On the Developmental Origin of Intrinsic Honesty

Tai-Sen He, Lili Qin*

Abstract

Contrary to the self-interestedness assumption, numerous economic studies have documented that people are intrinsically honest. However, little is known about this trait's developmental origin. This study examines whether and the extent to which children in early childhood incur the intrinsic lying cost. We modified the commonly used coin-flip task into a child-friendly ball-drawing task with 10 trials and conducted the experiment with 225 child participants aged three to eight years old. We found that—although young children, on average, told two lies in the task (an average winning rate of 71%)—they lied significantly less than the maximum level (i.e., lying 100% of the time). The pattern was largely similar across gender and the age range studied. Furthermore, our child subjects' propensity to lie dropped by approximately 9% when they were randomly assigned to the treatment condition with an increased "perceived" intrinsic cost of lying. Overall, our results provide further support for the innate morality hypothesis: young children, as young as three years old, are willing to give up pecuniary rewards in order to remain honest.

Keywords: intrinsic honesty, intrinsic lying cost, dishonesty, young children

JEL Codes: C91, D6

* He: Economics Programme, School of Social Sciences, Nanyang Technological University, ts.he@ntu.edu.sg; Qin: Department of Psychology, National University of Singapore, psyql@nus.edu.sg. Research funding by the Ministry of Education, Singapore (MOE Academic Research Fund Tier 1) is gratefully acknowledged.

"The moral sense, or conscience is as much a part of man as his leg or arm."

—Thomas Jefferson (1787)

I. Introduction

A unique trait differentiating human beings from other animals is the capacity to engage in extremely large-scale cooperation with unrelated humans. A prominent example is the global economic system, which coordinates virtually every single person in the world to produce goods and services to meet global human needs. Although self-interest and its corresponding incentive-compatible schemes ensure the smooth functioning of the system, trust is essentially the foundation of the human economic system. In the absence of adequate trust, people can barely engage in transactions—how many people would dare to deposit their money in banks or invest in stock markets if they did not trust the institutions (Guiso, Sapienza, and Zingales 2004)? Would people believe in "organic" food? Would anyone walk into a for-profit hospital for medical treatments? Would parents have any peace of mind while their young children were in a private childcare facility? Ultimately, how many market institutions could survive without people sufficiently believing that others would "do the right thing" even when it was not perfectly aligned with their own self-interest?

Importantly, then, economists must understand why the vast majority of people choose to behave honestly and morally. Tracing back to the 1980s, Robert Frank (1987) proposed a novel theoretical framework arguing that morality (e.g., blushing when lying) is an evolved human trait as it functions as a commitment device to enhance people's propensity to cooperate even if the ensuing action deviates from the Nash equilibrium. More recently, accumulating empirical evidence has demonstrated that people are intrinsically honest[1]: even in a tightly

[1] Intrinsic honesty and intrinsic lying cost are two sides of the same coin. Intrinsically honest individuals incur a high intrinsic cost when lying.

controlled condition in which subjects have a clear monetary incentive to make a false statement, which can neither be detected nor affect others' payoffs, they still lie very little (Abeler, Becker, and Falk 2014; Abeler, Nosenzo, and Raymond 2019; Erat and Gneezy 2012; Fischbacher and Föllmi-Heusi 2013; Gächter and Schulz 2016; Gneezy, Kajackaite, and Sobel 2018; Lundquist et al. 2009; Mazar, Amir, and Ariely 2008).[2]

Despite the prevalence of intrinsic honesty among the adult population and its importance in sustaining the exchange-based economy, very little attention has been paid to understanding the developmental origin of this behavioral inclination. Do people acquire it solely from social and cultural learning? Or is it an innate predisposition? Understanding the root cause of intrinsic honesty may influence how we think about the evolutionary basis of the human economic system.

This fundamental question can be partially addressed by examining whether young children, especially younger ones, also incur the intrinsic lying cost as they are less influenced than older children and adults by cultural and environmental factors (Gächter and Schulz 2016). For this purpose, we performed an incentive-compatible test of intrinsic honesty on the youngest children studied to date. More specifically, we modified the popular coin-flip paradigm into a young-child-friendly version and implemented it on a total of 225 child subjects aged three to eight years. In each of the 10 trials, participants privately drew a ball from an opaque box containing equal numbers of red and blue balls and were asked to report the ball color to the experimenter. If a red ball was reported, the child could choose a reward from a box of mixed tokens—which included candies, biscuits, and stickers—while reporting

[2] Notably, in a famous study, Cohn, Fehr, and Marechal (2014) reported on banking business culture having a negative effect on bank professionals' honesty. Even in the treatment condition where banking professionals were primed with the infamous banking business culture, their banker subjects remained fairly honest with a success rate of merely 58% in the coin-toss task. Likewise, in Gächter and Schulz's (2016) study investigating the prevalence of rule violation (PRV) on student participants' lying behavior, a relatively high level of intrinsic honesty among subjects from the low PRV countries was still observed.

a blue ball yielded no reward. Obviously, child participants could win rewards by misreporting the outcome of the private ball-drawing without being detected.

