

SUPPLEMENTARY INFORMATION

Infants Expect Some Degree of Positive and Negative Reciprocity between Strangers

Authors:

Kyong-sun Jin^{1*}, Fransisca Ting^{2,3*}, Zijing He^{4*}, and Renée Baillargeon^{5*}

Affiliations:

¹Department of Psychology, Sungshin Women's University, 2, Bomun-ro 34da-gil, Seongbuk-gu, Seoul, 02844, Republic of Korea.

²Department of Psychology, University of Toronto, 100 St. George Street, Toronto, ON M5S 3G3, Canada.

³Department of Psychological and Brain Sciences, Boston University, 64 Cummington Mall, Boston, MA 02215, United States.

⁴Department of Psychology, Sun Yat-sen University, No. 135, Xingang Xi Road, Guangzhou, 510275, P. R. China.

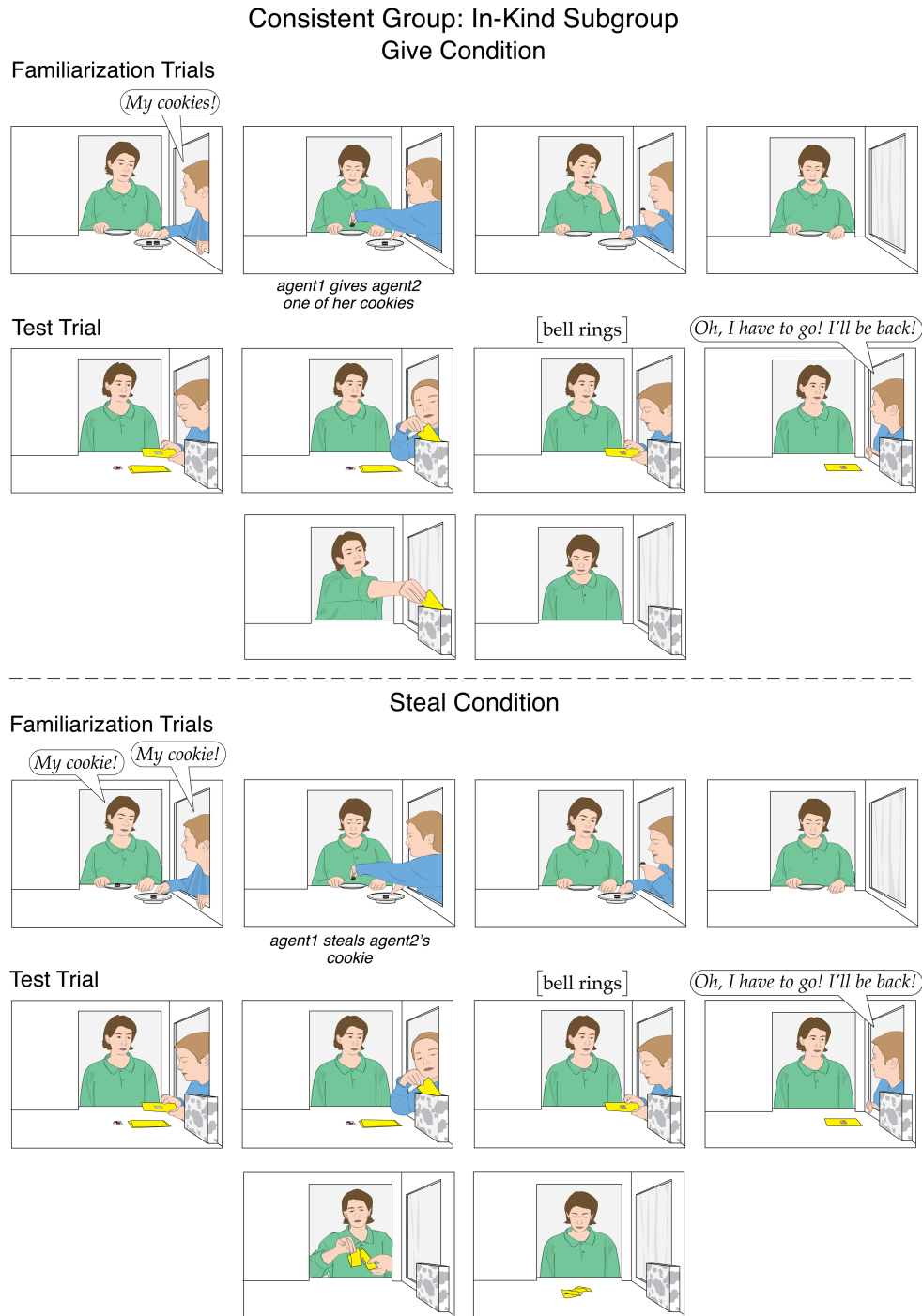
⁵Department of Psychology, University of Illinois Urbana-Champaign, 603 East Daniel Street, Champaign, IL 61820, United States.

