**1) Data collection.** Have any data been collected for this study already?

* Yes, we already collected the data.
* No, no data have been collected for this study yet.
* It's complicated. We have already collected some data but explain in Question 8 why readers may consider this a valid pre-registration nevertheless.

(Note: "Yes" is not an accepted answer.)

**2) Hypothesis**. What's the main question being asked or hypothesis being tested in this study?

In a previous study, we investigated the effects of social cognition on probabilistic reasoning in the Tsimane’, a farming-foraging group native to the Bolivian Amazon. Specifically, how does our knowledge of agents influence how we reason about the likelihood of events? We showed that participants underestimated the likelihood of events that appeared to be more “structured” (e.g., sampling a group of 7 blue and 9 yellow balls was judged to be more likely than sampling a group of 8 blue and 8 yellow balls). This study is a direct replication with U.S. participants.

Participants will be presented with a box with blue and yellow balls in equal proportions. Then, they will watch a brief video showing an agent drawing balls with their eyes closed (“agency”) condition, or an agent pouring balls from the box (“physical” condition; see Q4). Participants will first complete five warm-up trials (presented in a fixed order), followed by 15 test trials (presented in a random order). In each trial, participants will be shown two groups of balls and asked which of the two is more likely to have been sampled blindly (“agency” condition) or which of the two is more likely to appear first if an agent repeatedly poured the balls from the box (“physical” condition).

The warm-up trials are similar in style to the test trials, except that they emphasize a greater likelihood difference between the two groups (e.g., 7 blue and 9 yellow balls vs. 16 blue and 0 yellow balls). See Q6 for details on exclusions as a function of warm-up trial performance.

We predict that participants’ answers will correlate with the responses we obtained with the Tsimane’: their choices will track the relative probability of drawing each sample, with the exception of trials where at least one of the groups has “structure” that suggests that an agent generated it. Specifically, in cases where it’s more likely to sample a particular configuration of balls (e.g., 8 blue and 8 yellow balls vs. 7 blue and 9 yellow balls), we expect participants to sway away from the more structured option despite it being more statistically probable. We also predict that structure will have a larger effect on larger sample sizes (e.g., a group of 8 blue and 8 yellow balls will look less random than a group of 2 blue and 2 yellow balls).

**3) Dependent variable**. Describe the key dependent variable(s) specifying how they will be measured.

The dependent measure will be the group of balls that participants choose in each test trial as the group that is more likely to have been sampled blindly (“agency” condition; see Q4) or which of the two is more likely to appear first if an agent repeatedly poured the balls from the box (“physical” condition).

**4) Conditions.** How many and which conditions will participants be assigned to?

Participants will be divided into an “agency” and a “physical” condition. These conditions determine the sampling process that participants are introduced to in the training video. Participants in the “agency” condition will see an agent sample one group of balls with their eyes closed (1 blue ball and 3 yellow balls, 3 blue balls and 1 yellow ball, or 3 blue balls and 3 yellow balls; counterbalanced across participants) and arrange them in a line by color before putting them back into the box. During the test trials, participants will be told that the agent sampled one of these groups with their eyes open and the other with their eyes closed and will be asked which one of these groups they think the agent sampled with their eyes closed. Participants in the “physical” condition will see an agent pouring a group of balls onto the tray (4 blue balls and 3 yellow balls), putting them back into the box, shaking it, pouring another group of balls onto the tray (5 blue balls and 7 yellow balls), putting them back into the box, and shaking it again. During the test trials, participants will be told that the agent will continue this process for a long time and they will be asked which one of these groups they think would appear first if the agent continued this process. Participants in both conditions will see the same 15 test trials.

Participants will also be randomly split between two counterbalancing conditions. Participants in the “yellow” condition will only see groups that have the blue balls on the top side of the tray and the yellow balls on the bottom side. Participants in the “blue” condition will only see groups that have the yellow balls on the top side of the tray and the blue balls on the bottom side. Results from these two conditions will be collapsed and analyzed together.

**5) Analyses.** Specify exactly which analyses you will conduct to examine the main question/hypothesis.

Our first analysis will consist of correlating participant responses across the “agency” and “physical” conditions to see if participants give similar answers in the two conditions. We predict that participant responses will be roughly equal to or greater than the correlation among Tsimane’ participants ( and in the “blue” and “yellow” counterbalancing conditions, respectively; see Q4) due to increased comfort of U.S. participants in experiments.

Our second analysis will consist of correlating participant responses with the responses predicted by a several simple models that select groups proportional to their probability of being sampled. The first model computes the exact probability of sampling each group according to a binomial distribution. The second model is similar to the first, but will double the probability of uneven groups to account for people treating the color of the majority as irrelevant (e.g., people treating 7 blue and 9 yellow balls as belonging to the same class as 9 blue and 7 yellow balls). The third model will compute a probability from a two-tailed binomial test to account for people treating the group and more extreme samples as equivalent (e.g., people treating 2 blue and 6 yellow balls as belonging to the same class as 1 blue and 7 yellow balls and 0 blue and 8 yellow balls). We predict that these models will make incorrect predictions when the less likely sample has more structure.

Our final analysis will consist of correlating U.S. participant responses with the responses from our Tsimane’ participants.

We will also compute the percentage of choices in each trial along with 95% bootstrapped confidence intervals. In addition, the distribution produced by bootstrapping responses will be used to compute the probability that the effect size is in a certain range.

**6) Outliers and Exclusions.** Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

To make sure the participants understand the task, we employ five warm-up trials. The warm-up trials are similar in style to the test trials, except that they emphasize a greater likelihood difference between the two groups (e.g., 7 blue and 9 yellow balls vs. 16 blue and 0 yellow balls). Whenever a participant fails a warm-up trial, we will re-enact the agent in the video and re-explain the task, and the warm-up trial stage will be restarted from the beginning. If participants fail any of the warm-up trials a second time, they will be excluded and replaced.

**7) Sample Size.** How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.

We will recruit 40 participants for the “agency” condition and the “physical” condition for a total of 80 participants (see Q4). Within each condition, participants will be split on the counterbalancing “blue” and “yellow” condition ( participants in each counterbalancing condition). This sample size was determined by a previous experiment with the Tsimane’. Participants that are excluded (see Q6) will be replaced so that the final sample size is 80 participants.

**8) Other.** Anything else you would like to pre-register?

(e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

See training videos, warm-up trials, and test trials at:

<https://osf.io/kfxrm/?view_only=ddec69255f80473cb51a4199d2a2d08b>

**9) Name.** Give a title for this AsPredicted pre-registration.

Suggestion: use the name of the project, followed by study description.

Image Inference – Goal Inference