of cheating far easier and more low-cost than in real-world settings. For example, in the current situation, children could effortlessly access the answer key without running the risk of detection, thereby creating a tempting condition for cheating. In contrast, in real-world contexts, parents and teachers often implement preventive measures to ensure that cheating is difficult. However, even under the conditions of the present research, the trust effect remained effective in reducing cheating. This finding not only suggests the robustness of the trust effect but also invites inquiries into how trust influences honesty in more naturalistic situations.

Our findings regarding the causal link from trust to honesty dovetail well with the existing literature regarding the causal link from honesty to trust. The existing developmental research shows that the understanding of the honesty-to-trust linkage begins in infancy[16,17] and develops gradually throughout childhood[10,11,13–15,87–90]. Especially relevant to the present research are findings that although children younger than five have some rudimentary understanding that honesty is linked to trust, only five-year-olds are able to make clear and specific causal connections between honesty and trust and between dishonesty and distrust[11]. These and the current findings taken together suggest that children may come to understand the bidirectional causal links between trust and honesty around the same age. More broadly, they provide direct, empirical evidence to confirm a long-held belief: honesty leads to trust and vice versa.

The present research has several limitations. First, we did not measure children's individual characteristics, such as their sociocognitive abilities (for example, executive function and theory of mind[91]), their family dynamics or their basic trust in adults. Future research on these characteristics could provide important information about the factors that affect the links between trust and honesty. Second, this research only focused on the causal impact of trust on honesty. Because trust plays a central role in many aspects of human interactions, future research should examine how adult trust can shape the development of other prosocial behaviours, such as fairness and kindness[92–94]. Third, it would be informative to include a control condition in Studies 4 and 5 to enhance the reproducibility of the observed effects.

These limitations notwithstanding, the present research provides clear evidence that trust indeed causes people to be honest. Furthermore, it shows that this causal linkage already exists in early childhood. Our findings thus complement the finding that the honesty-to-trust link also emerges around age five to six. These findings together suggest that the causal links between trust and honesty are an early developmental milestone, perhaps due to their adaptive importance for developing cooperative relations with others who are not necessarily biological relatives[95,96]. Practically, our findings suggest that adults' trust in children to help, along with their expressions of trust, can serve as an effective method to reduce cheating in children, which provides an additional tool in the existing toolbox that parents and teachers can use to foster the development of honesty[50].

In summary, the present research with five preregistered studies tested the causal link from trust to honesty proposed by philosophers for centuries. Specifically, it examined whether children would become more honest and cheat less after they had been trusted by an adult to help. Our results support the reciprocity hypothesis, whereby being trusted by a person leads five- to six-year-olds to behave more honestly towards that specific person. These findings advance our theoretical understanding of the close causal links between trust and honesty. Practically, the present findings point to the potential of using adult trust as an effective method for promoting honesty in children.

Methods

This research was approved by the Scientific Research Ethics Committee of Hangzhou Normal University. Parents or legal guardians gave informed consent to allow children to participate, and the children also gave oral assent prior to participating in the studies.

Participants

We predetermined a sample size of 41 participants for each condition by conducting a priori power analyses based on existing research using similar paradigms[58] as well as our pilot testing. Specifically, we estimated a baseline cheating rate for the control condition of 55% and an average cheating rate for the experimental condition of 25%. The power analyses revealed that to achieve a significant condition effect with a power of 0.8, an α value of 0.05 and an enrolment ratio of 1, the appropriate sample size was 41 for each condition. Each of the studies was preregistered (https://aspredicted.org/k2q8h.pdf for Study 1 on 16 October 2021, https://aspredicted.org/hy8tp.pdf for Study 2 on 6 November 2021, https://aspredicted.org/ga2xb.pdf for Study 3 on 8 February 2023, https://aspredicted.org/v3kt2.pdf for Study 4 on 2 March 2023 and https://aspredicted.org/ig9cw.pdf for Study 5 on 21 February 2023). Data collection was conducted in sequence from Studies 1 through 5 over a span of one year and five months from 18 October 2021.