Contents of the Supplementary Information:

- Supplementary Table 1, which reports infants' mean looking times in the familiarization and test trials of Experiments 1–3.
- Supplementary Figures 1 and 2, which depict the events shown to infants in Experiment 3.
- Supplementary Methods, which report overall analyses of the test data obtained with infants in Experiments 2 and 3.
- Supplementary Note 1, which describes two online surveys conducted with adults and intended as manipulation checks for the events shown to infants in Experiments 1–3. The Note includes Supplementary Figure 3, which presents the results of the second survey, on adults' ratings of reciprocity scenarios adapted from the events shown to infants.
- Supplementary References.

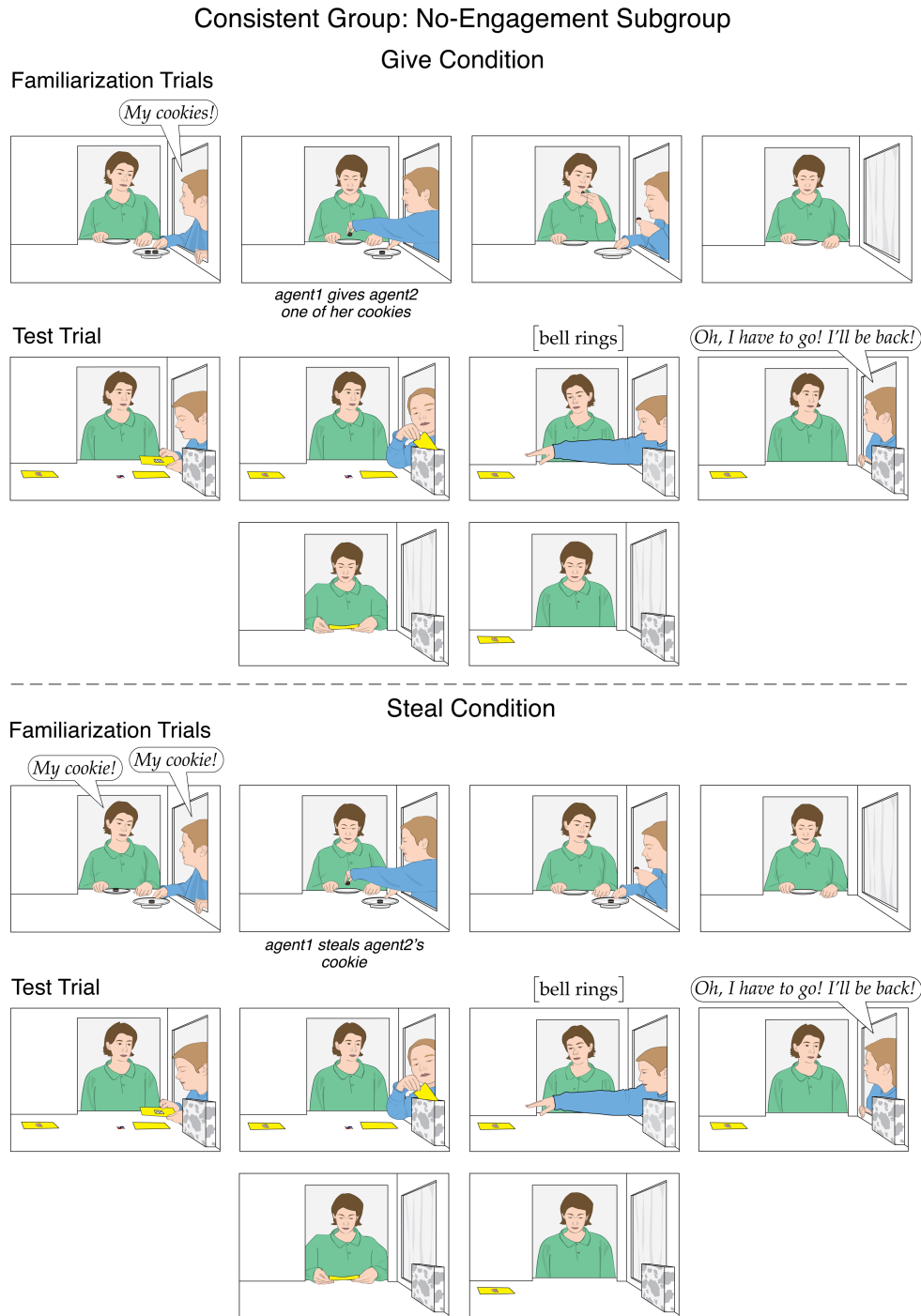
Supplementary Table 1. Mean looking times (and standard deviations) during the familiarization and test trials of Experiments 1–3, by group, subgroup, and condition. Subgroups within the control and consistent groups are listed in italics.