Moral and developmental psychology has a long history of exploring children's moral judgment and behaviors. However, unlike economists who analyze human behaviors from the self-interest perspective, psychological research in this area has yet to directly examine the reasons people, including young children, do not lie maximally. In fact, intrinsic honesty or intrinsic lying cost is not even commonly known by psychologists. As a result, although whether human morality is innate has been a decades-long question in several related fields—such as psychology, anthropology, biology, and even philosophy—and increasingly more recent evidence suggests that people are born with initial moral senses[3], none of the existing studies thus far serves as a satisfactory test of intrinsic honesty by meeting the strict conditions required by economists.[4] In view of this sense, the present study thus contributes to the extensive literature on innate morality by performing possibly the strongest test of intrinsic honesty on the youngest children studied to date.[5] Put simply, we test whether young children, as young as three years old, are willing to give up pecuniary rewards in order to remain honest.

[3] There is growing evidence suggesting that infants are endowed with compassion, empathy, and an initial sense of fairness (e.g., McAuliffe et al. 2017; Warneken and Tomasello 2006, 2009). Hamlin, Wynn, and Bloom (2007) even demonstrated that people are born with some sort of moral "instinct." In their experiment, six- and ten-month-old infants were presented with a display in which puppets helped or hindered each other. The child subjects were then given the option to choose one puppet to play with. Most of the subjects chose to play with the helpers and not the hinderers.

[4] The test of intrinsic honesty should meet four conditions: 1) subjects have a clear pecuniary incentive to lie; 2) lying cannot be detected; 3) lying does not affect others' payoffs; and 4) experimenters provide no hint of what is right or wrong behavior. To our best knowledge, the existing psychological experiments on children's lying violate at least one of the above conditions. For example, the most frequently used task by psychologists studying children's lying behavior is the temptation resistance paradigm developed by Sears, Rau, and Alpert (1965). In this paradigm, a child participant is typically told explicitly by a researcher not to peek at or play with a toy when left alone and then is asked if he or she peeked when the researcher returns. This paradigm violates both conditions 2) and 4) as subjects are told that peeking at the toys is a wrong behavior and that their lies can be subsequently detected, either through a follow-up question or via video recordings from a hidden camera. On the other hand, in several important psychological research studies examining very young children's moral senses—including Hamlin, Wynn, and Bloom (2007) and Bussey (1992)—subjects simply indicated their preference without making an "incentivized" decision, which thus violates condition 1). Even for several psychological works that involve pecuniary rewards to incentivize subjects' lying (e.g. hide and seek paradigm), conditions 2) and/or 3) are typically violated as subjects' lies can be subsequently detected and/or have impacts on others' payoffs.

[5] There are only a very limited number of studies in economics that have explored the lying behaviors among children, with the exception of Bucciol and Piovesan (2011), Glätzle-Rützler and Lergetporer (2015), and

Our results provide further empirical support for the innate morality hypothesis. First, although the experimenter cannot verify the outcome for each ball draw, the extent of lying can be measured by the difference between the reported distribution of red balls across the two conditions to the distributions of fair draw (50% red balls reported) and profit maximization (100% red balls reported). We found that, while the child subjects told an average of two lies in the form of misreporting the ball color (an average winning rate of 71%), they clearly lied substantially less than the maximum level of 100%. Notably, we did not find any significant age effect, meaning that the intrinsic cost of lying emerges from very early stages of life, and this pattern remains similar across the ages of three to eight years old. Second, our subjects' propensity to lie dropped by 9% when they were randomly assigned to the treatment condition with a heightened "perceived" intrinsic cost of lying. Overall, these findings constitute evidence that children in early childhood, like adults, also incur the intrinsic cost of lying.

II. Experimental Design and Procedures

2.1 Procedures

We recruited child participants aged three to eight years old by advertising with local parenting groups and by liaising with childcare centers. The majority (80.4%) of the parents who brought their children in were mothers as they are often the main caregiver of children. All sessions, which each took approximately a half-hour to complete, were conducted in a small multifunctional seminar room at a major research university in Singapore (see online Appendix A for more detail on the lab setting). Instructions were given verbally by female research assistants (see online Appendix B for experimental instructions).

Maggian and Villeval (2016). However, these studies focused on children in middle childhood and/or adolescents, and therefore, there remains no thorough understanding regarding the ontogenic development of the intrinsic lying cost in early childhood.

The experiment proceeded as follows. Verbal and written consent were obtained from parents and children before the study. Then, experimenters demonstrated and conducted the ball-drawing task, a modified version of the coin-flip paradigm. The ball-drawing task required child participants to randomly draw a ball from an opaque box containing equal numbers of red and blue balls. The child would report the result of his or her ball draw to an experimenter, whose view was blocked by a black screen to eliminate concerns about lie detection. If a red ball was reported, the child could choose a reward from a box of mixed tokens. However, if a blue ball was reported, no reward was given. The task consisted of ten trials for each child. We did not explicitly tell the subjects they could misreport the observed ball color to earn tokens.

A training phase was conducted to ensure that each child participant fully comprehended the task and to overcome potential learning effects. More specifically, experimenters checked whether the child subjects understood that they could receive a reward only if a red ball was reported and that they knew how to report the draw outcomes according to the treatment conditions. If a child failed the comprehension check, the training phase was repeated. Before the real ball-drawing task began, the child was asked two questions: "Can you see me?" and "Can you see Mummy/Daddy?" The child was then reassured that neither his/her parent nor the experimenter was able to see the child's action, eliminating concerns about lie detection. The child proceeded to complete ten successive ball draws with replacement (i.e. the drawn ball was placed back into the box for the next draw).[6] At the end of the experiment, the child's parents completed a post-experiment questionnaire on their demographic characteristics.