A total of 328 five- to six-year-old children were recruited from a kindergarten located in a metropolitan city in China (mean age, 71.30 months; s.d., 3.31 months; range, 63.35 to 77.88 months; 168 boys). Children participating in each of the five studies were randomly and blindly assigned to each condition in each study. There were 41 children in each of the eight conditions: the answer key experimental condition (mean age, 71.04 months; s.d., 4.17 months; range, 63.62 to 75.88 months; 22 boys) and the control condition (mean age, 70.73 months; s.d., 3.92 months; range, 63.35 to 75.85 months; 21 boys) in Study 1; the house key experimental condition (mean age, 70.42 months; s.d., 4.33 months; range, 63.62 to 75.55 months; 20 boys) and the house key control condition (mean age, 72.60 months; s.d., 2.07 months; range, 69.07 to 75.75 months; 21 boys) in Study 2; the modified house key experimental condition (Version 1) (mean age, 72.50 months; s.d., 2.35 months; range, 68.48 to 76.11 months; 21 boys) and the house key control condition (mean age, 69.88 months; s.d., 1.92 months; range, 66.94 to 74.43 months; 21 boys) in Study 3; the modified house key experimental condition (Version 2) (mean age, 72.54 months; s.d., 3.41 months; range, 68.05 to 77.88 months; 21 boys) in Study 4; and the modified house key experimental condition (Version 3) (mean age, 70.65 months; s.d., 2.10 months; range, 67.33 to 74.30 months; 21 boys) in Study 5.

According to the kindergarten records, all children were Han Chinese and from middle socio-economic status backgrounds. All children passed the comprehension checks (see below), and thus none was excluded. The participants received no compensation in any of the studies.

Study 1

A female experimenter brought children from their classroom to the testing room individually. The children were randomly assigned to either the answer key experimental condition or the control condition. The procedure and materials for these two conditions were identical except for the trust messaging manipulations.

Answer key experimental condition. On the way to the testing room, the experimenter asked the children to help hold a transparent plastic envelope containing a folded answer key to the math test that they were going to take and said the following: "Can you hold the envelope for me? I have all this stuff in my hands and my hands are full. This envelope is important because it has the answer key to the math test you are about to take and after you complete the test I need to use it to check whether you get all the answers right. Please take special care of it, and make sure you don't drop it on the ground. I trust that you will take good care of the answer key." The experimenter then handed over the envelope to the children. After arriving at the testing room, the experimenter said, "Thank you very much! You take very good care of things. Next time I will trust and ask you again if I need someone to look after something important."

Next, the experimenter told the children that they would be taking a math test that was designed to assess whether they were good at counting. Prior to the formal test, the experimenter first asked them to count from 1 to 20 to make sure that they had the basic numerical understanding to solve the problems. All children passed this pre-test. After that, the experimenter gave the children three practice problems, instructing them to count the shapes specified by the question to familiarize them with the test. The practice problems were also used to verify that all children had the necessary ability to count to complete the first four simple problems of the test. No child failed the practice problems.

The experimenter then presented the children with the test sheet of the math test, which included five problems, and told them that they should try to answer them all correctly in five minutes. She explained, "Now it's time for you to take the test. You will have up to five minutes to finish it. Here is a clock [indicating the digital countdown clock] that will remind you how much time is left for you to work on the test. It will sound an alarm when the time runs out. Remember that if you don't finish on time, no matter how many problems you got correct, it counts as zero." After instructing the children to complete the math test, the experimenter made an excuse to leave the room: "Sorry, I just remembered that I need to go to the neighboring classroom to deal with an emergency. I will not be able to come back in five minutes. While I am away, you should try to solve the problems by yourself. When you are done, you should leave your test sheet on your table and find me in the next room. Make sure that you finish the test before the time runs out. Please come to the next room to get me when five minutes are up." We asked the children to find the experimenter in the next room instead of her returning to the testing room to create a situation in which they could cheat on the test without having to worry about being caught by the experimenter.

The experimenter then asked the following questions as comprehension checks: (1) How much time do you have to finish the test? (2) When you are done, what should you do? (3) What will happen if you don't finish the test within five minutes? All children answered these questions correctly.

After completing the checks, the experimenter said, "Let me leave the answer key here [placing it in the middle of the nearby table; Fig. 2]. I will score your test according to it, and see if you answer all of the questions correctly. Remember: Don't peek at the answer key when I am away. I trust you not to peek at it." Then the experimenter started the timer and exited the room.

After the allotted time was up, the experimenter returned to the room after being called back by the children as they had been instructed to do. If the children did not come to retrieve the experimenter when the alarm went off, the experimenter returned to the room anyway. The children were then debriefed and sent back to their classrooms.

Control condition. The control condition was identical to the experimental condition except that on the way to the testing room, the experimenter did not ask the children to hold anything, and therefore there was no mention of trust. Before the experimenter left the room, she also said, "Remember: Don't peek at the answer key when I am away."

Study 2

Children were tested individually by the same female experimenter as in Study 1. The children were randomly assigned to either a house key experimental condition or a house key control condition. The procedure and materials for these two conditions were identical except for the trust messaging manipulations.

House key experimental condition. The house key experimental condition was identical to the answer key experimental condition in Study 1, except that the experimenter asked the children to help hold her house keys and said the following: "Can you hold my house keys for me? I have all this stuff in my hands and my hands are full. This set of house keys is important to me because they are the only keys that could open the door

of my house; otherwise, I can't get in the house when I get home. Please take special care of them and make sure you don't drop them on the ground. I trust that you will take good care of my house keys." The experimenter then handed over the house keys to the children. As in Study 1, after arriving at the testing room, the experimenter said, "Thank you very much! You take very good care of things. Next time I will trust and ask you again if I need someone to look after something important." The experimenter then gave the children the math test in the same way as in Study 1. As in Study 1, before leaving, the experimenter said, "Remember: Don't peek at the answer key when I am away. I trust you not to peek at it."

House key control condition. The house key control condition was identical to the experimental one except that (1) when giving the house keys to the children to hold, she did not say, "I trust that you will take good care of my house keys", and (2) after the children helped hold the house keys, the experimenter said, "Thank you very much! You take very good care of things", without saying, "Next time I will trust and ask you again if I need someone to look after something important." Also, before she left the room, she said, "Remember: Don't peek at the answer key when I am away," without saying, "I trust you not to peek at it."

Study 3

Children were tested individually by the same female experimenter as in the previous studies. They were randomly assigned to either the modified house key experimental condition (Version 1) or the house key control condition. The procedure and materials for these two conditions were identical except for the trust manipulations.

The modified house key experimental condition (Version 1) was nearly identical to the house key experimental condition in Study 2, in that the experimenter told the children that they were trusted to help look after the house key. However, unlike in the house key experimental condition of Study 2, before leaving, the experimenter instructed the children not to peek at the answer key but did not add the statement about their being trusted not to cheat. The house key control condition was identical to the house key control condition in Study 2.

Study 4

Children were tested individually by the same female experimenter as in the previous studies, and they were assigned to the single modified house key experimental condition (Version 2). This condition was nearly identical to the house key experimental condition in Study 2, in that the experimenter told the children that they were trusted not to cheat after they helped hold her house keys. However, unlike in the house key experimental condition of Study 2, the experimenter did not mention that the children were trusted to help. That is, the experimenter asked the children to help hold her house keys and gave the same instructions before handing over the house keys to them in the same way. However, after arriving at the testing room, the experimenter only said, "Thank you very much! You take very good care of things." The experimenter then gave the children the math test. Later, before leaving the room, as in the house key experimental condition of Study 2, the experimenter still said, "Remember: Don't peek at the answer key when I am away. I trust you not to peek at it."

Study 5

Children were assigned to the single modified house key experimental condition (Version 3). In this condition, we used the same design and procedure as the modified house key experimental condition (Version 1) in Study 3, except for the involvement of two experimenters, E1 and E2, both of whom were female graduate students.

Specifically, E1 (the trustor) began the session by asking the children to help hold her house key on the way to the testing room. After arriving outside of the testing room, E1 thanked the children for doing a good job of taking care of her house key and stated, "Next time I will trust and ask you again if I need someone to look after something important."

Then E1 said, "There is a teacher from another kindergarten whom I don't know, and she wants to give a test to you now." After that, E1 knocked on the door and had a brief conversation with E2 (the tester). E1 and E2 behaved as if they did not know each other in the presence of the children. E2 introduced herself to the children as a visitor from a distant city, who would go back to her city soon and wanted to test the children to find out whether they were good at counting. Then E1 left, and E2 presented the children with the math test by giving the same instructions as in Study 3. Specifically, before leaving, E2 said, "Remember: Don't peek at the answer key when I am away."

Coding of the dependent variable

In all five studies, the key dependent measure was cheating behaviour, defined as the child peeking at the answer key and copying an answer from it in the absence of the experimenter. Two graduate students who were naive to the study hypotheses independently coded the children's cheating behaviour on the basis of the recordings taken by the hidden video camera. All instances of cheating were further confirmed by ensuring that the children's answer to the fifth and crucial problem matched the corresponding answer on the answer key. The inter-coder agreement was 100%.