	<i>N</i>	Familiarization Trials			Test Trial		
		Give Condition	Steal Condition	Both Conditions	Give Condition	Steal Condition	Both Conditions
Experiment 1: Baseline Group	40 (20 per condition)	14.48 (4.83)	14.28 (5.41)	14.38 (5.06)	.	.	.
Experiment 2: Inconsistent Group	40 (20 per condition)	13.12 (4.48)	14.96 (6.62)	14.04 (5.66)	19.08 (7.79)	19.74 (7.30)	19.41 (7.46)
Experiment 2: Control Group	40 (20 per condition)	14.06 (7.14)	12.89 (5.83)	13.48 (6.46)	12.04 (5.42)	12.42 (5.64)	12.23 (5.46)
<i>New-Agent Subgroup</i>	20 (10 per condition)	15.18 (7.82)	13.88 (6.91)	14.53 (7.21)	12.15 (5.06)	11.79 (6.19)	11.97 (5.51)
<i>No-Knowledge Subgroup</i>	20 (10 per condition)	12.94 (6.61)	11.90 (4.66)	12.42 (5.59)	11.92 (6.03)	13.04 (5.28)	12.48 (5.55)
Experiment 3: Consistent Group	40 (20 per condition)	11.87 (3.47)	12.33 (5.68)	12.10 (4.65)	10.18 (4.52)	11.37 (4.97)	10.78 (4.73)
<i>In-Kind Subgroup</i>	20 (10 per condition)	11.41 (4.47)	14.52 (6.50)	12.97 (5.66)	10.02 (5.98)	12.18 (5.29)	11.10 (5.61)
<i>No-Engagement Subgroup</i>	20 (10 per condition)	12.32 (2.24)	10.14 (3.89)	11.23 (3.28)	10.34 (2.69)	10.56 (4.76)	10.45 (3.77)



Supplementary Fig. 1. Familiarization and test trials presented in the give and steal conditions of the in-kind subgroup of the consistent group in Experiment 3. Infants in each condition received the same three familiarization trials as in the give and steal conditions of the baseline group in Experiment 1, respectively, followed by a single test trial. In this trial, agent2 chose to respond in kind to agent1's actions. In the give condition, agent2 helpfully stored the decorated paper agent1 had deposited on the apparatus floor when she was called away by the bell; and in the steal condition, agent2 tore up the paper and dropped the pieces onto the apparatus floor. Infants in the two

conditions thus saw the opposite actions from those shown in the give condition (tearing up the paper) and steal condition (storing the paper) of the inconsistent group in Experiment 2. In each condition, the actions shown in the first two squares of the test trial were repeated as agent1 decorated and stored two papers. The test trial was preceded by a brief pretest trial (not shown) in which agent1 sat alone (agent2's window was closed) and prepared and stored two additional papers.



Supplementary Fig. 2. Familiarization and test trials presented in the give and steal conditions of the no-engagement subgroup of the consistent group in Experiment 3. Infants in each condition received the same three familiarization trials as in the give and steal conditions of the baseline group in Experiment 1, respectively, followed by a single test trial. In this trial, agent1 first decorated and stored two papers, as usual. However, her third and last decorated paper, already prepared, now lay across the apparatus, out of her reach but within agent2's reach; agent1 tried unsuccessfully to grasp it until she was called away by the bell. After she left, in both the give and steal

conditions, agent2 picked up the paper, looked at it briefly, and then returned it to its original position on the apparatus floor, out of agent1's reach. Agent2 thus opted not to engage with agent1 further: not to help her in the give condition, and not to harm her in the steal condition. In each condition, the actions shown in the first two squares of the test trial were repeated as agent1 decorated and stored two papers. The test trial was preceded by a brief pretest trial (not shown) in which agent1 sat alone (agent2's window was closed) and prepared and stored two additional papers.