2.2 Linguistic Manipulation

We complemented our design with a linguistic manipulation that aimed to heighten the "perceived" intrinsic lying cost. If subjects incur the intrinsic cost of lying, they are predicted

[6] That is, the drawn ball was placed back into the box for the next draw.

to reduce the propensity to lie in the treatment condition with a higher perceived lying cost. We designed two treatment conditions in which unobtrusive manipulation was naturally embedded into the ball-drawing task on a trial-by-trial basis. If child subjects were randomly assigned to the heighted perceived lying cost (HC) condition, they were instructed to report the outcome of the ball-drawing using a first-person pronoun by saying "*I got a red/blue ball.*" Otherwise, they simply said "*red/blue ball*" in the control condition.

Our linguistic manipulation was motivated by the existing studies in psychology and communications that have previously demonstrated that individuals use fewer self-oriented pronouns (e.g., "I" and "me") when lying than when telling the truth, possibly to distance themselves from the lies (Newman et al. 2003; Hancock et al. 2008). The psychological literature on self-awareness also supports our hypothesis. Self-awareness refers to individuals' capacity to take themselves as the object of thought—people can think, act, and experience, and they can also think about what they "themselves" are thinking, doing, and experiencing. Children typically demonstrate self-awareness by age two.[7] According to Duval and Wicklund's (1972) objective self-awareness theory, inducing self-awareness can increase one's self-evaluation and enhance sensitivity to social and moral norms, rules, and standards. Numerous studies in psychology provide extensive support for this theory with evidence from adults and children. For example, Beaman et al. (1979) reveal that children are more likely to adhere to the rule (taking only one candy from the bowl) when they are individuated (by being asked their names and where they live) and a mirror is placed in front of the candy bowl to induce their self-awareness. In engaging with lying behavior, individuals with greater self-awareness are more likely to evaluate their lying behavior with internal standards and thus come into greater conflict with moral values. Consequently, the mandatory use of the pronoun

[7] The Rouge test is commonly used to test children's development in self-awareness and the ability to recognize themselves as individuals. A red dot is secretly placed on children's faces before they are presented with their reflections in a mirror. Those who indicate an awareness of the red dot are considered to have passed the task.

"I" heightens the "perceived" intrinsic lying cost and is predicted to reduce children's propensity to lie.[8]

III. Results

3.1 Participants

A total of 240 children were recruited and participated in the experiment. Data from fifteen subjects was excluded.[9] Thus, data from 225 child participants was valid for analysis. The age of the children ranged from three to eight years, with an average of five (s.e. = 1.3) years. Half (49.8 %) of the participants were girls. Additionally, 91.1% of the participants were Chinese, and nearly all (99.1%) were Singaporean citizens or permanent residents. Table 1 provides summary statistics of the participants by treatment. There were no statistical differences for all of the demographic variables between the two treatment conditions except for a marginally significant difference in the age variable, which will be included as a control in regressions.

[Insert Table 1 about here]

3.2 Aggregate Results

In line with other studies in economic literature, we cannot verify participants' answers to directly observe lying for each individual trial. Instead, we compared the distribution of reported wins (red balls) to that of the expected rate to infer whether and the extent to which lying occurred. Across all treatments, our sample child subjects reported an average of seven

[8] Relatedly, Falk (2017) finds that people are less likely to inflict a painful electric shock on another subject to receive a monetary payment when they see their own faces on the decision screen in a real-time video feed, compared to two control conditions in which they see either no video at all or a video of an unrelated other.

[9] The data from one child subject was not saved successfully due to a technical glitch, while the remaining fourteen observations were excluded due to the child subjects' inability to adhere to the instructions given by experimenters, such as refusing to make any decisions, simply staying silent, and needing their parents to sit next to them. Among these fourteen observations, six were in the control condition, and eight were in the self-awareness condition: three were three years old, seven were four years old, and four were five years old.

wins (red balls), indicating that they told an average of two lies when reporting the ball color throughout the ten trials. The success rate was significantly higher than 50% ($p < 0.01$, two-sided one sample t-test) and significantly lower than 100% ($p < 0.01$, two-sided one sample t-test). The overall distribution of reported red balls for the total sample of 225 children is shown in Figure 1. Not surprisingly, the distribution is right-skewed as child subjects tended to over-report the number of red balls; however, most of them lied partially, and in fact, less than one-fifth (18.7%) of the subjects lied to the maximum level (i.e., 100% of the time) by reporting a total of ten red balls. These results suggest that, although young children tell lies by misreporting the ball color to increase their pecuniary benefit, they are likely to incur the intrinsic cost of lying. Hence, the average winning rate was substantially lower than 100%.

[Insert Figure 1 about here]

Result 1: The average number of wins was significantly lower than the maximum level of ten wins, suggesting that children in early childhood, like adults, incur the intrinsic cost of lying.

