An Instantiation Model of Category Typicality and Instability

Evan Heit
Department of Psychology
University of Michigan
330 Packard Road
Ann Arbor, MI 48104
Evan.Heit@um.cc.umich.edu

Lawrence W. Barsalou
Department of Psychology
University of Chicago
Chicago, IL 60637
L-Barsalou@uchicago.edu

Abstract

According to the instantiation principle, when we make a judgment about a relatively superordinate category, we follow a two-step process. First, we instantiate the category into one or more subordinates. Second, we make a judgment based on the subordinates. Instantiation theory applied to typicality judgments makes the following predictions. When subjects judge the typicality of a category A with respect to category B, their mean typicality judgment should equal the weighted mean typicality (with respect to B) of subordinate categories of A. Furthermore, typicality judgments for category A will be unstable (i.e., have a high standard deviation) to the extent that A has a large number of diverse subordinates. The instantiation principle was implemented in a computer simulation, which used production frequencies and typicality ratings for subordinates to predict ratings for superordinate-level categories. In two experiments, subjects judged the typicalities of various animal and food categories. The instantiation model successfully predicted the means and standard deviations for the observed distributions of responses for these categories. Extensions and other applications of the instantiation principle are also briefly discussed.

Introduction

Vertical Category Structure and Instantiation

Semantic categories may be described and related to each other in terms of their taxonomic, or vertical, structure (Rosch, 1978). Some categories, such as mammal and beverage, are relatively superordinate; they are at a high level of abstraction. Other categories, such as dog and milk, are subordinate; they are more specific than their respective superordinates.[1] The rationale behind the instantiation principle is that it is hard to think about relatively superordinate categories. Members of superordinate categories often do not share many features and or even look like each other. A superordinate category may be quite abstract or ambiguous. On the other hand, reasoning about more subordinate categories ought to be easier because members of these categories are more alike. The instantiation principle says that when we make a judgment about a relatively superordinate category (e.g., mammal), we perform this task by first instantiating the superordinate into one or more subordinate categories (e.g., dog and human), then second, making judgments about the subordinates.

There is some prior research bearing on instantiation. Contextual information has been shown to lead people to perform instantiation during reading comprehension. For example, after reading the sentence "The fruit was made into wine," the word "grape" will serve as a better retrieval cue for this sentence than the word "fruit" (Anderson, Pichert, Goetz, Schallert, & Stevens, 1976). The instantiation principle has also been successfully applied to reasoning. Osherson, Stern, Wilkie, Stob, and Smith (1991) proposed that when people evaluate the strength of an inductive argument such as

[1] The terms subordinate and superordinate are used here solely to describe taxonomic position *relative to each other*. These terms are not intended to describe taxonomic position relative to basic-level categories (as in Rosch, 1978).

"Canines have sesamoid bones; therefore mammals have sesamoid bones," they first instantiate the superordinate categories, canine and mammal. Osherson, et al. tested this proposal by finding subjects' subordinates of various superordinate categories. In the present example, they found that the argument's strength could be predicted from the similarity between instantiations of canine (e.g., dog and wolf) and instantiations of mammal (e.g., dog and human).

Typicality Structure and Instability

In addition, it has been found universally that categories have a typicality structure: some category members are judged to be better members than others. The typicality variable is an excellent predictor of how people perform other categorization tasks, such as the time to classify exemplars and the ease of learning category members. Interestingly, typicality structure does not remain fixed across contexts, between people, or even within a single person at different times (Barsalou, 1987). Categories may differ not only in their mean typicalities but also in terms of the standard deviations, reflecting their instabilities. One goal of this study is to investigate whether and how different categories differ systematically in their degree of instability.

Instantiation Model of Typicality Judgments

Following is a simple version of the instantiation principle, applied to typicality judgments. When a person judges the typicality of a category A with respect to category B, they do two things. First, they produce one subordinate of A, subA. Second, they judge the typicality of subA with respect to B. So, someone judging the typicality of mammal with respect to animal, might instead judge the typicality of dog with respect to animal.