In addition to cheating behaviour, a preregistered dependent variable, we performed post hoc coding of the temporal measures of children's behaviour during the experiment and reported the non-preregistered data analyses in the Supplementary Information. Specifically, we included cheating latencies for children who cheated by counting the number of seconds from when the experimenter left the room to the occurrence of the children's cheating behaviour. We also included waiting time for all children: for children who cheated, their waiting time was identical to their cheating latency (that is, the number of seconds from when the experimenter left the room to the occurrence of the cheating behaviour); for children who did not cheat, their waiting time was counted as 300 seconds. In addition, we calculated leaving time for children who finished the math test and retrieved the experimenter by counting the number of seconds from the time when the children finished the last problem of the math test to the time when they went to the nearby room to retrieve the experimenter. We did so to compare the time taken by children who cheated and that taken by children who did not cheat to call the experimenter after they finished the math test. The above-mentioned two graduate students independently coded the children's behaviour on these temporal measures on the basis of the video recordings. The time was classified as the same if the two coders reported numbers of seconds that differed by no more than one second. There was 100% inter-coder agreement for all these measures.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data that support the findings of this study are available at https://osf.io/yqdhp/.

Code availability

We used R software (version 4.3.1) to perform the binary logistic analyses. We estimated Bayes factors using the Bayesian loglinear regression models with mixture Dirichlet priors and performed generalized estimating equations using SPSS (version 26). The code for all analyses that support the findings of this study is available at https://osf.io/yqdhp/.

References

1. Fehr, E. & Gächter, S. Altruistic punishment in humans. *Nature* **415**, 137–140 (2002).
2. Rotenberg, K. J. *The Psychology of Interpersonal Trust: Theory and Research* (Routledge, 2020).