Next, we considered the impact of gender and age on lying behavior. Gender differences in ethical preferences among adult populations have been well-documented in previous studies (Dreber and Johannesson 2008), but the results are rather ambiguous among child subjects.[10] In our study, we did not find a significant gender difference in lying among our sample child subjects across the two treatment conditions ($p > 0.1$, two-sided t-test). Regarding the age effect, we followed Fehr, Bernhard, and Rockenbach (2008) and categorized our child subjects into three age groups: 3/4 ($n = 84$), 5/6 ($n = 111$), and 7/8 ($n = 30$). The average number of wins is 7.25 in the 3/4 group, 6.99 in the 5/6 group, and 7.10 in the 7/8

[10] For instance, Glätzle-Rützler and Lergetporer (2015) did not note any gender difference, while Maggian and Villeval (2016) show that this gender gap disappears among older child subjects.

group, respectively. Again, we did not find any significant difference in lying among any two of the three age groups ($p > 0.1$, two-sided t-test).[11]

Result 2: The pattern of intrinsic lying cost was similar across gender and the age range from three to eight years old.

3.3 The Heightened Perceived Lying Cost Treatment Effect

We predicted that, if our sample child subjects incurred the intrinsic cost of lying, the mandatory use of the pronoun "I" in reporting ball color would heighten the perceived intrinsic lying cost and thus reduce their propensity to lie, compared to the control condition. Figure 2 illustrates the treatment effect. Consistent with our prediction, those in the HC group reported an average of 6.8 wins, while those in the control group reported an average of 7.4 wins ($p = 0.03$ using two-sided Mann-Whitney test). Such outcomes mean that the subtle linguistic manipulation reduced the winning rate by roughly 9% (or six percentage points). In addition, we compared the distributions of reported red balls in the HC treatment condition vs. the control condition, as shown in Figure 3. The difference was statistically significant at 5% using the Kruskal-Wallis equality-of-populations rank test. Notably, the proportion of complete liars dropped largely from 25.4% in the control condition to 11.7% in the HC condition. These results reveal the effectiveness of using the first-person pronoun "I" in reducing young children's propensity to lie. Importantly, the fact that young children respond positively to this manipulation provides another piece of evidence that they incur intrinsic or moral costs when telling lies.

Result 3: The heighted perceived lying cost through mandatory use of the pronoun "I" significantly reduced young children's propensity to lie.

[11] The success rate for each subgroup stratified by gender and age was significantly higher than 50% ($p < 0.01$, two-sided one sample t-test) and significantly lower than 100% ($p < 0.01$, two-sided one sample t-test).

[Insert Figure 2 about here]

3.4 Probit Regression Analysis

The probit regression reported in Table 2 provides additional evidence for Results 1–3. Here, the dependent variable is reporting a win (1 for red ball and 0 otherwise). First, we regressed on the treatment dummy only and observe lower frequency of lying in the HC condition (coefficient of -0.06, p-value = 0.02), as illustrated in Column 1. Column 2 shows the results from exploring the period effect. The coefficient of *Period* is significantly positive, suggesting the learning effect. Perhaps child participants better understand the material benefits of lying throughout the experiment and therefore are more likely to lie with periods. Alternatively, once a child reports a red ball and receives a token, he or she may become more tempted to receive more tokens and thus lie more often. Even more, the subjects could potentially adapt themselves to lying with repetition (Garrett et al. 2016). Despite this, the winning rate of 76.9% in the last trial remained substantially lower than the maximum level of 100%. We further controlled for demographic variables, including gender (Female = 1 if the child subject was a girl; otherwise, Female = 0), age (the child subject's age in years), race (Chinese = 1 if the parent reported the child subject's race to be Chinese; otherwise, Chinese = 0), and immigration status (Citizen or PR = 1 if the parent reported the child's immigration status to be Singapore citizen or permanent resident; otherwise, Citizen or PR = 0), as shown in Column 3. None of the coefficients of these demographic controls were statistically significant, meaning that the intrinsic lying cost was not associated with these demographic characteristics, corroborating Result 2.

[Insert Table 2 about here]

VI. Conclusion

Although several recent economic studies have documented that adults incur a high intrinsic cost of lying, very limited research has examined whether and the extent to which children, especially very young ones, incur the intrinsic lying cost. Using an economics approach, this study revealed the prevalence of the intrinsic lying cost among young children between 3 and 8 years old. Furthermore, child subjects' propensity to lie was significantly reduced when they were randomly assigned to the treatment condition where subjects are required to report ball color using the pronoun "I" to aggravate the "perceived" intrinsic lying cost. Overall, our results demonstrate that the intrinsic lying cost emerges at very early stages of one's life, and its pattern remains similar across the age range between 3 and 8 years old. We suggest that future research should systematically explore the prevalence of the intrinsic cost of lying in separate developmental stages to gain a fuller understanding of the progressive trajectory of the intrinsic lying cost.

In sum, this is the first study in economics and related fields that performs the test of intrinsic honesty on the youngest children studied to date. Although we cannot conclude that humans are born to be intrinsically honest, our results shed important light on the developmental origin of intrinsic honesty. In particular, we tested the decades-old innate morality hypothesis from an economist's perspective, and our results provide further and stronger empirical evidence supporting this hypothesis: human beings as young as three years old have started to exhibit intrinsic honesty in an economically meaningful way—they are willing to give up pecuniary rewards in order to remain honest.

References

Abeler, Johannes, Anke Becker, and Armin Falk. 2014. "Representative Evidence on Lying Costs." *Journal of Public Economics* 113 (May): 96–104. https://doi.org/10.1016/j.jpubeco.2014.01.005.

Abeler, Johannes, Daniele Nosenzo, and Collin Raymond. 2019. "Preferences for Truth-Telling." *Econometrica*. https://doi.org/10.3982/ecta14673.

Beaman, Arthur L., Bonnel Klentz, Edward Diener, and Soren Svanum. 1979. "Self-Awarenesss and Transgression in Children: Two Field Studies." *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.37.10.1835.

Bucciol, Alessandro, and Marco Piovesan. 2011. "Luck or Cheating? A Field Experiment on Honesty with Children." *Journal of Economic Psychology*. https://doi.org/10.1016/j.joep.2010.12.001.

Bussey, Kay. 1992. "Lying and Truthfulness: Children's Definitions, Standards, and Evaluative Reactions." *Child Development*. https://doi.org/10.1111/j.1467-8624.1992.tb03601.x.

Cohn, Alain, Ernst Fehr, and Michel Andre Marechal. 2014. "Business Culture and Dishonesty in the Banking Industry." *Nature* 516 (729). https://doi.org/10.1038/nature13977.

Dreber, Anna, and Magnus Johannesson. 2008. "Gender Differences in Deception." *Economics Letters*. https://doi.org/10.1016/j.econlet.2007.06.027.

Duval, Shelley, and Robert A Wicklund. 1972. "A Theory of Objective Self Awareness." *A Theory of Objective Self Awareness*.

Erat, Sanjiv, and Uri Gneezy. 2012. "Whites Lies." *Management Science*. https://doi.org/10.1177/1745691612459060.

Falk, Armin. 2017. "Facing Yourself: A Note on Self-Image." *CESifo Working Paper*.

Fehr, Ernst, Helen Bernhard, and Bettina Rockenbach. 2008. "Egalitarianism in Young Children." *Nature*. https://doi.org/10.1038/nature07155.

Fischbacher, Urs, and Franziska Föllmi-Heusi. 2013. "Lies in Disguise-an Experimental Study on Cheating." *Journal of the European Economic Association*. https://doi.org/10.1111/jeea.12014.

Frank, Robert. 1987. “If Homo Economicus Could Choose His Own Utility Function, Would He Want One with a Conscience?” *American Economic Review* 77 (4): 593–604. www.jstor.org/stable/1814533.

Gächter, Simon, and Jonathan F. Schulz. 2016. “Intrinsic Honesty and the Prevalence of Rule Violations across Societies.” *Nature*. https://doi.org/10.1038/nature17160.

Garrett, Neil, Stephanie C. Lazzaro, Dan Ariely, and Tali Sharot. 2016. “The Brain Adapts to Dishonesty.” *Nature Neuroscience*. https://doi.org/10.1038/nn.4426.

Glätzle-Rützler, Daniela, and Philipp Lergetporer. 2015. “Lying and Age: An Experimental Study.” *Journal of Economic Psychology*. https://doi.org/10.1016/j.joep.2014.11.002.

Gneezy, Uri, Agne Kajackaite, and Joel Sobel. 2018. “Lying Aversion and the Size of the Lie.” *American Economic Review*. https://doi.org/10.1257/aer.20161553.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2004. “The Role of Social Capital in Financial Development.” *American Economic Review*. https://doi.org/10.1257/0002828041464498.

Hamlin, J. Kiley, Karen Wynn, and Paul Bloom. 2007. “Social Evaluation by Preverbal Infants.” *Nature*. https://doi.org/10.1038/nature06288.

Hancock, Jeffrey T., Lauren E. Curry, Saurabh Goorha, and Michael Woodworth. 2008. “On Lying and Being Lied to: A Linguistic Analysis of Deception in Computer-Mediated Communication.” *Discourse Processes*. https://doi.org/10.1080/01638530701739181.

Lundquist, Tobias, Tore Ellingsen, Erik Gribbe, and Magnus Johannesson. 2009. “The Aversion to Lying.” *Journal of Economic Behavior and Organization*. https://doi.org/10.1016/j.jebo.2009.02.010.

Maggian, Valeria, and Marie Claire Villeval. 2016. “Social Preferences and Lying Aversion in Children.” *Experimental Economics*. https://doi.org/10.1007/s10683-015-9459-7.

Mazar, Nina, On Amir, and Dan Ariely. 2008. “The Dishonesty of Honest People: A Theory of Self-Concept Maintenance.” *Journal of Marketing Research*. https://doi.org/10.1509/jmkr.45.6.633.

McAuliffe, Katherine, Peter R. Blake, Nikolaus Steinbeis, and Felix Warneken. 2017. “The Developmental Foundations of Human Fairness.” *Nature*

Human Behaviour. https://doi.org/10.1038/s41562-016-0042.

Newman, Matthew L., James W. Pennebaker, Diane S. Berry, and Jane M. Richards. 2003. "Lying Words: Predicting Deception from Linguistic Styles." *Personality and Social Psychology Bulletin*. https://doi.org/10.1177/0146167203029005010.