More complex variants of this model are possible, with additional assumptions about how or how many instantiations are produced. Also, this model does not make any claims about how the typicality of subA is judged; it might involve the same instantiation process, or it might involve a simpler, non-recursive process. This model only specifies the relation between typicality judgments on A and typicality judgments on subordinates of A. Nonetheless, this simple instantiation model is worth testing, to see what it reveals about category structure.

What does the instantiation principle predict about typicality judgments? In general, people's responses when they judge the typicality of A with respect to B should resemble people's responses when they judge the typicality of instantiations of A with respect to the same B. If mammals are typical animals, then particular mammals, such as dogs and humans, should also be typical animals. Imagine that a group of subjects makes typicality judgments for A and for instantiations of A. (A set of instantiations of A may be obtained by asking another group of subjects to each name a single subcategory of A.) The two distributions of typicality judgments, for A and for its instantiations, should be alike. In particular, the respective means and standard deviations of the two distributions should be the same.

Some superordinate categories may be particularly ambiguous, leading people to make highly variable, unstable judgments using the different possible instantiations. According to the instantiation principle, instability of typicality judgments has a few sources. First, the typicality of A will be more unstable as the typicality ratings for the subordinates of A are more unstable. However, even if A has fairly stable subordinates, a second source of instability may be the diversity of its subordinates. If it has two subordinates that are very different from each other, then typicality judgments about A would show a lot of variation. In general, the instantiation principle predicts that typicality judgments about categories with more subordinates will be more unstable, unless these subordinates are all alike in terms of how typical they are.

Experiment 1

Testing the instantiation principle for typicality judgments was accomplished with two groups of subjects. The Production group produced subordinates for a set of relatively superordinate categories (e.g., mammal, reptile). The Rating group made typicality judgments about the superordinate and subordinate categories, always with respect to the category "animal."

Method

Subjects. The Production and Rating groups each consisted of 20 Stanford University undergraduates, recruited in dormitories. The Production subjects were all run before the Rating subjects.

Stimuli. For the Production group, the stimuli were 7 superordinate categories: amphibian, bird, fish, insect, mammal, microorganism, and reptile. The Production subjects produced 63 different subordinate categories in response to these superordinates. For the Rating group, the stimuli were 63 subcategories produced by the first group, plus the 7 superordinate

Table 1. Data for Mammal Category, Experiment 1.

Instantiation	Prod. Freq.	Typicality with Respect to Animal
Human	5	10, 8, 10, 10, 1, 5, 10, 1, 10, 7, 10, 10, 9, 6, 10, 8, 2, 9, 6, 2
Bear	3	10, 9, 10, 10, 6, 10, 10, 9, 10, 9, 4, 10, 7, 8, 10, 8, 9, 10, 10, 10
Kangaroo	3	8, 8, 10, 9, 4, 10, 10, 5, 2, 7, 4, 7, 7, 6, 5, 6, 5, 7, 8, 8
Whale	2	8, 7, 10, 7, 6, 10, 10, 1, 5, 4, 8, 10, 6, 8, 5, 6, 10, 5, 9, 4
Ape	1	10, 8, 10, 10, 7, 10, 10, 9, 10, 9, 8, 10, 7, 6, 8, 6, 9, 10, 7, 8
Cat	1	10, 8, 10, 9, 6, 10, 8, 10, 8, 9, 1, 6, 10, 7, 10, 10, 10, 10, 10
Cow	1	10, 9, 10, 9, 6, 10, 10, 7, 10, 8, 10, 9, 8, 9, 9, 9, 10, 10, 10, 10
Dog	1	10, 9, 10, 10, 10, 10, 10, 9, 10, 9, 10, 8, 8, 10, 9, 10, 10, 10, 10, 10
Elephant	1	10, 10, 10, 10, 5, 10, 10, 8, 10, 9, 7, 9, 7, 9, 7, 7, 10, 10, 10, 10
Horse	1	10, 9, 10, 10, 7, 10, 10, 10, 10, 8, 10, 9, 7, 10, 10, 10, 10, 10, 10, 10
Platypus	1	4, 6, 10, 6, 4, 10, 10, 3, 7, 7, 5, 8, 5, 4, 5, 8, 5, 7, 10, 3

Superordinate	Typicality with Respect to Animal
Mammal	9, 6, 10, 10, 1, 5, 10, 10, 10, 9, 10, 10, 9, 7, 10, 10, 9, 10, 10, 5

categories. Each subcategory appeared once as a Rating stimulus, regardless of the number of mentions by Production subjects.

Procedure. Production subjects were instructed to produce one subordinate for each of the 7 categories. Rating subjects were instructed to rate the 7 superordinate categories and the 63 subordinates on typicality with respect to "animal," using a 1-10 scale, for which higher numbers meant greater typicality. Sample questions would be "How typical is a mammal for the category animal?" and "How typical is a dog for the category animal?"

The Instantiation Model

The predicted distributions of typicality ratings for the superordinate categories relied on two sources of data: the production frequencies from the Production group and the subordinate typicality ratings from the Rating group. The predictions for each superordinate category were created by a computer simulation consisting of an iterated two-step process. First, a subordinate was chosen, according to the production frequencies for that superordinate. For example, 5 of 20 subjects produced "human" as an instantiation of "mammal," so "human" was used 25% of the time on this step, when predicting the distribution of responses for "mammal." Second, a typicality rating for the subordinate was chosen from the 20 typicality judgments that the Rating subjects made for this item. In this case, one of the 20 ratings that people made for "human" was chosen. This simulation did not operate stochastically, rather it exhaustively produced a distribution of 400 ratings for each predicted distribution of superordinate category responses. For each of the 20 productions of subordinates, each of its 20 corresponding typicality ratings were added to the predicted distribution. Thus, for each superordinate category, the predicted distribution represented 400 simulated subjects, with each simulated subject assumed to have made one production of a subordinate then one typicality rating on this subordinate. The observed distributions for the 7 superordinate categories were simply the 20 ratings made for each superordinate by the Rating subjects.

Table 1 illustrates how this model was applied to the data collected for the superordinate category "mammal." The 20 Production subjects responded with 11 distinct subordinates of this category, with production frequencies as shown in the table. Then the 20 Rating subjects judged the typicality of each subordinate, as well as the category "mammal," with respect to "animal." The 20 typicality ratings for

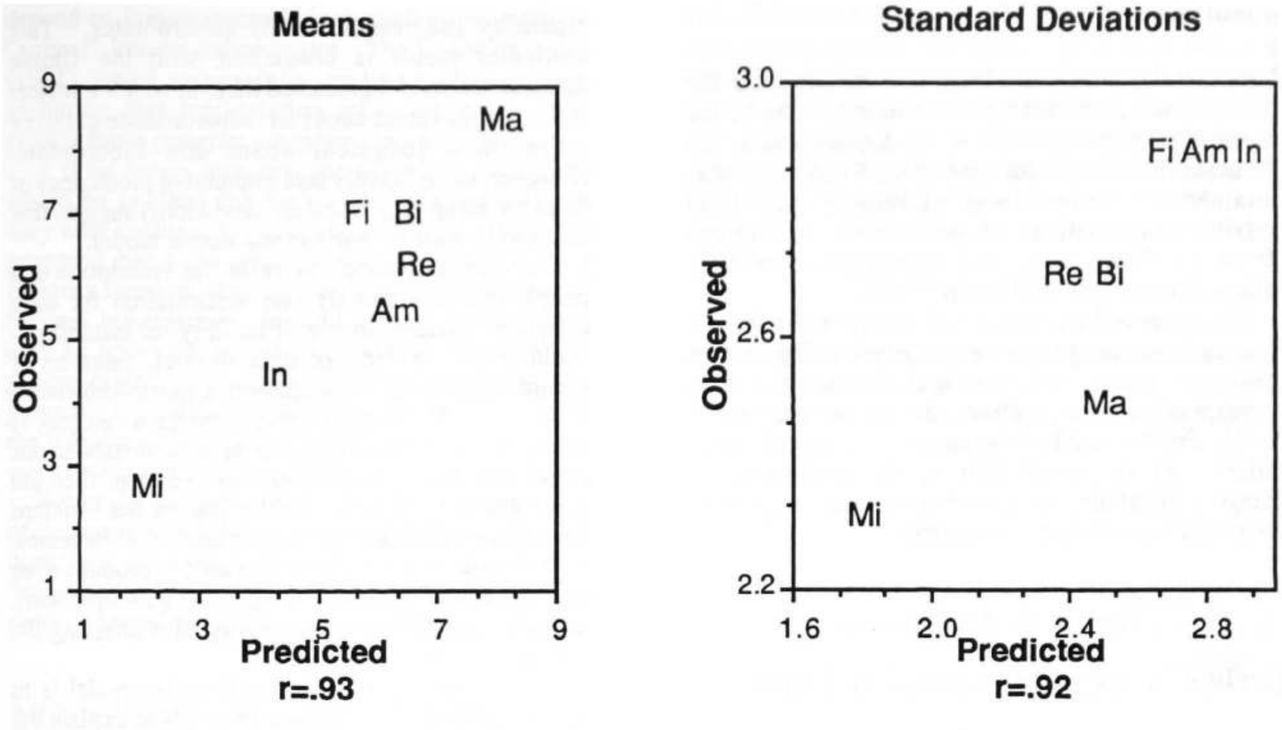

Figure 1. Statistical comparisons of observed and predicted distributions, Experiment 1.

each category are shown in the table. The observed distribution for "mammal" was simply the 20 typicality ratings for this item. The predicted distribution contained 400 ratings: each rating for "human" taken 5 times, each rating for "bear" and "kangaroo" taken 3 times, each rating for "whale" taken twice, and each rating for the other subordinates taken once.

Results

Of primary concern is whether the 7 observed distributions of typicality ratings for the superordinate categories correspond to the 7 predicted distributions, obtained by application of instantiation theory. The observed and predicted distributions are described in Figure 1, in terms of mean and standard deviation. (Each data point is shown as the first two letters of the superordinate category, e.g., Ma for mammal.) These categories did indeed differ in terms of their means and standard deviations. The correspondence between the observations and the predictions is excellent; $r=.93$ for the means and $r=.92$ for the standard deviations, one-tailed $p<.01$ in both cases. The predicted standard deviations generally underestimate the observed standard deviations, suggesting that additional sources contributed to instability of the ratings on the superordinates.

Experiment 2

To provide generality, this experiment was a replication of Experiment 1, with typicality judgments with respect to a different category (food) and more subjects.

Method

Subjects. The Production group consisted of 40 University of Michigan undergraduates, recruited in public places in Ann Arbor. The Rating group consisted of 40 Michigan undergraduates, who participated as part of a course requirement.

Stimuli. For the Production group, the stimuli were 9 superordinate categories of food: beverages, dairy products, desserts, fish, fruits, meats, poultry, seasonings, and vegetables. For the Rating group, the stimuli were the 9 superordinate categories plus 88 subordinates produced by the first group.

Procedure. The procedure was like Experiment 1, except Rating group subjects rated the categories on a scale from 1 to 9, in terms of typicality with respect to "food."

Results

Again, the primary analysis was to compare the observed and predicted typicality distributions for the superordinate categories. As in Experiment 1, the predicted distribution for each of the 9 categories was obtained by exhaustively simulating the 1600 possible combinations of subordinate productions (from the first group) and subordinate typicality ratings (from the second group).