3. Rotter, J. B. Generalized expectancies for interpersonal trust. *Am. Psychol.* **26**, 443–452 (1971).
4. Rotter, J. B. Interpersonal trust, trustworthiness, and gullibility. *Am. Psychol.* **35**, 1–7 (1980).
5. Van Lange, P. A., Joireman, J., Parks, C. D. & Van Dijk, E. The psychology of social dilemmas: a review. *Organ. Behav. Hum. Decis. Process.* **120**, 125–141 (2013).
6. McLeod, C. in *The Stanford Encyclopedia of Philosophy* Fall 2021 Edn (ed. Zalta, E. N.) (Stanford Univ., 2021); https://plato.stanford.edu/archives/fall2021/entries/trust/
7. Connelly, B. L., Crook, T. R., Combs, J. G., Ketchen, D. J. Jr & Aguinis, H. Competence- and integrity-based trust in interorganizational relationships: which matters more? *J. Manage.* **44**, 919–945 (2018).
8. Levine, E. E. & Schweitzer, M. E. Prosocial lies: when deception breeds trust. *Organ. Behav. Hum. Decis. Process.* **126**, 88–106 (2015).
9. Schweitzer, M. E., Hershey, J. C. & Bradlow, E. T. Promises and lies: restoring violated trust. *Organ. Behav. Hum. Decis. Process.* **101**, 1–19 (2006).
10. Lane, J. D., Wellman, H. M. & Gelman, S. A. Informants' traits weigh heavily in young children's trust in testimony and in their epistemic inferences. *Child Dev.* **84**, 1253–1268 (2013).
11. Li, Q. G., Heyman, G. D., Xu, F. & Lee, K. Young children's use of honesty as a basis for selective trust. *J. Exp. Child Psychol.* **117**, 59–72 (2014).
12. Ma, F. et al. Modesty can promote trust: evidence from China. *Soc. Dev.* **28**, 218–233 (2019).
13. Rotenberg, K. J. (ed.) *Children's Interpersonal Trust: Sensitivity to Lying, Deception and Promise Violations* (Springer, 1991).
14. Sher, I., Koenig, M. & Rustichini, A. Children's strategic theory of mind. *Proc. Natl Acad. Sci. USA* **111**, 13307–13312 (2014).
15. Xu, F. et al. The role of honesty and benevolence in children's judgments of trustworthiness. *Int. J. Behav. Dev.* **37**, 257–265 (2013).
16. Poulin-Dubois, D. & Chow, V. The effect of a looker's past reliability on infants' reasoning about beliefs. *Dev. Psychol.* **45**, 1576–1582 (2009).
17. Poulin-Dubois, D., Brooker, I. & Polonia, A. Infants prefer to imitate a reliable person. *Infant Behav. Dev.* **34**, 303–309 (2011).
18. Lount, R. B. Jr & Pettit, N. C. The social context of trust: the role of status. *Organ. Behav. Hum. Decis. Process.* **117**, 15–23 (2012).
19. Mayer, R. C., Davis, J. H. & Schoorman, F. D. An integrative model of organizational trust. *Acad. Manage. Rev.* **20**, 709–734 (1995).
20. Misztal, B. A. Trust: acceptance of, precaution against and cause of vulnerability. *Comp. Sociol.* **10**, 358–379 (2011).
21. Rousseau, D. M., Sitkin, S. B., Burt, R. S. & Camerer, C. Not so different after all: a cross-discipline view of trust. *Acad. Manage. Rev.* **23**, 393–404 (1998).
22. Siegrist, M. Trust and risk perception: a critical review of the literature. *Risk Anal.* **41**, 480–490 (2021).
23. Whitener, E. M., Brodt, S. E., Korsgaard, M. A. & Werner, J. M. Managers as initiators of trust: an exchange relationship framework for understanding managerial trustworthy behavior. *Acad. Manage. Rev.* **23**, 513–530 (1998).
24. Blau, P. M. *Exchange and Power in Social Life* (Wiley, 1964).
25. Cook, K. S., Cheshire, C., Rice, E. R. & Nakagawa, S. in *Handbook of Social Psychology* (eds DeLamater, J. & Ward, A.) 61–88 (Springer Dordrecht, 2013); https://doi.org/10.1007/978-94-007-6772-0_3
26. Cropanzano, R., Anthony, E. L., Daniels, S. R. & Hall, A. V. Social exchange theory: a critical review with theoretical remedies. *Acad. Manage. Ann.* **11**, 479–516 (2017).
27. Homans, G. C. *Social Behavior and Its Elementary Forms* (Harcourt, Brace and World, 1961).
