Simulation-Based Inference Worksheet

PSTAT 5LS - Spring 2025

Investigation: Friend or Foe?[1]

In a study reported in the November 2007 issue of *Nature*, researchers investigated whether infants take into account an individual's actions towards others in evaluating that individual as appealing or aversive, perhaps laying for the foundation for social interaction (Hamlin, Wynn, and Bloom, 2007[2]). In other words, do children who aren't even yet talking still form impressions as to someone's friendliness based on their actions?

In one component of the study, 10-month-old infants were shown a "climber" character (a piece of wood with "googly" eyes glued onto it) that could not make it up a hill in two tries. Then the infants were shown two scenarios for the climber's next try, one where the climber was pushed to the top of the hill by another character (the "helper" toy) and one where the climber was pushed back down the hill by another character (the "hinderer" toy).

Helper Scenario	Hinderer Scenario

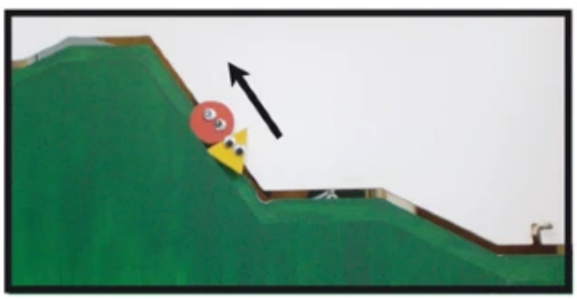
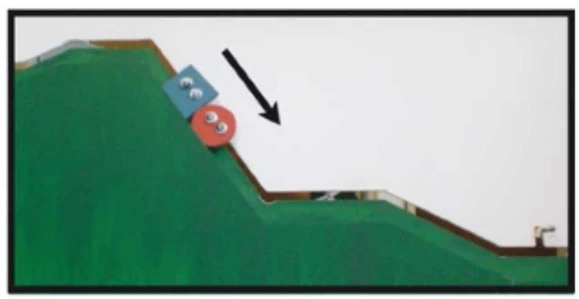

	

Figure 1.a from Hamlin, Wynn, and Bloom (2007)	
helper video	hinderer video

The infant was alternately shown these two scenarios several times. Then the child was presented with both pieces of wood (the helper and the hinderer characters) and asked to pick one to play with. More videos demonstrating this component of the study can be found at http://campuspress.yale.edu/infantlab/media/.

[1] Based on Investigation 1.1 from Beth Chance and Allan Rossman's *Investigating Statistical Concepts, Applications, and Methods* (4e)

[2] Hamlin, J., Wynn, K. & Bloom, P. Social evaluation by preverbal infants. *Nature* 450, 557–559 (2007). https://doi.org/10.1038/nature06288

Collecting the Data

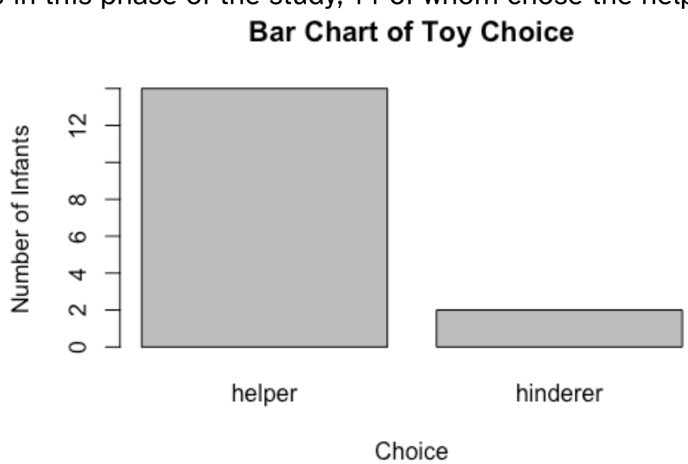
A. Identify the observational units in this study.

B. What is the variable we are measuring about each infant? Is it a quantitative variable or a categorical variable?

One important design consideration to keep in mind is that in order to equalize potential influencing factors such as shape, color, and position, the researchers varied the colors and shapes of the wooden characters and even on which side the toys were presented to the infants.

Summarizing the Observed Data

There were 16 infants in this phase of the study, 14 of whom chose the helper toy.