The observed and predicted distributions for the superordinate categories are described in Figure 2, in terms of mean and standard deviation. The correspondence is excellent for the means, $r=.89$, $p<.01$. For the standard deviations, $r=.64$, $p<.05$ one-tailed. As in Experiment 1, the predictions of category instability are good but not quite as good as the predictions of mean typicality.

General Discussion

Evaluation of the Instantiation Model

These experiments demonstrate a constancy of categorical structure: judgments about a category tend to resemble judgments about its instantiations. Most strikingly, the mean typicality judgment of a superordinate was quite close to the weighted mean typicality judgments for its subordinates. This particular result is consistent with the simple instantiation model presented here, in which a subject making a judgment about an superordinate category relies on a judgment about one subordinate. However, the relatively less impressive predictions of category instability (standard deviations) suggest that extensions must be made to the simple model.

One extension would be relax the assumption that people evaluate exactly one instantiation for each category. Another source of stability--or instability--could be that subjects produce different numbers of subordinates for different categories, due to knowledge differences. If a subject's judgment for a category is based on the mean value for several instantiations rather than based on a single instantiation, then the judgment will be more stable, just as the standard error of a mean decreases as the sample size increases. And if some subjects find it difficult to produce even one subordinate of some category (e.g., amphibian), then they might respond randomly, also affecting the standard deviation of responses.

Another direction for extending this model is to further apply the instantiation principle to explain the typicality judgment process. The model, as presented, assumes that to judge how typical mammals are of the category "animal," a subject compares instantiations of "mammal" to the category "animal." This comparison process itself might