Supplementary Methods

In a final overall analysis ($N = 120$), we compared the looking times of infants in the give and steal conditions of the inconsistent (Experiment 2), control (Experiment 2), and consistent (Experiment 3) groups during the final phase of the test trial. To this end, we conducted an ANOVA with group (inconsistent, control, consistent) and condition (give, steal) as between-subjects factors. The only significant effect was the main effect of group, $F(2, 114) = 23.270$, $p = .000$, $\eta_p^2 = 0.290$, 90% CI = [.171, .384].

We followed up this main effect with six planned comparisons, using a Bonferroni-adjusted α -level of $.05/6 = 0.008$. First, overall contrasts confirmed that infants in the inconsistent group looked significantly longer than those in the control and consistent groups, $F(1, 114) = 45.395$, $p = .000$, $\eta_p^2 = 0.285$, 90% CI = [.173, .385], who did not differ significantly, $F(1, 114) = 1.146$, $p = .287$, $\eta_p^2 = 0.010$, 90% CI = [.000, .060]. Next, we conducted the same two contrasts within each condition, with similar results. Thus, in the give condition ($N = 60$), the inconsistent group looked significantly longer than the other two groups, $F(1, 114) = 23.058$, $p = .000$, $\eta_p^2 = 0.168$, 90% CI = [.075, .268], who did not differ significantly, $F(1, 114) = 0.937$, $p = .335$, $\eta_p^2 = 0.008$, 90% CI = [.000, .056]. Similarly, in the steal condition ($N = 60$), the inconsistent group looked significantly longer than the other two groups, $F(1, 114) = 22.340$, $p = .000$, $\eta_p^2 = 0.164$, 90% CI = [.072, .263], who did not differ significantly, $F(1, 114) = 0.297$, $p = .587$, $\eta_p^2 = 0.003$, 90% CI = [.000, .039].

As discussed in the main text, these results support several conclusions. First, in each group, infants in the give and steal conditions showed the same response pattern, indicating that, in this situation at least, infants were equally adept at reasoning about direct positive and negative reci-

procity. Second, infants in the inconsistent group detected a reciprocity violation in agent2's actions: After agent1 gave cookies to agent2, infants viewed it as unacceptable for agent2 to harm her; similarly, after agent1 stole agent2's cookies, infants viewed it as unacceptable for agent2 to help her. Third, infants in the control group detected no violation in agent2's actions; even though these actions were the same as in the inconsistent group, they were no longer knowingly directed at agent1, so that considerations of reciprocity did not apply. Lastly, infants in the give and steal conditions of the consistent group also detected no violation in agent2's actions, suggesting that they viewed the test outcomes they were shown (with agent2 choosing to respond in kind to agent1 or to not engage with her further) as broadly consistent with reciprocity.

Supplementary Note 1

When describing the actions we showed infants in Experiments 1–3, we made claims about the value of individual actions (e.g., moderately positive or negative in value) and about the acceptability of reciprocal actions (e.g., inconsistent or consistent with reciprocity). As manipulation checks for these descriptions, we tested adults in two surveys created using the online survey tool Qualtrics (Provo, UT). The first survey asked adults to rate the value of individual actions similar to those produced by agent1 and agent2 in our infant experiments. The second survey asked adults to rate the plausibility of various actions produced by agent2 in response to positive or negative actions by agent1, again using scenarios similar to those in our infant experiments. We speculated that obtaining ratings in line with our descriptions and with infants' responses would support our interpretations of these responses.

As an important caveat, we stress that our surveys were not intended to offer a thorough, in-depth investigation of adults' reasoning about reciprocity; for one thing, the child-friendly scenarios we used were hardly appropriate for this purpose. As noted above, our surveys were intended solely to provide manipulation checks for our descriptions of the events we showed infants. For this reason, we limited our analyses to ones designed to identify overall patterns in adults' ratings.

As mentioned in the Methods of the main manuscript, the data generated in our online surveys with adults have been deposited on the Open Science Framework under the name “Jin et al. (2024) adult dataset”¹ and are publicly available at <https://doi.org/10.17605/OSF.IO/FDQS2>.