28. Hartshorne, H. & May, M. S. *Studies in the Nature of Character: Studies in Deceit* (Macmillan, 1928).
29. Meyer, A. E. The lies that children tell. *Sci. Mon.* **23**, 519–528 (1926).
30. Piaget, J. *The Moral Judgment of the Child* (Harcourt Brace, 1932).
31. Voelker, P. F. *The Function of Ideals and Attitudes in Social Education* (Columbia Univ. Press, 1921).
32. Harris, P. L., Koenig, M. A., Corriveau, K. H. & Jaswal, V. K. Cognitive foundations of learning from testimony. *Annu. Rev. Psychol.* **69**, 251–273 (2018).
33. Harris, P. L. & Corriveau, K. H. Young children's selective trust in informants. *Phil. Trans. R. Soc. B* **366**, 1179–1187 (2011).
34. Koenig, M. A. & Harris, P. L. The basis of epistemic trust: reliable testimony or reliable sources? *Episteme* **4**, 264–284 (2008).
35. Tong, Y., Wang, F. & Danovitch, J. The role of epistemic and social characteristics in children's selective trust: three meta-analyses. *Dev. Sci.* **23**, e12895 (2020).
36. Corriveau, K. H. & Harris, P. L. Preschoolers continue to trust a more accurate informant 1 week after exposure to accuracy information. *Dev. Sci.* **12**, 188–193 (2009).
37. Corriveau, K. H. & Kurkul, K. E. 'Why does rain fall?': children prefer to learn from an informant who uses noncircular explanations. *Child Dev.* **85**, 1827–1835 (2014).
38. Corriveau, K. H., Meints, K. & Harris, P. L. Early tracking of informant accuracy and inaccuracy. *Br. J. Dev. Psychol.* **27**, 331–342 (2009).
39. Koenig, M. A. & Harris, P. L. Preschoolers mistrust ignorant and inaccurate speakers. *Child Dev.* **76**, 1261–1277 (2005).
40. Koenig, M. A., Clément, F. & Harris, P. L. Trust in testimony: children's use of true and false statements. *Psychol. Sci.* **15**, 694–698 (2004).
41. Johnston, A. M., Mills, C. M. & Landrum, A. R. How do children weigh competence and benevolence when deciding whom to trust? *Cognition* **144**, 76–90 (2015).
42. Miyoshi, M. & Sanefuji, W. Focusing on different informant characteristics by situation: the dimensions of benevolence and competence in children's trust judgment. *Soc. Dev.* **31**, 1231–1239 (2022).
43. Clément, F., Koenig, M. & Harris, P. L. The ontogenesis of trust in testimony. *Mind Lang.* **19**, 360–379 (2004).
44. Harris, P. L. et al. in *Foundations of Metacognition* (eds Beran, M. J. et al.) 193–210 (Oxford Univ. Press, 2012); https://doi.org/10.1093/acprof:oso/9780199646739.003.0013
45. Kinzler, K. D., Corriveau, K. H. & Harris, P. L. Children's selective trust in native-accented speakers. *Dev. Sci.* **14**, 106–111 (2011).
46. Fu, G., Heyman, G. D., Chen, G., Liu, P. & Lee, K. Children trust people who lie to benefit others. *J. Exp. Child Psychol.* **129**, 127–139 (2015).
47. Koenig, M. A., Tiberius, V. & Hamlin, J. K. Children's judgments of epistemic and moral agents: from situations to intentions. *Perspect. Psychol. Sci.* **14**, 344–360 (2019).
48. Liu, D., Vanderbilt, K. E. & Heyman, G. D. Selective trust: children's use of intention and outcome of past testimony. *Dev. Psychol.* **49**, 439–445 (2013).
49. Margoni, F. & Nava, E. The development of intent-based epistemic trust. Preprint at *PsyArXiv* https://doi.org/10.31234/osf.io/btazs (2022).
50. Evans, A. D. & Lee, K. in *Handbook of Moral Development* 3rd edn (eds Killen, M. & Smetana, J. G.) 289–304 (Routledge, 2022).
51. Lee, K. Little liars: development of verbal deception in children. *Child Dev. Perspect.* **7**, 91–96 (2013).
52. Evans, A. D., Xu, F. & Lee, K. When all signs point to you: lies told in the face of evidence. *Dev. Psychol.* **47**, 39–49 (2011).
53. Evans, A. D. & Lee, K. Emergence of lying in very young children. *Dev. Psychol.* **49**, 1958–1963 (2013).

