The Effects of Mechanistic and Functional Explanations on Categorization

Tania Lombrozo (lombrozo@berkeley.edu)
Department of Psychology, UC Berkeley, 3210 Tolman Hall, Berkeley, CA 94720 USA

Keywords: explanation; categorization; functions; causal status

Explanation and Conceptual Structure

In a 1985 paper, Murphy and Medin suggested that concepts are coherent by virtue of the theories in which they're embedded, where theories are “any of a host of mental ‘explanations’” (Murphy & Medin, 1985). The idea that explanation and categorization are intimately related has since received direct empirical support: explanatory knowledge facilitates category learning, influences typicality judgments, and makes for more ‘coherent’ categories (see Lombrozo, 2006 for review).

Recent work also suggests that causal beliefs influence categorization. One proposal, the “causal status hypothesis,” is that causally ‘deep’ features are more important than their effects for category membership. For example, if members typically have F1 and F2, and F1 causes F2, then an item with F1 but not F2 is more likely to be judged a member than an item with F2 but not F1 (Ahn & Kim, 2000). While this effect and its interpretation are controversial, it provides a basis for exploring the role of explanation in mediating the influence of causal beliefs on categorization. In the experiments below, I find a relationship between explanatory construal—that is, whether a category’s features are explained mechanistically (by appeal to proximate causes) or teleologically (by appeal to functions)—and categorization judgments.

Exp 1: Explanations Correlate with Categorization

Ninety-six undergraduates completed the study in exchange for course credit. They were taught about a novel category (biological kind or artifact) with features F1 and F2 (e.g. eats blueberries, has blue fur), where F1 causes F2 and F2 subserves a function (e.g. camouflage). Participants were then asked to explain why category members have F2, and to estimate the probability that items missing one feature (F1 or F2) were category members. These probabilities were subtracted to yield a difference score with positive differences reflecting a causal status effect, and negative differences its reversal. The prediction is that participants who provide a mechanistic explanation in response to the “why?” prompt (“category members have F2 because they have F1”) should show a greater causal status effect than those who provide a teleological explanation (“category members have F2 in order to bring about the function”). As predicted, participants who provided a mechanistic explanation had significantly higher difference scores than participants who provided a teleological explanation (main effect of condition, F(1, 92) = 12.09, p < .05; main effect of domain n.s., interaction n.s., see Fig. 1A). This finding demonstrates a relationship between explanation and categorization, but can explanatory construal causally influence categorization?

Exp 2: Explanations Influence Categorization

One-hundred-ninety-two undergraduates completed the study in exchange for course credit; half each in the mechanism and function conditions. The task was identical to Experiment 1, but participants were not told the function of F2, and those in the function condition were prompted to generate a function by answering whether F2 might have a purpose. This manipulation successfully increased teleological explanations from 1% in the mechanism condition to 81% in the function condition. And as predicted, participants in the mechanism condition exhibited a stronger causal status effect than those in the function condition (main effect of condition, F(1,188) = 8.27, p < .05; main effect of domain n.s., interaction n.s., see Fig. 1B), suggesting that explanatory construal is not merely correlated with categorization, but causally affects categorization.

Conclusions

These findings demonstrate that explanations influence categorization. Because participants were provided with the same information across comparisons, the effects cannot be reduced to differences in the causal relationships between features. Rather, the findings suggest that explanations influence the interpretation and use of causal information to determine which relationships are conceptually central.

References