Value survey

Our first survey asked adults to rate the value (valence and magnitude) of actions similar

to those we showed infants in our experiments; we told participants that we were preparing scenarios for a study with young children. Each scenario described a brief interaction between two strangers who happened to be sitting near each other in a waiting room. In each scenario, one stranger produced a particular action toward the other; across scenarios, different letters were used to refer to the two strangers (e.g., A and B, C and D, and so on), but for ease of description we use agent1 and agent2 here. There were eight scenarios in total. Two involved cookie stories: In one, agent1 gave a cookie to agent2, who had none; in the other, agent1 and agent2 each had a cookie, and agent1 stole agent2's cookie. Three scenarios involved sticker stories: While agent1 watched, agent2 affixed stickers to a page in a sticker book; one sticker fell out of her reach under agent1's chair, just before agent2 was called away; after she left, leaving her things behind, agent1 picked up the sticker and chose to help (she placed the sticker near the sticker book), ignore (she looked at the sticker then dropped it back under her chair), or harm (she tore up the sticker then dropped the pieces back under her chair). The final three scenarios involved coloring stories: Agent1 and agent2 were both coloring in their coloring books, and one of them was missing a green crayon to finish the page she was working on; agent1 again chose to help (she gave her green crayon to agent2, who lacked one), ignore (she did nothing to help agent2), or harm (she herself lacked a green crayon and stole that of agent2 after she was called away). Scenarios were presented in random order. After reading a scenario, participants rated agent1's action on a 1-7 scale ranging from very negative to very positive, with 4 as a neutral point (neither negative nor positive). Ratings for the help, ignore, and harm scenarios were averaged across the sticker and coloring stories.

Participants were 115 adults (46 self-reported as male, 63 as female, and 6 as other; mean age = 26.41 years, range = 18–35 years). They were recruited through Prolific and paid \$1.25 for their participation. Another 4 participants were tested but excluded because the difference in their

mean ratings of the moderately positive and moderately negative scenarios was over 3 standard deviations from the group mean. Each participant provided informed consent prior to completing the survey, and the protocol was approved by the Institutional Review Board at the University of Illinois Urbana-Champaign (We did not conduct a power analysis to determine how many adults to include in each of our two surveys. In a recent report² on infant sociomoral reasoning where we also used an online survey with adults as a manipulation check, we tested 110 adults; in line with this number, we chose to include at least 100 participants in each survey, and ended up with 115 in the first and 130 in the second).

The give-cookie story was rated as moderately positive ($M = 6.63$, $SD = 0.61$), as were the help scenarios ($M = 6.30$, $SD = 0.84$). Conversely, the steal-cookie ($M = 1.20$, $SD = 0.44$) and harm scenarios ($M = 1.54$, $SD = 0.54$) were rated as moderately negative. Finally, the ignore scenarios were rated as mildly negative ($M = 3.23$, $SD = 0.64$). Participants were highly consistent in their rankings, with 95% giving higher ratings to (a) the help vs. the ignore scenarios and (b) the ignore vs. the harm scenarios. Finally, 98% of participants gave ratings above 4 (i.e., in the positive range) to the help scenarios, 87% gave ratings below 4 (i.e., in the negative range) to the ignore scenarios, and 100% did the same for the harm scenarios. These results dovetailed well with our descriptions of the valences and magnitudes of the actions shown to infants in our experiments.

Reciprocity survey

Our second survey built on the first and presented a new group of adults with scenarios depicting a sequence of two brief interactions between two strangers. As before, different letters were used to refer to the two strangers in each scenario (e.g., A and B, C and D, and so on), but here we use agent1 and agent2 for ease of description. In each scenario, agent1 first either gave a cookie to agent2 or stole her cookie. Next, agent2 produced the help, ignore, or harm actions from

the sticker and coloring stories used in the first survey.

Adults were presented with 12 scenarios in total, in four blocks of three scenarios. The first block always involved the give-sticker scenarios, as pilot data suggested that seeing give scenarios first helped anchor adults' ratings. The other three blocks (the give-crayon, steal-sticker, and steal-crayon scenarios) were presented in random order. Within each block, the help, ignore, and harm scenarios were presented in random order. After reading each scenario, adults were asked to rate the plausibility of agent2's action, in the context of agent1's action, on a 1-7 scale ranging from very plausible to very implausible (reverse-scored), with 4 as a neutral point (neither plausible nor implausible); reciprocity was never mentioned. As before, ratings were averaged across the sticker and coloring stories.