54. Fu, G., Evans, A. D., Xu, F. & Lee, K. Young children can tell strategic lies after committing a transgression. *J. Exp. Child Psychol.* **113**, 147–158 (2012).
55. Fu, G., Heyman, G. D., Qian, M., Guo, T. & Lee, K. Young children with a positive reputation to maintain are less likely to cheat. *Dev. Sci.* **19**, 275–283 (2016).
56. Kotaman, H. Impact of rewarding and parenting styles on young children's cheating behavior. *Eur. J. Dev. Psychol.* **14**, 127–140 (2016).
57. Talwar, V. & Lee, K. Development of lying to conceal a transgression: children's control of expressive behavior during verbal deception. *Int. J. Behav. Dev.* **26**, 436–444 (2002).
58. Zhao, L. et al. The moral barrier effect: real and imagined barriers can reduce cheating. *Proc. Natl Acad. Sci. USA* **117**, 19101–19107 (2020).
59. Bar-Tal, D. Sequential development of helping behavior: a cognitive-learning approach. *Dev. Rev.* **2**, 101–124 (1982).
60. Grueneisen, S. & Warneken, F. The development of prosocial behavior—from sympathy to strategy. *Curr. Opin. Psychol.* **43**, 323–328 (2022).
61. Hepach, R., Vaish, A. & Tomasello, M. Young children are intrinsically motivated to see others helped. *Psychol. Sci.* **23**, 967–972 (2012).
62. Liszkowski, U., Carpenter, M., Striano, T. & Tomasello, M. 12- and 18-month-olds point to provide information for others. *J. Cogn. Dev.* **7**, 173–187 (2006).
63. Warneken, F. How children solve the two challenges of cooperation. *Annu. Rev. Psychol.* **69**, 205–229 (2018).
64. Warneken, F. & Tomasello, M. Altruistic helping in human infants and young chimpanzees. *Science* **311**, 1301–1303 (2006).
65. Warneken, F. & Tomasello, M. Helping and cooperation at 14 months of age. *Infancy* **11**, 271–294 (2007).
66. Zhao, L. et al. Subtle alterations of the physical environment can nudge young children to cheat less. *Dev. Sci.* **25**, e13190 (2022).
67. Greitemeyer, T. & Osswald, S. Effects of prosocial video games on prosocial behavior. *J. Pers. Soc. Psychol.* **98**, 211–221 (2010).
68. Lyubomirsky, S. & Layous, K. How do simple positive activities increase well-being? *Curr. Dir. Psychol. Sci.* **22**, 57–62 (2013).
69. House, B., Henrich, J., Sarnecka, B. & Silk, J. B. The development of contingent reciprocity in children. *Evol. Hum. Behav.* **34**, 86–93 (2013).
70. Malti, T., Gasser, L. & Buchmann, M. Aggressive and prosocial children's emotion attributions and moral reasoning. *Aggress. Behav.* **35**, 90–102 (2009).
71. Malti, T., Eisenberg, N., Kim, H. & Buchmann, M. Developmental trajectories of sympathy, moral emotion attributions, and moral reasoning: the role of parental support. *Soc. Dev.* **22**, 773–793 (2013).
72. Nelson, L. J., Padilla-Walker, L. M. & Son, D. in *The Oxford Handbook of Parenting and Moral Development* (eds Laible, D. J. et al.) 355–373 (Oxford Univ. Press, 2019).
73. Padilla-Walker, L. M., Nielson, M. G. & Day, R. D. The role of parental warmth and hostility on adolescents' prosocial behavior toward multiple targets. *J. Fam. Psychol.* **30**, 331–340 (2016).
74. Williams, K. E. & Berthelsen, D. The development of prosocial behaviour in early childhood: contributions of early parenting and self-regulation. *Int. J. Early Child.* **49**, 73–94 (2017).
75. Merritt, A. C., Effron, D. A. & Monin, B. Moral self-licensing: when being good frees us to be bad. *Soc. Pers. Psychol. Compass* **4**, 344–357 (2010).
76. Miller, D. T. & Effron, D. A. in *Advances in Experimental Social Psychology* Vol. 43 (eds Zanna, M. P. & Olson, J. M.) 115–155 (Academic Press, 2010).
77. Cameron, S., Wilks, M. & Nielsen, M. Does helping now excuse cheating later? An investigation into moral balancing in children. *R. Soc. Open Sci.* **8**, 202296 (2021).
78. Maftei, A. & Holman, A. C. Moral in the future, better now: moral licensing versus behavioral priming in children and the moderating role of psychological distance. *Curr. Psychol.* **42**, 18904–18915 (2023).
79. Mills, C. M. & Keil, F. C. The development of cynicism. *Psychol. Sci.* **16**, 385–390 (2005).
80. Clément, F. To trust or not to trust? Children's social epistemology. *Rev. Phil. Psychol.* **1**, 531–549 (2010).
81. Harris, P. L. *Trusting What You're Told: How Children Learn from Others* (Harvard Univ. Press, 2012).
82. Heyman, G. D., Fu, G. & Lee, K. Evaluating claims people make about themselves: the development of skepticism. *Child Dev.* **78**, 367–375 (2007).
83. Isella, M., Kanngiesser, P. & Tomasello, M. Children's selective trust in promises. *Child Dev.* **90**, e868–e887 (2019).
84. Koenig, M. A. & Sabbagh, M. A. Selective social learning: new perspectives on learning from others. *Dev. Psychol.* **49**, 399–403 (2013).
85. Li, Q., Zhang, W., Heyman, G. D., Compton, B. J. & Lee, K. Susceptibility to being lured away by a stranger: a real-world field test of selective trust in early childhood. *Psychol. Sci.* **31**, 1488–1496 (2020).
86. Mills, C. M. Knowing when to doubt: developing a critical stance when learning from others. *Dev. Psychol.* **49**, 404–418 (2013).
87. Heyman, G. D., Sritanyaratana, L. & Vanderbilt, K. E. Young children's trust in overtly misleading advice. *Cogn. Sci.* **37**, 646–667 (2013).
88. Lee, K. & Cameron, C. A. Extracting truthful information from lies: emergence of the expression–representation distinction. *Merrill Palmer Q.* **46**, 1–20 (2000).
89. Mascaro, O. & Sperber, D. The moral, epistemic, and mindreading components of children's vigilance towards deception. *Cognition* **112**, 367–380 (2009).
90. Vanderbilt, K. E., Liu, D. & Heyman, G. D. The development of distrust. *Child Dev.* **82**, 1372–1380 (2011).
91. Rakoczy, H. Foundations of theory of mind and its development in early childhood. *Nat. Rev. Psychol.* **1**, 223–235 (2022).
92. Blake, P. R. & McAuliffe, K. 'I had so much it didn't seem fair': eight-year-olds reject two forms of inequity. *Cognition* **120**, 215–224 (2011).
93. Decety, J. & Cowell, J. M. The complex relation between morality and empathy. *Trends Cogn. Sci.* **18**, 337–339 (2014).
94. McAuliffe, K., Blake, P. R., Steinbeis, N. & Warneken, F. The developmental foundations of human fairness. *Nat. Hum. Behav.* **1**, 0042 (2017).
95. Fehr, E., Fischbacher, U. & Gächter, S. Strong reciprocity, human cooperation, and the enforcement of social norms. *Hum. Nat.* **13**, 1–25 (2002).
96. Gintis, H. Strong reciprocity and human sociality. *J. Theor. Biol.* **206**, 169–179 (2000).

