nature portfolio

Corresponding author(s): Kyong-sun Jin

Last updated by author(s): Nov 15, 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	Infants in Experiments 1-3 were shown live events in a puppet-stage apparatus; these events were acted out by experimenters using second-by-second scripts. Hidden observers monitored infants' looking behavior; they pressed a button linked to a computer when infants attended to the events presented. Custom-made computer software (which the Baillargeon lab has been using for many years) recorded the observers' responses and determined when a trial had ended. The criteria used by the computer to end trials are described in the Methods section. As manipulation checks for the events we showed infants, we conducted two online Qualtrics surveys with adults, using written scenarios describing similar events. The results of these surveys are reported in the Supplementary Online Material (SOM).
Data analysis	We used SPSS 26 for all frequentist statistical analyses and JASP (version 0.18.1) for all Bayesian tests. Confidence intervals for effect sizes were calculated using an online calculator.

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

All of the infant data reported in this article can be found at: https://osf.io/fdqs2/?view_only=16453d82dcb14e55b85e38c45f631643.

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	As described in the Methods section, participants in Experiments 1-3 were 160 infants, 83 male and 77 female. Participants in the online surveys (described in the SOM) were 245 adults, 126 male, 107 female, and 12 other.
Reporting on race, ethnicity, or other socially relevant groupings	As described in the Methods section, participants in Experiments 1-3 were term 15-month-old infants (M = 15 months, 22 days; range = 15 months, 3 days to 16 months, 16 days). Infants were from English-speaking families residing in or near Urbana and Champaign, Illinois, two adjacent small towns in the Midwestern United States. Demographic information was not collected, but the infants came from a range of socioeconomic backgrounds, and most were European American. We chose to test infants because we wanted to examine whether infants already possess an expectation of reciprocity. Participants in the online surveys (described in the SOM) were English-speaking adults residing in the Unites States (M = 26.8 years; range = 18–35 years). Demographic information was not collected.
Population characteristics	See above.
Recruitment	Infants were recruited from a Department-maintained database of parents living in or near Urbana and Champaign, Illinois, who are interested in participating in developmental research. All parents in this database who were contacted and agreed to participate in our research did so. Adults were recruited through Prolific.
Ethics oversight	Institutional Review Board of the University of Illinois at Urbana-Champaign.

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	Our experiments used the violation-of-expectation method to examine whether 15-month-old infants already possess an expectation of reciprocity. Infants watched live interactions between two women who appeared to be strangers to one another, agent1 and agent2. We tested whether initial positive or negative actions by agent1 toward agent2 would influence infants' expectations about how agent2 might subsequently act toward agent1. In particular, would infants be surprised (as indexed by longer looking times) if agent2 chose to harm agent1 after being the beneficiary of her positive actions, or chose to help her after being the victim of her negative actions? As manipulation checks for the events we showed infants, we conducted two online surveys with adults, using written scenarios describing similar events. These surveys are described in the SOM. All of our data are quantitative.
Research sample	As described above, participants in Experiments 1-3 were 160 term infants from English-speaking families residing in or near Urbana and Champaign, Illinois, two adjacent small towns in the Midwestern United States (83 males and 77 females; M = 15 months, 22 days; range = 15 months, 3 days to 16 months, 16 days). Infants were tested because our research goal was to examine whether infants already possess an expectation of reciprocity.

Participants in the online surveys (described in the SOM) were 245 English-speaking adults residing in the United States (126 male, 107 female, and 12 other; M = 26.8 years; range = 18–35 years).

Sampling strategy

Across Experiments 1-3, infants were randomly assigned to one of four groups (baseline, inconsistent, control, and consistent). As explained in the Methods section, a power analysis using G*Power based on previous experiments on sociomoral reasoning in infants suggested that the minimal number of participants per group was 36 infants. In line with this analysis, we tested 40 infants per group, for a total of 160 infants.

We did not carry out a formal power analysis for the online adult surveys, which were intended merely as manipulation checks. Instead, we chose to include at least 100 adults per survey, and ended with 115 adults in the first survey and 130 in the second. These numbers are similar to other recent manipulation checks we have conducted using adult surveys (see SOM).

Data collection

As described in the text and Methods of our article, each infant sat on a parent's lap in front of a large puppet-stage apparatus. To avoid influencing infants' responses, parents were instructed to remain silent and neutral throughout the testing session, and to close their eyes for the test trial in Experiments 2 and 3 (infants in Experiment 1 received only familiarization trials).

At the start of each trial, a hidden supervisor lifted a curtain at the front of the apparatus; the infant then watched live events acted out by two experimenters using second-by-second scripts; a metronome beat softly to help the experimenters adhere to the scripts. As events unfolded, the experimenters never made eye contact with the infant: They only looked at each other, at the objects they acted on, or at a neutral point on the apparatus floor.

During each testing session, one camera captured an image of the events, and another camera captured an image of the infant; the two images were combined, projected onto a monitor located behind the apparatus, and watched by the supervisor to confirm that the events followed the prescribed scripts. The images were also recorded, and recorded sessions were checked off-line for experimenter and observer accuracy.

In the pretest and test trials of Experiments 2 and 3, two observers (hidden behind panels on either side of the apparatus) monitored the infant's looking behavior. Each observer pressed a button linked to a computer when the infant attended to the events presented. As noted above, the computer used long-established custom-made software to record the observers' responses and to determine when a trial had ended. The looking times recorded by the primary observer were used in the data analyses of these trials. In the familiarization trials that preceded the pretest and test trials, the primary observer was absent from the testing room, to ensure that the observer was blind to the infant's condition. In Experiment 1, an offline observer coded each infant's recorded testing session from silent video, with the portion of the computer screen showing the events hidden from view, to ensure that the observer was blind to the infant's condition.

In our adult surveys (described in the SOM), participants completed an online Qualtrics survey.

Timing

Data collection for Experiments 1-3 (infants) began in 2008 and ended in 2020. Data collection for the online surveys with adults began and ended in 2022.

Data exclusions

As described in the Methods section, in addition to the 160 infants included in Experiments 1-3, another 25 infants were tested but excluded: 16 were distracted (6), fussy (5), inattentive (3), or drowsy (2); 1 was the subject of parental interference; 1 was uninterested and crawled away; and 7 (3 in the control group and 4 in the consistent group) had a test looking time over 2.5 standard deviations from the group mean. This percentage of exclusion (25/185, or 14%) is well within the norms for infancy research, and the criteria we used for exclusion are typical for infancy research and correspond to those we routinely use in our experiments.

In our first survey with adults (Value survey, N = 115), another 4 participants were tested but excluded because the difference in their mean ratings of the moderately positive and negative scenarios was over 3 standard deviations from the group mean. In our second survey (Reciprocity survey, N = 130), 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.

Non-participation

No parent asked to terminate their infant's testing session. However, a few testing sessions ended, as noted above, because the infant became fussy or did not want to watch the events. All adults who initiated the survey completed it.

Randomization

Infants in each experiment were randomly assigned to the groups, subgroups, and/or conditions tested in the experiment, with the restriction that the numbers of male and female infants were kept approximately equal in each cell.

As explained in the SOM, adults participating in the first, value survey read eight scenarios presented in random order. Adults participating in the second, reciprocity survey read 12 scenarios presented in four blocks of three scenarios. The first block was always fixed and the other three blocks were presented in random order; within each block, the three scenarios were presented in random order.

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

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.