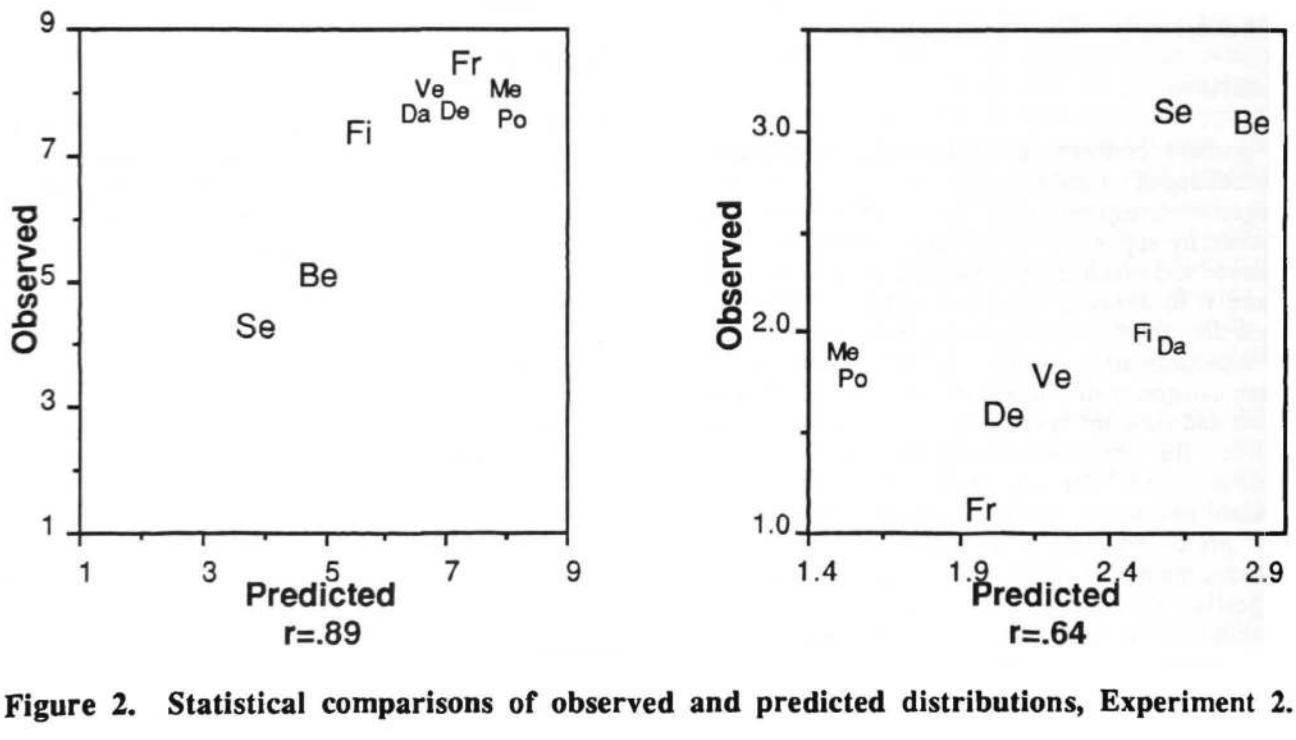

Figure 2. Statistical comparisons of observed and predicted distributions, Experiment 2.

depend on further instantiations, such as instantiating "animal" into its subordinates. The Osherson, et al. (1991) model successfully applied such a scheme, assuming that instantiation takes place for every superordinate category considered in a judgment.

A third future direction for extending this model would be to allow the basis of judging typicality to vary with context. In particular, Barsalou (1987) has proposed that when people make a judgment about a category member, they do not consider every piece of featural information. Instead, certain features may only be activated in certain contexts. For example, people might not usually consider cost when they judge the typicality of various foods, but the costs of the foods could indeed influence judgments if costs were made more relevant or salient.

While the preliminary tests of the simple version of the instantiation model were encouraging, the clear directions for extending this approach suggest that further experiments and more advanced models will be needed, as well as comparison to alternate approaches.

Applications of the Instantiation Principle

More generally, the instantiation principle sheds light on other results about instability of category structure. Barsalou (1987) identified three kinds of instability. Both within-subjects instability and between-subjects instability may be explained in terms of different instantiations being produced at different times. Recent experiences and differences in knowledge are likely to contribute to these differences in instantiations. Context-dependent instability (such as categories having different structures when they are considered from different perspectives or in different environments) can be partly explained in terms of different instantiations for different contexts. However, the typicality judgments on the instantiations could also reflect different knowledge about the different contexts, even if the same instantiations are made for both contexts. For example, someone might judge fish to be a typical "ocean animal" but an atypical "desert animal." For both judgments, "fish" might be instantiated the same way, say as "shark," but sharks would then be judged as typical for ocean animals and atypical for desert animals, based on knowledge about oceans and deserts.

Much more generally, the instantiation principle's successful applications to comprehension, inductive reasoning, and typicality judgment suggest further applications to other cognitive abilities. For example, to make a decision about whether to eat at restaurant X or restaurant Y, someone might compare specific meals that they have had at each restaurant (cf., Kahneman & Miller, 1986). This person would consistently choose restaurant X over restaurant Y if the meals at X are better overall than meals at Y, and the particular meals do not vary much. However, the decision-making process would be more unstable to the extent that the particular instances of restaurant X meals and restaurant Y meals were more variable.

Acknowledgments

We are grateful to Douglas Medin, David Thau, and Edward Wisniewski for comments on this paper, and to Rachael Jeffries and William Krooss for help in collecting the data.

References

Anderson, R. C., Pichert, J. W., Goetz, E. T., Schallert, D. L., Stevens, K. V., & Trollip, S. R. (1976). Instantiation of general terms. *Journal of Verbal Learning and Verbal Behavior, 15*, 667-679.

Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization*. Cambridge: Cambridge University Press.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93*, 136-153.

Osherson, D. N., Stern, J., Wilkie, O., Stob, M., & Smith, E. E. (1991). Default probability. *Cognitive Science, 15*, 251-269.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and Categorization*. New York: Wiley.