Participants were 130 adults (63 self-reported as male, 61 as female, and 6 as other; mean age = 27.19 years, range = 18–35 years). They were recruited through Prolific and paid \$1.25 for their participation. Another 47 participants were tested but excluded: 26 failed to understand or comply with the survey's instructions and rated agent2's action without taking account of agent1's action, 13 completed the survey very quickly, and 8 gave identical answers throughout the survey. Each participant provided informed consent, and the protocol was approved by the Institutional Review Board at the University of Illinois Urbana-Champaign.

To facilitate comparisons between our infant and adult data, we organized the scenarios presented to adults in two conditions (give, steal) and three groups: inconsistent scenarios (give/harm, steal/help), consistent in-kind scenarios (give/help, steal/harm), and consistent no-engagement scenarios (give/ignore, steal/ignore). Next, we analyzed adults' ratings using an ANOVA with group (inconsistent, consistent in-kind, consistent no-engagement) and condition

(give, steal) as within-subjects factors (see Supplementary Fig. 3). This analysis yielded a significant main effect of group, $F(2, 258) = 334.088, p = .000, \eta_p^2 = 0.721, 90\% \text{ CI} = [.676, .754]$, a significant main effect of condition, $F(1, 129) = 8.652, p = .004, \eta_p^2 = 0.063, 90\% \text{ CI} = [.012, .139]$, and a significant Group $\times$ Condition interaction, $F(2, 258) = 43.632, p = .000, \eta_p^2 = 0.253, 90\% \text{ CI} = [.177, .319]$.

To follow up on these results, we conducted eight contrasts, using a Bonferroni-adjusted α -level of $.05/8 = 0.006$. The first six contrasts indicated that (a) overall, adults rated the inconsistent scenarios as significantly more implausible than either the consistent in-kind scenarios, $F(1, 129) = 478.173, p = .000, \eta_p^2 = 0.788, \text{ CI} = [.735, .822]$, or the consistent no-engagement scenarios, $F(1, 129) = 259.903, p = .000, \eta_p^2 = 0.668, \text{ CI} = [.591, .722]$; (b) in the give condition alone, the inconsistent scenarios were again rated as significantly more implausible than either the consistent in-kind scenarios, $F(1, 129) = 658.896, p = .000, \eta_p^2 = 0.836, \text{ CI} = [.795, .863]$, or the consistent no-engagement scenarios, $F(1, 129) = 179.124, p = .000, \eta_p^2 = 0.581, \text{ CI} = [.490, .647]$; and (c) the same pattern also held in the steal condition alone, with the inconsistent scenarios receiving significantly higher ratings than either the consistent in-kind scenarios, $F(1, 129) = 94.585, p = .000, \eta_p^2 = 0.423, \text{ CI} = [.316, .508]$, or the consistent no-engagement scenarios, $F(1, 129) = 135.167, p = .000, \eta_p^2 = 0.512, \text{ CI} = [.412, .587]$.

The last two contrasts followed up on the significant Group $\times$ Condition interaction found in the ANOVA. Inspection of Supplementary Fig. 3 suggested that adults responded somewhat differently to the two types of consistent scenarios in the give and steal conditions. Indeed, in the give condition, the no-engagement (give/ignore) scenarios received significantly higher ratings than the in-kind (give/help) scenarios, $F(1, 129) = 239.270, p = .000, \eta_p^2 = 0.650, \text{ CI} = [.569, .706]$.

In the steal condition, however, the no-engagement (steal/ignore) and in-kind (steal/harm) scenarios tended to receive similar ratings, $F(1, 129) = 1.105, p = .295, \eta_p^2 = 0.008, CI = [.000, .053]$. Supporting this last result, a BF test comparing the in-kind and no-engagement scenarios in the steal condition yielded a BF of 5.98 in favor of the null hypothesis. Thus, whereas adults tended to view the steal/harm and steal/ignore scenarios as equally plausible, they viewed the give/ignore scenarios as somewhat less plausible than the give/help scenarios. Nevertheless, one-sample t -tests confirmed that adults rated both the give/ignore scenarios, $t(129) = -7.193, p = .000$, Cohen's $d = -0.631$, 95% $CI = [-.819, -.442]$, and the give/help scenarios, $t(129) = -42.577, p = .000, d = -3.734$, $CI = [-4.213, -3.246]$, as significantly below 4, the neutral point on the scale, and thus as broadly plausible.

This last result could be said to echo prior findings from a study examining direct positive reciprocity in adults following a small, unnecessary gift from a stranger³. In this study, each participant was paired with an unfamiliar same-sex confederate, and the two were asked to independently rate how much they liked art slides presented using a projector. Halfway through the slides, the participant and the confederate were given a break. During the break, the confederate briefly left the room, returned with two packages of M&Ms ostensibly from a nearby vending machine, and offered one to the participant (“I thought you might want some too”). Later, the participant and the confederate were separated to complete additional questionnaires in privacy. After a brief period, the participant was told that the confederate had had to leave to get to work and had asked the experimenter to give the participant an envelope with “some kind of charity thing to fill out”. The envelope contained a pledge form requesting a donation for a charity, the Run for the Kids Foundation, and participants could choose to pledge money or not. When they

could do so privately, without the confederate finding out what they chose, about 50% of participants pledged money; this number decreased to about 30% when no candy was offered. These results seem consistent with the notion that for adults as for infants, a small, unnecessary gift by a stranger may create only a weak obligation to reciprocate.

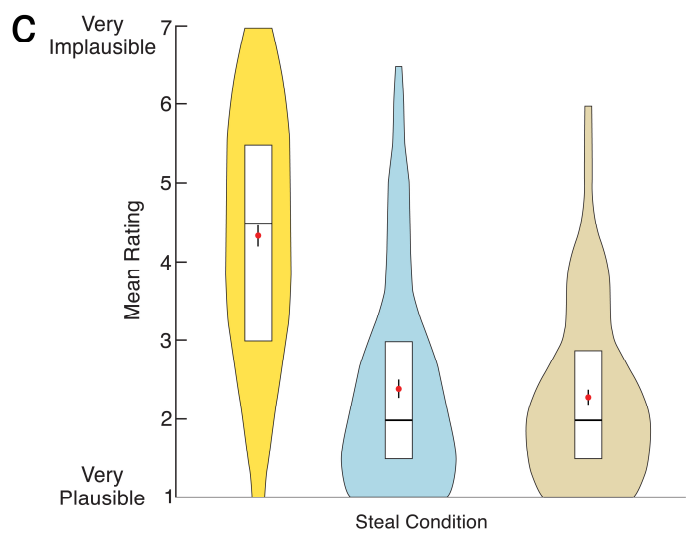
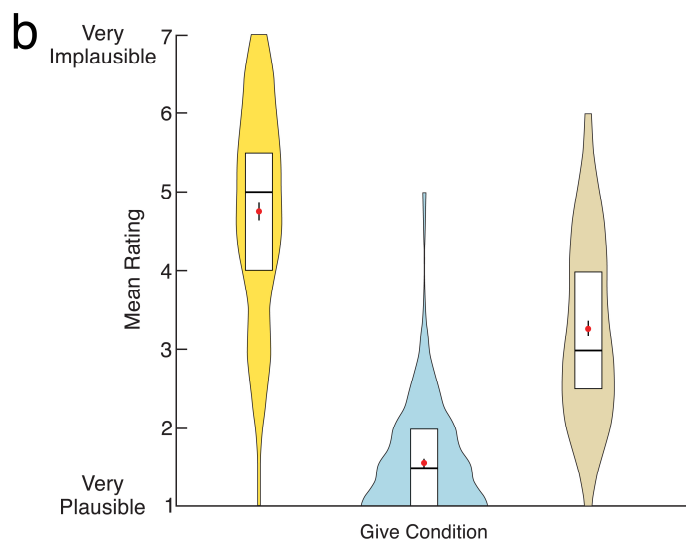
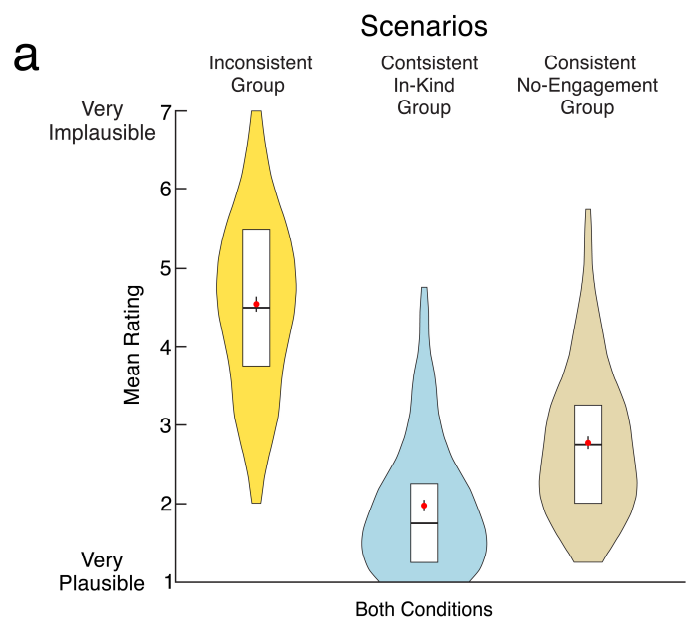
Discussion

If we weave together the results of our value and reciprocity surveys, the following patterns arise. First, after agent1 gave agent2 a cookie (a moderately positive action), adults viewed it as implausible for agent2 to harm agent1 (a moderately negative action), but they viewed it as plausible for agent2 to help her (a moderately positive action) and, to a lesser extent, to ignore her need for help (a mildly negative action). Second, after agent1 stole agent2's cookie (a moderately negative action), adults viewed it as implausible for agent2 to help agent1 (a moderately positive action), but they viewed it as equally plausible for agent2 to harm her (a moderately negative action) or to ignore her need for help (a mildly negative action). Thus, in each condition, adults did not rate only reciprocal actions that exactly matched agent1's initial action in valence and magnitude as plausible: They considered a range of actions by agent2 as broadly plausible.

Finally, if we compare our results with adults to those we obtained with infants in Experiment 2 (inconsistent group) and Experiment 3 (in-kind and no-engagement subgroups of the consistent group), clear parallels emerge. Adults rated the inconsistent scenarios (give/harm, steal/help) as more implausible than the consistent in-kind (give/help, steal/harm) and no-engagement (give/ignore, harm/ignore) scenarios, presumably because they viewed the inconsistent scenarios as violating reciprocity. Moreover, this finding held true in the give and steal conditions separately. Adults' responses thus echoed those of infants, supporting our interpretations of infants' looking-time patterns. The main difference between the two age groups involved the give/ignore scenarios,

and it is open to at least two interpretations. One is that this difference reflects a qualitative change with age and experience in what actions are and are not acceptable in response to a friendly overture that one has no desire to pursue. It could be that for infants, only moderate negative actions are unacceptable in this situation, whereas for adults, mild negative actions are also less than ideal, even if not entirely unacceptable. Another, related interpretation is that in deciding whether a mild negative action is acceptable or not in that situation, adults take into account the costs associated with different courses of action. In the give/ignore scenarios, the cost to agent2 of helping agent1 was minimal. The actions needed—picking up agent1’s fallen sticker and moving it closer to her sticker book, or loaning her a green crayon—were so trivial and required so little effort that to adults, agent2’s decision not to produce them bordered on churlishness. In this view, had helping agent1 required significantly more cost or effort (e.g., pledging money to a charity for her, giving her a ride across town), adults might have seen it as more acceptable for agent2 not to do so.

That difference aside, adults’ and infants’ responses were markedly similar overall, hinting at underlying similarities in their intuitions about direct positive and negative reciprocity between strangers. Nevertheless, as already noted, caution is needed in drawing this conclusion as our scenarios were specifically designed to study infants’ rather than adults’ intuitions about reciprocity.



Supplementary Fig. 3. Violin plots showing adults' mean plausibility ratings. **a** Overall mean ratings for the inconsistent scenarios (yellow plot), the consistent in-kind scenarios (blue), and the consistent no-engagement scenarios (beige). **b** Mean ratings for the same groups of scenarios in the give condition alone. **c** Mean ratings for the same groups of scenarios in the steal condition alone. The same adults ($N = 130$) rated all scenarios. Dots represent means, and error bars represent standard errors. In each violin plot, the width of the shaded area represents the proportion of responses observed at each value, smoothed by a kernel density estimator. Overlaid on each violin plot is a boxplot ranging from the 25th to the 75th percentile, with a line drawn at the median.

Supplementary References

1. Jin, K., Ting, F., He, Z. & Baillargeon, R. Infants expect some degree of positive and negative reciprocity between strangers. Open Science Framework, <https://doi.org/10.17605/OSF.IO/FDQS2> (2024).
2. Bian, L. & Baillargeon, R. When are similar individuals a group? Early reasoning about similarity and in-group support. *Psychol. Sci.* **33**, 752-764 (2022).
3. Whatley, M. A., Webster, J. M., Smith, R. H. & Rhodes, A. The effect of a favor on public and private compliance: How internalized is the norm of reciprocity? *Basic Appl. Soc. Psychol.* **21**, 251-259 (1999).