Drawing Conclusions Beyond the Sample

The descriptive analysis above tells us about this particular sample of 16 infants. Do these results convince us that infants in general are more likely to pick the helper toy? That is, is there more than a 50% chance that a 10-month-old infant picks the helper toy? Or, was there a 50% chance that a 10-month-old infant would pick the helper toy and we observed these results by random chance alone?

We have two explanations for the majority we observed:

- Infants choose equally between the two toys in the long run and we happened to get “lucky” and had an unusual sample where most of the infants in our sample picked the helper toy.
- There is something to the theory that infants are genuinely more likely to pick the helper toy (for some reason).

C. So for the two possibilities we are still considering, how might you choose between them? In particular, how might you convince someone whether or not option (2) is plausible based on this study?

Our analysis approach is going to be to assume the first explanation is true (similar to how in a legal trial we assume a defendant is innocent), and then see whether our data are consistent or inconsistent with that assumption. To do this, we need to investigate the values we expect to see for the number choosing the helper toy when 16 infants are equally choosing between the two toys. We can simulate the outcomes of a random process to help us determine which outcomes are more or less likely to occur.

Modeling the Situation

For a 50-50 simulation model, we can flip a fair coin. We can arbitrarily define “heads” to be choosing the helper toy and “tails” to be choosing the hinderer toy. We will repeat the random process 16 times to represent the 16 infants, and we will count how many times heads come up, representing an infant choosing the helper toy. Complete the table below to show the correspondence between the real world, which we saw one instance of, and the simulation model, which we can easily repeat many times. Keep in mind that in the simulation model, we know the probability of heads is 0.50.

	Real World	**Simulation Model**
Observational unit		
Variable		
“Success”		
Probability of “success”		
Sample size		

Coin Tossing Model

We already have a sense that 14 heads in 16 tosses would be unusual, but to really investigate the probability, we need to run the simulation many more times. This would be very tedious and time-consuming with coins, so let's turn to technology.

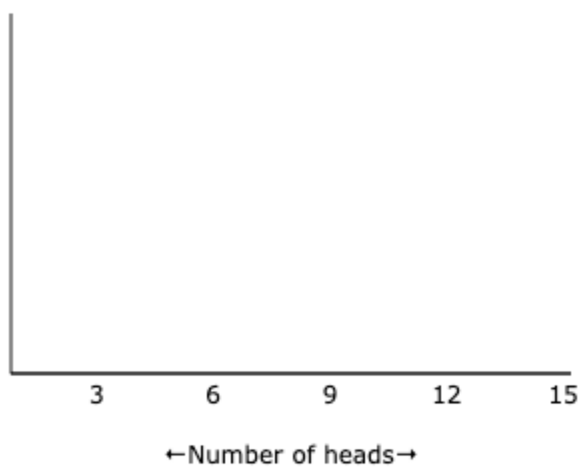
Navigate to https://www.rossmanchance.com/applets/2021/oneprop/OneProp.htm.

- Keep the Probability of heads set to 0.5.
- Set the Number of Tosses to 16 to match the number of infants in the study.
- Keep the Number of repetitions at 1 for now.
- Press Draw Samples. How many heads did you get?

- Uncheck the "Show animation" box and press Draw Samples four more times, do you get the same number of heads each time?

- Now change the Number of repetitions to 995 and press Draw Samples, to produce a total of 1000 repetitions of this random process of tossing a coin 16 times.

D. Sketch your results below.



E. Where is the distribution of the simulation results centered? Briefly explain why this makes sense.

Measuring "Rareness"

The distribution you created, although beautiful, by itself really tells us nothing. We need to use it to evaluate the actual observed study result (14 of 16). We could report the proportion of times we observed 14 heads in the 1000 repetitions, but as we have larger sample sizes (number of coin tosses), any individual outcome (number of heads) will be increasingly unlikely. So to give us a more “standard” number, that we can compare across studies, we will look at the proportion of results “as or more” extreme. In this study, “more extreme” means even more of the 16 tosses landed heads.

F. Go back to the applet and enter 14 in the “As extreme as” box and click on the “Count” button. How many of your 1000 repetitions had 14 or more heads? Convert this count to a decimal. This is the **estimated p-value**.

Formulate Conclusions

G. If you were publishing these results in a research article, would you say the data provide convincing evidence that infants evaluate the behavior of others? Or, do you think it is more reasonable to say that infants have no preference and the study result was just a random fluke? Explain your reasoning.