# Information, Uncertainty, and the Utility of Categories

Mark A. Gluck *Stanford University*
James E. Corter *Columbia University*

## *INTRODUCTION*

Do categories which encode certain types of regularities or invariances in the world have a special psychological status? Many studies have shown that some categories or groupings of instances are easier than others to learn and recall as coherent concepts or generalizations. For example, within a hierarchically nested set of categories (such as a taxonomy of animals), there is some level of abstraction--called the "basic level"--that is most natural for people to use (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). For example, in the hierarchy *animal-bird-robin*, *bird* is the basic level category. The preferrential status of basic level categories can be measured in a variety of ways. Basic level names are generally learned earlier by children (Rosch et al., 1976; Daehler, Lonardo, and Bukatko, 1979), and arise earlier in the development of languages (Berlin, Breedlove, & Raven, 1973). People tend to spontaneously name pictured objects at the basic level, and can name them faster at this level than at subordinate or superordinate levels (Rosch et al., 1976; Jolicoeur, Gluck, & Kosslyn, 1984).

### *Structural Explanations of the Optimality of Categories*

Recent findings suggest that the superiority of basic level categories is due to *structural* properties of the categories, that is, to the distribution of features across instances and non-instances (Murphy & Smith, 1982; Hoffmann & Ziessler, 1983). Previous attempts to characterize these structural regularities have focused on measures of the degree to which basic level categories capture regularities or invariances in the world. For example, it has been suggested that basic level categories maximize the average *cue validity* of features for a category (e.g. the conditional probability of a category given a feature) or the average *category validity* (e.g. the conditional probability of features given a category), or some combination of the two, such as the product of the cue validity and category validity (Rosch & Mervis, 1975; Medin, 1983; Jones, 1983). These measures of category structure, however, are purely extensional and ignore the contexts and needs of people who are creating and using concepts. Recent evidence suggests that conceptual structures are highly unstable and vary greatly with context (Murphy & Medin, 1985; Barsalou, 1983). We suggest here a *context-sensitive* measure of the utility of categorizations. Our primary goal is to develop a model of the structural basis of basic levels effects; we present evidence that this measure does predict the occurance of basic levels in several experiments. We are motivated in this effort by the belief that the existence of basic levels reflects an underlying cognitive heuristic that people use for organizing information about the world.

---

For their guidance and comments, we are indebted to Gordon Bower, Misha Pavel, W. K. Estes, Ed Smith, Doug Medin, Joachim Hoffmann, and Greg Murphy. The assistance of Katie Albiston and Audrey Weinland is also gratefuly acknowledged. Please address correspondence to: Mark A. Gluck, Department of Psychology, Stanford University; Bldg. 420, Stanford, CA 94305.

*The Informational Value of Categories*

We suggest that the degree to which certain concepts are favored over others may be related to how useful these concepts are for encoding and communicating information about the properties of things in the world. In other words, the most useful categories are those that are, on the average, optimal for communicating information (hence reducing uncertainty) about the properties of instances. We will show how to formalize this idea in situations where the relevant attributes are well defined.

We consider two specific definitions of uncertainty and show the implications of each for Category Utility. First, we utilize the standard definition of uncertainty from information theory (Shannon & Weaver, 1949), and show what it implies about Category Utility. Second, we consider a hypothetical communication game in which one person attempts to transmit information about an item's attributes to another person. Within this game, we interpret uncertainty as an inability to predict attributes, and analyze how category membership information can be used to transmit information about the attributes of objects or events.

We will describe our theory of Category Utility within the context of a finite population of items, each of which is describable in terms of a set of multi-valued nominal attribute dimensions. Each attribute dimension (e.g. *eye color*) is assumed to have a set of possible values (e.g. *green, brown, blue*), one of which occurs in every instance. A category of instances can be described by specifying the *distributions* of attribute values for instances in the category. For example, a specific category of faces may have 40% green eyes, 50% brown eyes, and 10% blue eyes.

In information theory (Shannon & Weaver, 1949), the *uncertainty* of a set, $F$, of $n$ messages (i.e. $F = f_1, f_2, \cdots, f_n$) is given by

$$U(F) = -\sum_{i=1}^{n} P(f_i)\log P(f_i).$$

We consider an attribute dimension to be a set of messages regarding the possible values of the attribute dimension. Consider also a partition, $C$, of a population of objects into two sets: those which are members of a category $c$ and those which are not. Given information that an item is a member of category $c$, the uncertainty of the values of attribute dimension $F$ will be:

$$U(F|c) = -\sum_{i=1}^{n} P(f_i|c)\log P(f_i|c),$$

where $P(f_i|c)$ is the conditional probability that a member of category $c$ has value $f_i$ on attribute dimension $F$. If instances of $c$ occcur with probability $p(c)$ and instances of *not–c* occur with probability $(1-p(c))$, then the expected reduction in uncertainty when one is told the category or not-category information is:

**Category Utility(C,F)**

$$= \left[P(c)\sum_{i=1}^{n} P(f_i|c)\log P(f_i|c) + (1-P(c))\sum_{i=1}^{n} P(f_i|\ not\ c)\log P(f_i|\ not\ c)\right] - \sum_{i=1}^{n} P(f_i)\log P(f_i).$$

This measure of Category Utility is identical to the standard notion of the *information transmitted* between the message sets C and F.

In certain applications, we may be interested in defining the informational value of category $c$ separately from that of *not–c*. The Category Utility of category $c$ alone is given by:

$$\textbf{Category Utility(c,F)} = P(c)\left[-\sum P(f_i)\log P(f_i) \; - \; -\sum_{i=1}^{n} P(f_i|c)\log P(f_i|c)\right].$$

*The Guessing-Game Measure of Category Utility*

The information-theoretic measures of Category Utility given in the preceding section have close connections to expected performance in a feature prediction task. If we consider the expected score of someone guessing the values of each attribute dimension of an item, we can compare their expected score when they know nothing about the item to their expected score when they are told whether the items belong to $c$ or *not–c*.

Assuming that the receiver adopts a ***probability-matching strategy*** (e.g. the receiver guesses value $f_i$ with a probability equal to his expectation of the likelihood of $f_i$ occuring given $c$ or *not–c*) their expected increase in score given the category message can be shown to be given by:

$$\textbf{Category Utility(C,F)} = \left[P(c)\sum_{i=1}^{n} P(f_i|c)^2 + (1-P(c))\sum_{i=1}^{n} P(f_i|\ not\ c)^2\right] - \sum_{i=1}^{n} P(f_i)^2.$$

If the receiver is assumed to have no information about the *not–c* distribution, then the expected increase in score is

$$\textbf{Category Utility(c,F)} = P(c)\left[\sum_{i=1}^{n} P(f_i|c)^2 - \sum_{i=1}^{n} P(f_i)^2\right].$$

The information-theoretic and expected-score measures of Category Utility are closely related both in mathematical form and in terms of how they order categories as to relative goodness because $log(p)$ approximates to $p$ for small numbers. Futhermore, assumptions about alternate guessing strategies have little effect on the predicted orderings of categorization utility as long as the strategy predicts that the receiver will do best when one attribute value is certain and will do worst when all attribute values are equally likely. In all our empirical applications to date, the most significant discrepancy between results of the information-theoretic and the feature-prediction versions are a few cases in which a tie in the goodness of two categories was broken by use of the other version.

## *APPLICATIONS TO CATEGORIZATION EXPERIMENTS*

Hoffmann and Ziessler (1983) replicated the basic level phenomena using a series of three artificial category hierarchies. The hierarchies were differentiated by the degree to which exemplars of categories at different levels share common attribute values. Thus a different level was expected to be basic in each of the hierarchies.

Subjects were assigned to learn one of the three hierarchies. They were taught to associate each item with category names at three levels of generality (e.g. exemplar, intermediate, superordinate). Following this, subjects were presented with a picture of one of the items, paired with a concept name. They were asked to verify, as quickly as possible, whether or not the picture was an example of the named category. In a second task, they were asked to recall the correct name at a given level of abstraction. Reaction times for both the verification and naming studies indicated that the basic level was at the superordinate level for one hierachy, at the intermediate level for another, and at the exemplar level for the third hierarchy.

We evaluated the ability of our Category Utility index to predict, for each hierarchy, which level would be optimal or basic. The optimal level is operationally defined to be the level at which people are quickest to verify that an object is a member of the category. According to our theory, the average Category Utility of the categories at a given level should be highest for this optimal level. Utilizing a straightforward encoding of the drawings using three attribute dimensions (outline, edge, and bottom) with two, four, and four attribute values respectively, we calculated both the average Category Utility(c,F) of the individual categories and the average Category Utility(C,F) of the partitions induced by each individual category. For comparison, we also calculated the average cue validity, category validity, and the product of these two (G. Jones' (1983) "collocation" measure).

In these studies, cue validity and the collocation measure invariably identified the highest level as best. Category validity failed to distinguish between any of the levels. In summary, these measures were insensitive to the manipulation of attributes across the three hierarchies; each failed to predict the basic level in at least two out of three studies. The average Category Utility(c,F) index correctly predicted the ordering of reaction times for the three levels in *each* of the three hierarchies, with the exception of giving equal ratings to the basic and intermediate levels in the first hierarchy. The average Category Utility(C,F) index correctly predicted the ordering of reaction times in all three hierarchies.

In addition, in the analyses of a similar experiment by Murphy & Smith (1982), Category Utility was also the only measure which successfully predicted the basic level.

*DISCUSSION*

The results from these experiments indicate that Category Utility is able to predict the psychologically preferred level of categorization in these verification and naming experiments. None of the alternative measures did nearly as well. An additional advantage to the measure is that it is context sensitive: Category Utility is computed as an expected *decrease* in uncertainty given some context population. Thus, this affords a way of measuring how the utility of a category or generalization can change depending on the context in which it is analyzed. This is particularly important from a psychological standpoint because of evidence indicating that the usefulness of categories and concepts is highly context dependent (Murphy & Medin, 1985; Barsalou, 1982, 1983).

## References

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition, 10,* 82-93.

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition, 11,* 211-227.

Berlin, B., Breedlove, D. E., & Raven, P. H. (1973). General principles of classification and nomenclature in folk biology. *American Anthropology, 75,* 214-242.

Daehler, M. W., Lonardo, R., & Bukatko, D. (1979). Matching and equivalence judgments in very young children. *Child Development, 50,* 170-179.

Hoffmann, J. & Ziessler, C. (1983). Objectidentifikation in kunstlichen Begriffshierarchien. *Zeitschrift fur Psychologie, 194,* 135-167.

Jolicoeur, P., Gluck, M., & Kosslyn, S. (1984). Pictures and names: Making the connection. *Cognitive Psychology, 16,* 243-275.

Jones, G. V. (1983). Identifying basic categories. *Psychological Bulletin, 92,* 174-177.

Medin, D. (1983). Structural principles of categorization. In B. Shepp & T. Tighe (Eds.), *Interaction: Perception, Development and Cognition.* Hillsdale: Erlbaum.

Murphy, G. (1982). Cue validity and levels of categorization. *Psychological Bulletin, 91,* 174-77.

Murphy, G. & Medin, D. (1985). The role of theories in conceptual coherence. *Manuscript.*

Murphy, G. L. & Smith, E. E. (1982). Basic level superiority in picture categorization. *Journal of Verbal Learning and Verbal Behavior, 21,* 1-20.

Rosch, E. & Mervis, C. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573-603.

Rosch, E., Mervis, C., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382-439.

Shannon, C. & Weaver, W. (1949). *The Mathematical Theory of Communication.* Chicago: University of Illinois Press.