Acknowledgements

The present research was supported by funding from the National Natural Science Foundation of China (grant no. 32171060) and the Major Project of Philosophy and Social Science Research of the Ministry of Education of China (grant no. 22JZD044) to L.Z., and from the Social Sciences and Humanities Research Council of Canada to K.L. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the paper.

Author contributions

L.Z. and K.L. designed the studies. H.M. collected the data. L.Z., K.L., P.L.H. and H.M. analysed the data and wrote the paper. L.Z. managed the project.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41562-024-01837-4.

Correspondence and requests for materials should be addressed to Li Zhao.

Peer review information *Nature Human Behaviour* thanks Jellie Sierksma and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

nature portfolio

Corresponding author(s): Li Zhao

Last updated by author(s): Nov 7, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used to collect data in this study.
Data analysis	We used R software (Version 4.3.1) and SPSS software (Version 26) for data analysis in this study. Analysis scripts are publicly available via an OSF repository at https://osf.io/yqdhp/.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The data sets are publicly available via an OSF repository at https://osf.io/yqdhp/.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	We included gender information for each condition in the Methods section. Gender-based analyses are provided in the Supplemental information section.
Reporting on race, ethnicity, or other socially relevant groupings	*Please specify the socially constructed or socially relevant categorization variable(s) used in your manuscript and explain why they were used. Please note that such variables should not be used as proxies for other socially constructed/relevant variables (for example, race or ethnicity should not be used as a proxy for socioeconomic status).* *Provide clear definitions of the relevant terms used, how they were provided (by the participants/respondents, the researchers, or third parties), and the method(s) used to classify people into the different categories (e.g. self-report, census or administrative data, social media data, etc.)* *Please provide details about how you controlled for confounding variables in your analyses.*
Population characteristics	A total of 328 5- to 6-year-old children were recruited from a kindergarten located in Eastern China in this study (mean age = 71.30 months, SD = 3.31 months, age range = 63.35 to 77.88; 168 boys). All children were Han Chinese from middle-class backgrounds living in a metropolitan city in P. R. China.
Recruitment	Children were recruited from a kindergarten located in Eastern China. Parents or legal guardians gave informed consent to allow their children to participate, and children also gave their oral assent prior to participating in the experiments. All eligible children in each classroom who were in attendance on testing days were given the opportunity to participate. The only exception was that data collection was stopped as soon as we reached the planned number of participants in each study. Participants received no compensations in any of the studies. Any potential self-selection biases, such as achievement motivation related to the task, time pressure of the task, or reputation concerns, are not likely to impact the generalization of our results, since our study used a randomized selection process and assignment to different conditions.
Ethics oversight	The current research was approved by the Scientific Research Ethics Committee of Hangzhou Normal University.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This is a quantitative experimental study using a naturalistic field design. We observed whether kindergarten children cheated or not after they had been entrusted by an adult to help her with a task.
Research sample	A total of 328 5- to 6-year-old children were recruited from a kindergarten located in Eastern China in this study (mean age = 71.30 months, SD = 3.31 months, age range = 63.35 to 77.88; 168 boys). All children were Han Chinese from middle-class backgrounds living in a metropolitan city in P. R. China. This sample is not representative for all humans due to it consisting of 5- to 6-year-old children from the kindergarten.
Sampling strategy	We predetermined a sample size of 41 participants for each condition by conducting a priori power analyses based on existing research using similar paradigms as well as our pilot testing. Specifically, we estimated a baseline cheating rate for the control condition of 55% and an average cheating rate for the experimental condition of 25%. The power analyses revealed that to achieve a significant condition effect with a power of 0.8, an alpha at 0.05, and an enrolment ratio of 1, the appropriate sample size was 41 for each condition.
Data collection	A female experimenter brought children from their classroom to the testing room individually. Children were randomly assigned to each condition in each of the five studies. Testing sessions were were recorded with a hidden video camera to be coded by two graduate students who were naïve to the study hypotheses.
Timing	Study 1: 10/2021; Study 2: 11/2021; Studies 3-5: 03/2023
Data exclusions	All participants passed the comprehension checks, and thus none were excluded.
Non-participation	No participants dropped out or declined the participation.
Randomization	Participants were randomly assigned to each condition in each of the five studies.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging