Sociology Study: Naughty or Nice?
We all recognize the difference between naughty and nice, right? What about children less than a year old—do they recognize the difference and show a preference for nice over naughty? 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 the foundation for social interaction (Hamlin, Wynn, and Bloom, 2007). In one component of the study, sixteen 10-month-old infants were shown a “climber” character (a piece of wood with “google” eyes glued onto it) that could not make it up a hill in two tries. Then they 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 (“helper”) and one where the climber was pushed back down the hill by another character (“hinderer”). 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) and asked to pick one to play with. The color and shape and order (left/right) of the toys were varied and balanced out among the 16 infants.

(a) In order for you to be reasonably convinced that infants in general are not just choosing blindly but genuinely choose the helper toy more often, how many of the 16 would have to choose the helper toy? Or would no outcome convince you? Explain your answer.

Suppose for the moment that the researchers’ conjecture is wrong, and infants do not really have any preference for either type of toy. In other words, infants just blindly pick one toy or the other, without any regard for whether it was the helper toy or the hinderer. Put another way, the infants’ selections are just like flipping a fair coin: choose the helper if the coin lands heads and the hinderer if the coin lands tails.

(b) If this is really the case (that infants show no preference between the helper and hinderer), what would be the most likely outcome (for number of infants choosing the helper toy) when this study is conducted on 16 infants?

(c) Still assuming that infants show no preference between the helper and hinderer, what kind of results (for number of infants choosing the helper toy) would you not be surprised to see when this study is conducted on 16 infants?
The researchers actually found that 14 of the 16 infants in the study selected the helper toy.

(d) Calculate the proportion of these infants who chose the helper toy. Is this more than half (a majority)?

(e) If it is really the case that infants show no preference between the helper and hinderer, which word do you believe best completes the sentence: It would be __________ for 14 out of 16 infants to choose the helper toy just by chance. (Circle your answer below.)

1) impossible
2) very surprising
3) somewhat surprising
4) not at all surprising

A key question is how to determine whether the observed result is surprising under the assumption that infants have no real preference. (We will call this assumption of no genuine preference the null model.) To answer this question, we will replicate the infants’ selection process over and over, assuming that infants have no genuine preference and were essentially flipping a coin in making their choices (i.e., knowing the null model to be true). In other words, we’ll simulate the process of 16 hypothetical infants making their selections, where we know those selections are by random chance (coin flip), and we’ll see how many of them choose the helper toy. Then we’ll do this again and again, over and over. Every time we’ll see how many of the 16 infants choose the helper. Once we’ve repeated this process a large number of times, we’ll have a pretty good sense for whether 14 of 16 is very surprising, or somewhat surprising, or not so surprising under the null model.

**Hands-on simulation analysis:**
Now the practical question is, how do we simulate this random selection (with no genuine preference)? One answer is to go back to the coin flipping analogy. Let’s literally flip a coin for each of the 16 hypothetical infants: heads will mean to choose the helper, tails to choose the hinderer.

(f) Flip a coin 16 times. Count how many of your 16 hypothetical infants chose the helper toy (represented by heads):

(g) Combine your results with your classmates. Do this by producing a dotplot (of the number of infants who choose the helper toy) on the board, where you contribute a dot corresponding to your simulation result from (f). How many simulated experiments altogether are represented in the resulting plot (i.e., how many students put a dot on the board)?
(h) Granted, we have not conducted a large number of simulated experiments, but how is it going so far? Does it seem like the results actually obtained by these researchers would be surprising under the null model that infants do not have a genuine preference for either toy? Explain.

Let’s make this question of surprising-ness more specific by quantifying how often the experimental result would occur.

(i) In how many of you and your classmates’ simulated experiments did 14 or more infants choose the helper toy? In other words, in how many of the simulated experiments did randomness alone produce a result at least as extreme as the researchers found in the actual study? What proportion of the simulated experiments is this?

*Computer simulation analysis:*

We can use the computer to simulate this random process of 16 infants making this helper/hinderer choice, still assuming the null model that infants have no real preference and so are equally likely to choose either toy. The computer enables you to conduct many more repetitions much more efficiently.

(j) Open the applet at: http://statweb.calpoly.edu/bchance/applets/BinomDist3/BinomDist.html. Make sure that the probability of heads is set to .5 (as our null model for the infants stipulates) and the number of tosses is 16 (corresponding to the number of infants in the study). Keep the number of repetitions at 1 for now, and click on “16 Tosses” to simulate 16 coin tosses, determining the number of heads and adding this result to the dotplot on the right. Then repeat this four more times (five total), noting how the number of heads in 16 tosses varies from repetition to repetition. Then change the number of repetitions to 995 (which will bring the total number of repetitions to 1000) and click on “16 Tosses” again.

Now enter 14 in the “as extreme as $\geq$” box and press the “Count” button. Below the button, the applet will report the proportion of your 1000 repetitions that produced 14 or more infants choosing the helper toy. Record this number here.
(k) Based on this simulation analysis which assumes the null model that infants have no preference and so choose blindly, the actual result obtained by the researchers (14 of 16 infants choosing the helper toy) is _______ (circle one),

1) impossible
2) very surprising
3) somewhat surprising
4) not at all surprising

**Terminology**: The long-run proportion of times that an event happens when its random process is repeatedly indefinitely is called the **probability** of the event. We can approximate a probability empirically by simulating the random process a large number of times and determining the proportion of times that the event happens.

More specifically, the probability that randomness would produce data as (or more) extreme as an actual study, assuming the null model to be true, is called a **p-value**. Our analysis above approximated this p-value by simulating the random process a large number of times (under the null model) and finding how often we obtained results at least as extreme as the actual data. You can obtain better and better approximations of this p-value by using more and more repetitions in your simulation.

A small p-value indicates that the observed data would be surprising to occur by randomness alone, if the null model were true. Such a result is said to be **statistically significant**, providing evidence against the null model.

(l) Click on the applet’s “Exact values” button, and report the exact (to three decimal places) p-value.

Did your simulation analyses produce a close approximation to this exact p-value?

(m) Based on this simulation analysis, what conclusion should the researchers draw: does their experimental result provide strong evidence that infants in general are not just choosing blindly but genuinely choose the helper toy more often? Explain the reasoning process behind your conclusion, being sure to explain the role of the simulation analysis, as if to a classmate who is not in class today.

Conclusion:

Reasoning process:
(n) If the actual study had instead found that 9 of the 16 infants chose the helper toy, then what
decision should the researchers make based on this result? Explain the reasoning process behind
your answer.

Conclusion:

Reasoning process: