Selecting Evidence to Limit Hypotheses

Alexandra Kincannon (kincannon@virginia.edu)
Barbara A. Spellman (spellman@virginia.edu)
Department of Psychology; University of Virginia
Charlottesville, VA 22903 USA

Introduction

When people want to *generalize* a hypothesis, they recognize the importance of getting evidence from different sources. This preference, the "diversity principle", is both assumed as normative by philosophers of science and is a strategy used by participants in scientific reasoning tasks (Popper, 1962; Spellman, López, & Smith, 1999). To *limit* a hypothesis, however, it seems that evidence from similar sources would be more informative. Using an example from our first experiment, "Suppose you know for a fact that elephants have an ostic vesicle. What organism would you examine to test whether or not ONLY mammals have an ostic vesicle?" The response choices were hippopotamus (inside category and similar), fox (inside category and dissimilar), crocodile (outside category and similar), and snake (outside category and dissimilar). It seems to us that the best answer is crocodile—the non-mammal that is the most similar to mammals. Most people agree that no information about hippos or foxes is relevant to the question of whether only mammals have this property, but the choice of crocodile over snake is less obvious. Crocodiles are more similar to elephants than snakes are, and perhaps the property in question (ostic vesicle) has something to do with that similarity (e.g., four-leggedness). Thus, although finding that crocodiles or snakes do have the property is equally conclusory (i.e., that the property is not limited to only mammals), finding that crocodiles don't have the property is more informative than finding that snakes don't have the property (because if snakes don't have it, crocodiles still might).

Experiments

In the first experiment, we tested participants' ability to limit hypotheses using stimuli like that from the elephant example above. Similarity was established by the results of an earlier experiment in which 80 participants made pairwise ratings on a 6-point scale. As expected, participants were more likely to test an animal outside the category of the target than inside the category, $X^2(1,N=63)=21.7$, $p<0.001$ for the elephant question. Contrary to our prediction, the dissimilar animal was chosen more often than the similar one, $X^2(1,N=50)=6.5$, $p=0.01$ for the elephant question. Similar results were found with the other target (robin).

In the next experiment, the test questions had a premise/conclusion format like that used by Osherson, et al. (1990). Participants saw one first premise (Robins have a condyloid canal) and one conclusion (Therefore ONLY birds have a condyloid canal) and two different second premises (Bats DO NOT have a condyloid canal. / Gorillas DO NOT have a condyloid canal.). Participants chose which second premise gave stronger support to the conclusion. Here, they tended to choose the similar animal premise (bats) more than the dissimilar animal premise (gorillas); however, this preference was not significant. When we compared across experiments, a 2x2 contingency table revealed a statistically significant difference between the frequencies of similar versus dissimilar animal choices depending on the test question format, $X^2(1,N=128)=10.24$, $p=0.001$ for the robin question. The frequency of choosing the outside-category similar animal, compared to the frequency of choosing the outside-category dissimilar animal, was greater in the premise choice task than in the evidence choice task.

Discussion

We have found that some conditions seem to promote better reasoning in hypothesis limitation tasks; what needs more exploration is the process behind that reasoning. In ongoing experiments, we are delving further into how people select evidence for limiting hypotheses. Participants are conducting multiple tests of their hypothesis, given several animals in several different categories to choose from. One factor that could play a role in our participants' behavior is the desire for a quick confirmation of their hypothesis (as in Wason, 1960). Perhaps people start by picking the most dissimilar animal, but if given the option to continue testing, they may see the value of working their way towards the more similar animal on subsequent tests. Under many real-life circumstances, such a strategy might be viewed as quite rational.

References

Osherson, D. N., Smith, E. E., Wilkie, O., López, A., & Shafir, E. (1990). Category-based induction. *Psychological Review, 97,* 185-200.

Popper, K. R. (1962). *Conjectures and refutations: The growth of scientific knowledge.* New York: Basic Books.

Spellman, B. A., López, A., & Smith, E. E. (1999). Hypothesis testing: Strategy selection for generalising versus limiting hypotheses. *Thinking & Reasoning, 5,* 67-91.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *The Quarterly Journal of Experimental Psychology, 12,* 129-140.