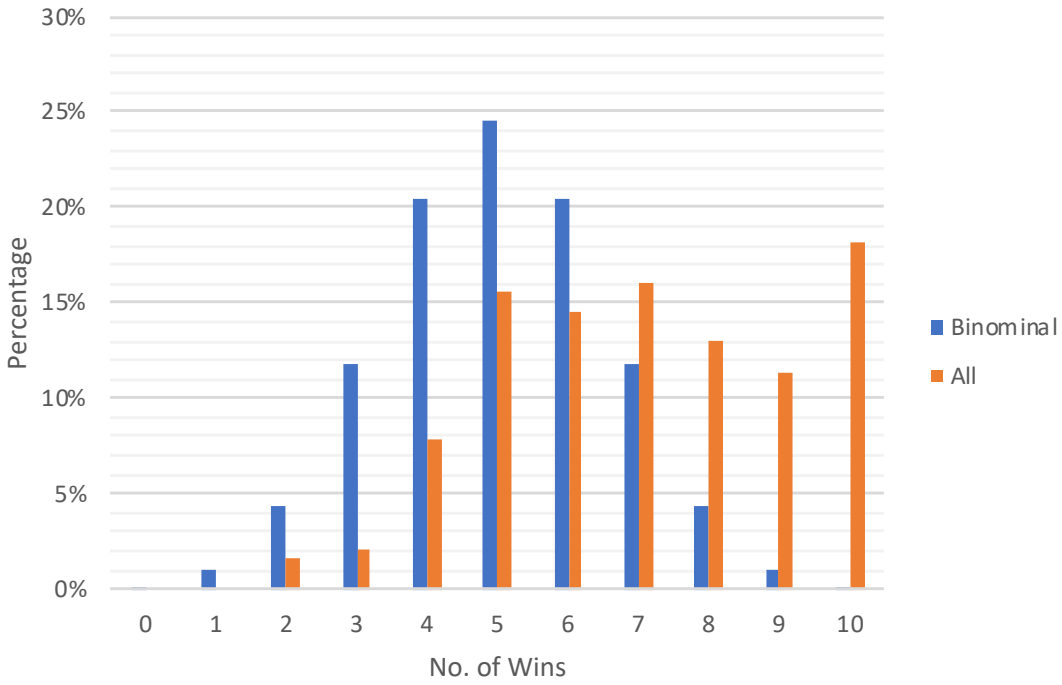
Sears, Robert R., Lucy Rau, and Richard Alpert. 1965. *No Identification and Child Rearing*. New York: John Wiley.

Warneken, Felix, and Michael Tomasello. 2006. "Altruistic Helping in Human Infants and Young Chimpanzees." *Science*. https://doi.org/10.1126/science.1121448.

———. 2009. "Varieties of Altruism in Children and Chimpanzees." *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2009.06.008.

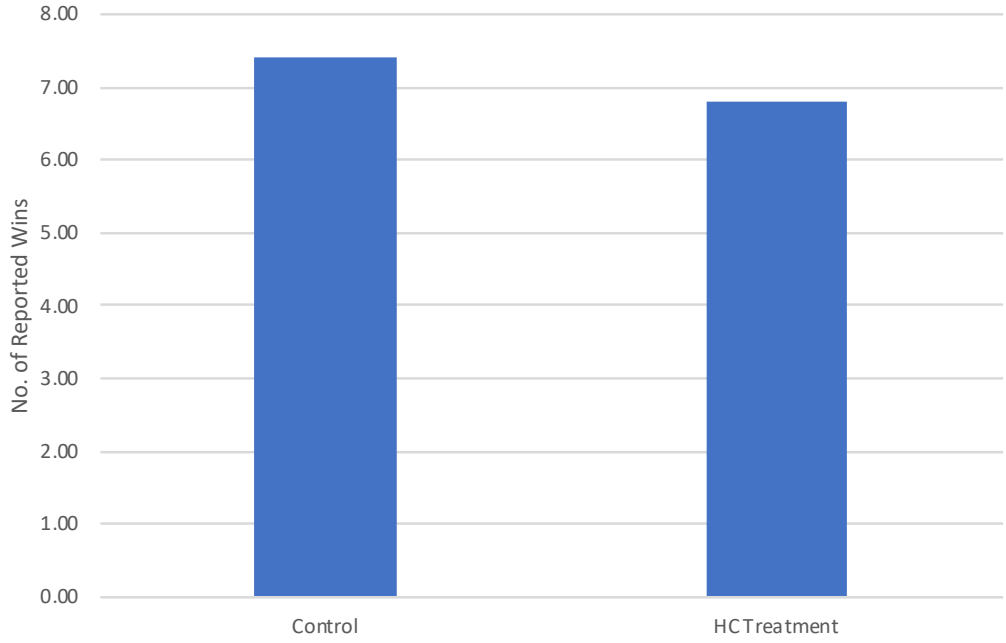
Tables and Figures

Figure 1. Distribution of wins reported by child subjects



Note: Distribution of wins reported by subjects in both control and treatment conditions in comparison to the binomial distribution implied by honest reporting.

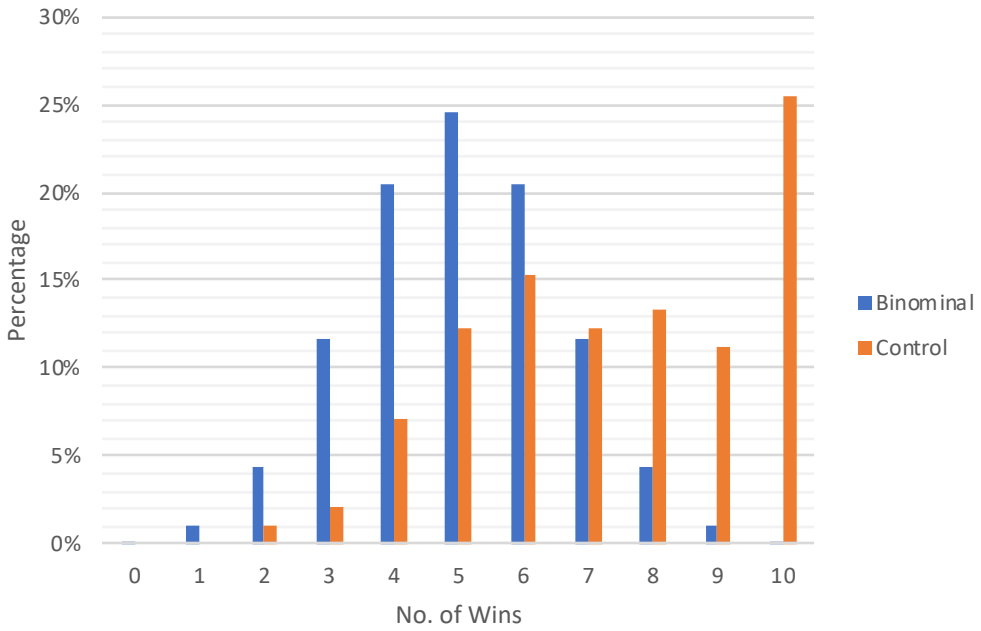
Figure 2. The impact of heightened perceived lying cost treatment on dishonesty



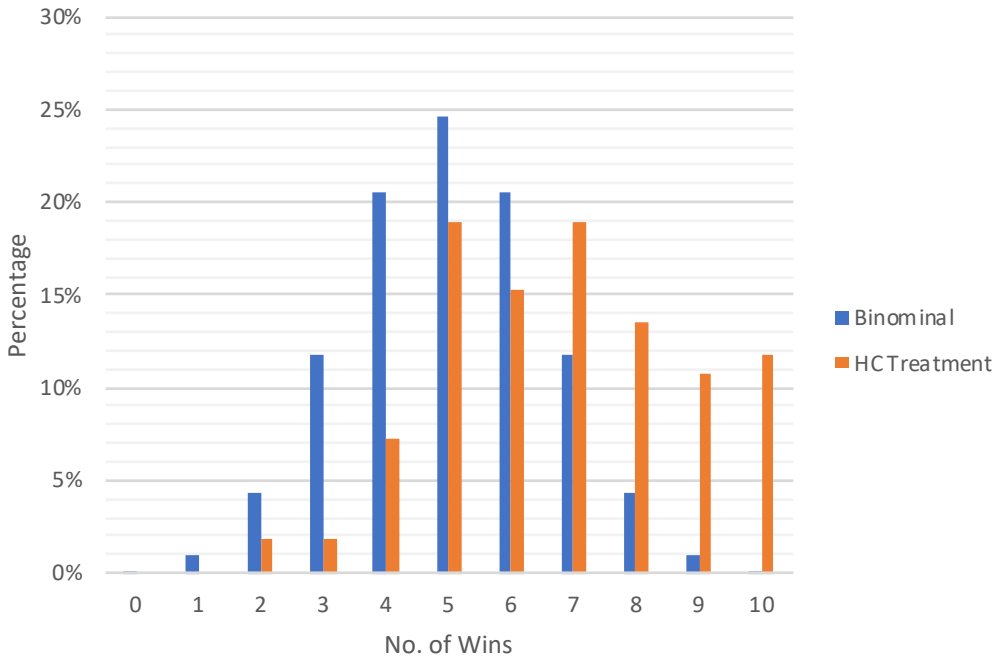
Note: The child subjects in the HC treatment condition reported an average of 6.8 wins, which is significantly lower than the that of the control group (P = 0.03, two-sided Mann-Whitney test; n = 225).

Figure 3. Comparison of the distribution of wins reported by child subjects by treatment

A



B



Note: (A) Distribution of reported wins in the control condition in comparison to the binomial distribution implied by honest reporting. (B) Distribution of reported wins in the HC treatment condition in comparison to the binomial distribution.

Table 1. Summary statistics

Treatment	Control	HC Treatment	All Subjects
Female	0.5088	0.4865	0.4978
	(0.5021)	(0.5020)	(0.5011)
Age	0.4877	0.5207	5.0400
	(1.2701)	(1.3220)	(1.3036)
Chinese	0.9035	0.9189	0.9111
	(0.2966)	(0.2742)	(0.2852)
Singapore Citizen/PR	0.9825	1.0000	0.9911
	(0.1319)	(0.0000)	(0.0941)
No. of Observations	114	111	225

Note: Standard deviations are reported in parentheses. Proportion tests (in the case of binary variables) and t-tests comparing demographics by treatment do not show statistically significant differences except for a marginally significant difference in age. These demographic variables are controlled in regression analysis.

Table 2. Effect of heighted lying cost on dishonesty

	(1)	(2)	(3)
Dependent Variable:	1 if Successful Outcome		
HC Treatment	-0.0611**	-0.0610**	-0.0610**
	(0.0273)	(0.0273)	(0.0283)
Period		0.0072**	0.0075**
		(0.0029)	(0.0030)
Female			0.0363
			(0.0265)
Age			-0.0089
			(0.0104)
Chinese			- 0.0065
			(0.0423)
Citizen or PR			0.1260
			(0.1211)
Observations	2250	2250	2250
No. of clusters	225	225	225

Note: Probit estimates. Reported results are average marginal effects. Robust standard errors corrected for clustering on the individual level are in parentheses. The decision to report a win is regressed on a dummy for the HC treatment condition in Column (1). Column (2) includes period effect. Column (3) further controls for individual characteristics.
*** Significant at the 1% level.
** Significant at the 5% level.
* Significant at the 10% level.

Online Appendix A: Lab Setup

1. Experimenter and child subject sit on different sides of the black screen.

2. A black cloth is attached to the opening of the box to prevent children from seeing into the box and "finding" a red ball.

3. Inside of the ball-drawing box, there are six red and six blue balls.

4. Tokens includes candies, biscuits, and stickers.

Online Appendix B: Experimental Instructions for the Ball-Drawing Task

Pre-task Parent Briefing
Give briefing on procedure and to refrain from giving comments to the child.

A. Training Phase:

1. Experimenter demonstration

SAY	DO
Briefing by Experimenter 1 "Today we will play a fun game. Let me and (E2) show you how to play ok? Now sit together with (E2) and learn how to play the game". E2 will take a ball from this box, tell me the colour of the ball, and put it back into the box. There are two different coloured balls in the box: red and blue. If E2 gets a red ball, E2 can choose 1 thing from this box (reward) If E2 gets a blue ball, she would get nothing. We show you how to play ok?	E1 sits **behind** screen. E2 sits with child on the same side.
Red Ball Experimenter 1: "Take a ball" Experimenter 2: **"(I got a) red ball"** Experimenter 1: Ok now put the ball back into the box. Experimenter 1: "Yay! Choose a present from the box." Experimenter 2: (Explain to the child) "I need to say **"(I got a) red ball"**. Because the ball is red, I can take 1 present."	E2 draws ball, shows child E2 puts ball back E2 takes reward
Blue Ball Experimenter 1: "Take a ball" Experimenter 2: **"(I got a) blue ball"** Experimenter 1: Now put the ball back into the box. Experimenter 1: "Sorry, you get nothing." Experimenter 2: (Explain to the child) "I need to say **"(I got a) blue ball"**. Because the ball is blue, I cannot take anything."	E2 draws ball, shows child E2 puts ball back

2. Comprehension check

SAY	DO
Experimenter 1: "Take a ball" **Experimenter 2: (ask child) "What should I say?"** • **The child should report "(I got a) red/blue ball" (check condition)** • Correct the child if the child merely says "red/blue".	E2 draws ball (Wait for answer)
Experimenter 1: "Put the ball back into the box." **Experimenter 2: (ask child) "What should I do now?"** • **The child should say "put the ball back."** • If correct – encourage child: "Yes, that's correct. I put the ball back" • If wrong – remind child	(Wait for answer)

Experimenter 1:
• If red: "Yay! You can select a present from the box."
• If blue: "Sorry, you get nothing."
Experimenter 2: (ask child) "What should I do now?"
• **Child should say take/not take a reward.**
• If correct – Encourage the child:
  ○ If Red: "Yes that's correct. Because the colour is red, I can take 1 present";
  ○ If Blue: "Yes that's correct. Because the colour is blue, I cannot take a present."
• If wrong – correct the child, show correct behaviour
	(Wait for answer)

Repeat step 2 until the child can succeed <u>3 times in a row</u>.
(i.e. Can correctly report colour + knows in what situation she/he could get a reward or not)
 	
• If **pass**:
  ○ Stop the training: "This is pretty fun, right? Do you want to play the game?"
  ○ **Continue to Study Phase**
• If **fail**: Repeat Training Phase	

B. Study Phase:

1. Briefing

SAY	DO
Remind parents not to make any comments or actions during the full course of the Study Phase, except when prompted by the experimenter. Ensure child is seated behind screen.	E2 leaves
<u>Parent View Check (done by E1)</u> "You have learned how to play this game! Now I have a question for you." Q: "When you sit behind this screen, can you see <u>mommy/daddy</u>?" Q: "Do you think <u>mommy/daddy</u> can see you?" Q: Let's check. "<u>Mommy/daddy</u>, can you see (child's name)'s face?" • Ensure parent responds verbally • If parent can see child, adjust accordingly. "You cannot see mommy, mommy cannot see you. So, mommy doesn't know what coloured ball you have. Only you know."	Wait for child's answer
<u>Experimenter View Check</u> *Sit behind screen* Q: "When you sit behind this screen, can you see me?" Q: "Do you think I will know what ball you get or not?" • Response: "That's right/Nope, I cannot see you!" "You cannot see me, I cannot see you. So, I don't know what coloured ball you have. Only you know."	Wait for child's answer
<u>Reminder</u> • If the colour is RED, you will GET a present • If the colour is BLUE, you will NOT get a present • You should say loudly **"I got a red/blue ball"** OR **'Red Ball/Blue Ball'**. Okay? • Now we will start.	

2. Study Phase

SAY	Remarks
"Draw a ball" *Child draws ball.* Q: What did you get? *Experimenter notes: Record reported colour* If child does not follow condition, record response as null and remind child again. "Ok, put the ball back into the box."	Ensure child responds, **"I got a red/blue ball"** *OR* **"red/blue ball"** Ensure child puts ball back.
If red: "Yay! Choose a present from the box". If blue: "Sorry, you get nothing" Let's try again!	Ensure child takes 1 reward only. *shake box again
Repeat steps for 10 times	