UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
On Knowing the Category Before Knowing the Features

Permalink
https://escholarship.org/uc/item/29r5h678

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 27(27)

ISSN
1069-7977

Authors
Kusumi, Takashi
Nakamoto, Keiko

Publication Date
2005

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

On Knowing the Category Before Knowing the Features

Kenneth J. Kurtz (kkurtz@binghamton.edu)
Department of Psychology, PO Box 6000
Binghamton University (State University of New York)
Binghamton, NY 13902 USA

Abstract

An assumption of all major accounts of categorization is that the system operates in a Features-First manner: a stimulus is mentally encoded in terms of observable properties which are then evaluated for fit to known categories. A testable prediction of this view is that people must know the features of an object before knowing what category it belongs to. Experimental results using a speeded verification task clearly show the opposite: people verify a category label more quickly than they verify a physical or functional feature. A theoretical groundwork for interpreting this finding is suggested. Categorization can be viewed as a means for constructing featural representations, rather than as the result of a comparison process between a "received" featural encoding and generic concept representations.

Introduction

Categorization is the central, ubiquitous process by which we make sense of the world. By categorizing, we interpret the stimuli in our immediate experience as examples of generic knowledge structures stored in long-term memory. Reseachers have proposed a variety of theoretical accounts and models to explain categorization, but there is little consensus on the major questions (Murphy, 2002).

One might be tempted to consider the possibility that there has been a misstep in the field. The theory or knowledge-based view of concepts (Murphy & Medin, 1985: Medin, 1989) issued a powerful critique of certain core assumptions widely held across the class of models designated as 'probabilistic.' In particular, the reliance on independent feature lists for item representation was argued to be inadequate and the reliance on a similarity computation between inputs and generic representations was argued to be fatally unconstrained. Despite these concerns, the fits of major models to behavioral data from laboratory studies have been robust and compelling (e.g., Kruschke, 1992). The theory view has failed to give rise to a process model that is competitive on these grounds. Therefore, the field stands in the position of offering a set of models that impressively account for only a highly managed portion of the problem of understanding categorization. It is possible that we are a little bit stuck.

In the present investigation, the idea is to take a step back and experimentally evaluate an assumption common to all major accounts of categorization. From the perspective of rules, similarity, prototypes, exemplars, probabilities, even theories, the problem has been articulated as follows: Find the best account for a set of input features in terms of known categories. Many models assume that some form of perceptual pre-processing serves to deliver a set of feature values as input to the categorization system. But as some researchers have noted over the years: the features of a stimulus do not arrive objectively from the bottom-up or merely for the asking (e.g., Schyns, Goldstone, & Thibaut. 1998; Wisniewski & Medin, 1994).

Implementations of the exemplar view such as ALCOVE (Kruschke, 1992) – representative of what is considered by many to offer the best available account of human category learning – address the issue of input representation in terms of *psychological dimensions*. Every stimulus is represented as a point in a space; as in a multidimensional scaling solution. Presumably this is intended to stand in for, rather than to explain at the process level how a physical stimulus is encoded in psychological terms. In actual practice (Kruschke, 1992), the step is passed over. As the author states, "It was assumed that the three physical dimensions of the stimuli [Shepard, Hovland, and Jenkins' (1961) geometric figures] had corresponding psychological dimensions" (p. 27).

To further emphasize the widespread commitment to the Features-First assumption, the power of models such as ALCOVE and SUSTAIN (Love, Medin, & Gureckis, 2004) rests to a considerable extent on a mechanism by which the degree of attention to each input feature is selectively updated based on the diagnosticity of that feature for classification purposes. With such an operation as an integrated part of the categorization system, it is clear that the features have to be available and subject to processing as a preliminary to categorization.

To summarize, major models of category learning rest on the assumption that inputs arrive to the categorization system in the form of feature lists or dimension values. This assumption does not prove damaging to models in part because most artificial laboratory studies use stimuli that actually *are* graphical instantiations of small sets of underlying binary-valued feature lists that are carefully packaged for easy access by experimental participants.

The focus of the present study is to evaluate the Features-First assumption. As discussed, the dominant approach to categorization defines the problem in terms of performing a set of computations on a fixed, available feature set. There is, however, an alternative perspective: the construal view of categorization (Kurtz & Dietrich, in preparation; Kurtz, 1997) in which the very goal of the categorization process is to construct an encoding of the elements of meaning that comprise a stimulus. This idea shares something in common with the theory view as well as with the notion of strong mutual dependency and flexible interactivity between perceptual and conceptual systems (e.g., Goldstone, 2003).

However, the construal view makes no commitment to causal/explanatory principles underlying concepts and is generally compatible with a modular perceptual front-end; placing the focus instead on the processes of recoding and enriching perceptually-derived initial representational content. To be clear, the construal view does not challenge the very idea of semantic features as units of representation; instead, the claim is that such features are the product of, rather than the input to, categorization. The construal view therefore makes the unusual-sounding prediction that we know what category an input belongs to before we have a meaningful description of the properties of that input.

Experiment 1

There has been no clear experimental test of the assumption that categorization begins with a featural item description. The Features-First assumption makes the testable prediction that people must encode the features of a stimulus in order to (and *prior* to) determining its category membership.

A speeded verification task is used to evaluate this prediction. Photographic images of highly familiar everyday objects were presented to participants – whose task was to evaluate whether a verbal descriptive matched the image. The manipulation in this within-subjects design was the type of descriptive that appeared. For each image in the set, three possible descriptives were prepared: 1) a Category label chosen as the expected basic level name for the object depicted; 2) a Functional feature chosen as an archetypal use or action associated with the object; and 3) a Physical feature chosen as an archetypal structural or perceptual attribute of the object.

The Features-First assumption implicit in all major accounts of categorization makes the prediction that the features of the stimulus are encoded initially and used as the basis from which to compute the best fitting category. A further specification of this standard view is that perceptually-available features are encoded initially while more conceptual features (such as functional features) are not immediately available, but must be inferred from an activated category representation. The Features-First view can therefore be summarized as: 1) Physical features encoded initially; 2) Category determined based on features; and 3) Function features inferred from Category.

Some theorists might suggest that Functional features are read off directly from the stimulus as part of the input to the categorization system (that is, it requires no top-down information processing to encode a chair as for *sitting*). The ecological approach with its focus on affordances is certainly not far removed from such a view and therefore might generate a prediction that functional verification occurs quickly and in a direct fashion rather than mediated through a category representation.

The construal views rejects the notion of an initially, fully-featured input and suggests that a category is activated (via heuristic methods which generate candidate categories based on raw visual information and situational context) and then leveraged to build an actual semantic encoding of the stimulus (Kurtz & Dietrich, in preparation).

Table 1: Theoretical predictions

Theoretical stance	Predicted Fastest Verification
Features-first	Physical
Ecological	Functional/Physical
Construal	Category

Method

Subjects A total of 82 undergraduates at Binghamton University participated in the experiment in order to receive course credit.

Materials Images (see Figure 1) were collected by searching the Internet for clear, representative photographs of everyday objects. Images were manipulated in order to show each object in isolation or presented on a generic surface. The size of the images varied in a range of approximately 3 to 5 inches in height and width. The most obvious descriptive was sought in all cases, though two additional constraints were applied: 1) no repetition of a descriptive across the item set; and 2) maximal avoidance of difficult, ambiguous, unusual, or low frequency words.

Table 2: Descriptives used for the critical items.

Category	Function	Physical
Banana	Eating	Peel
Baseball Bat	Hitting	Wood
Book	Reading	Paper
Calculator	Computing	Numbers
Camera	Photographing	Lens
Candle	Burning	Wax
Chair	Sitting	Legs
Clock	Timekeeping	Hands
Fork	Dining	Prongs
Glasses	Seeing	Fragile
Hammer	Pounding	Heavy
Ice Cream	Snacking	Cold
Lamp	Lighting	Bulb
Paintbrush	Painting	Bristles
Scissors	Cutting	Sharp
Stapler	Attaching	Metal
Telephone	Ringing	Buttons
Tennis Racket	Playing	Strings
Toothbrush	Cleaning	Plastic
Umbrella	Protecting	Handle
Vase	Containing	Delicate

Procedure Participants were given a thorough set of instructions that explained their task. They were asked to respond as quickly as possible as to whether or not the verbal descriptive matched the pictured everyday object. The instructions fully explained the three types of descriptives and gave clear examples. The Category descriptives were described as the "common name for the object." The Function descriptives were described as the "Function or Activity associated with the object." The Physical descriptives were described as a "Physical aspect of the object" and was further explained as a "feature, property, substance, characteristic, or part."

Figure 1: Two examples of object stimuli.

In addition, it was pointed out that the descriptives would not include "tricks." The instructions emphasized the importance of responding as quickly as possible without sacrificing accuracy. Participants were asked to keep their fingers in position above the response keys and were told they ought to be able respond in under one second. A practice phase was conducted in which participants acclimated to the task, practiced responding with a keypress as quickly as possible, and gained additional exposure to the three types of descriptives.

Participants were randomly assigned to one of three item counterbalancing forms. Each group was shown the exact same set of images, however the groups varied in the assignments of the images to descriptions. For each image, one group was asked the Category question about that image, a second group was asked the Functional question about that image, and the remaining group was asked the Physical question about that image. The assignments of descriptives to form were fixed and arbitrary aside from the criteria that each image appear once for each group and with a different descriptive in each group.

In addition to the set of Critical items, a set of Filler items were used to ensure that 50% of the presented items were paired with an accurate descriptive during test. The Filler items were the same for all three groups and they were always mismatches between the image and descriptive. The Filler items were photos of everyday objects just like the Critical items. The descriptives were chosen to be clearly wrong, but not distractingly so (care was taken to avoid near misses or humorously inappropriate descriptives). The Filler items were evenly distributed across the three types of descriptives. The combined set of Critical and Filler images was presented to each participant in a random order.

On each trial, participants were given 3s to prepare while a "Get Ready" prompt was shown on the computer screen. A fixation point appeared for 750ms and was replaced by the stimulus image. After a delay of 500ms, the verbal descriptive appeared below the image. The image was intentionally presented first so that the task consisted of processing the object stimulus and then evaluating the descriptive. Alternatively, it would be possible for participants to use the descriptive to guide their processing of the depicted image. With the delay, the initial processing of the image is neutral. At the same time as the descriptive appears, the cue words "Yes" (on the left) and "No" (on the right) also appeared in locations on the screen corresponding spatially to sticker-labeled response keys. No feedback was provided at any point during the task.

Results and Discussion

Latency data usually require a procedure to protect against the distorting effects of outliers. In the present data set, we applied a pruning process in which any single response time that was more than 2.5 standard deviations from the mean was removed. This procedure left the vast majority of data points intact, but a total of forty individual response latencies out of the entire data set were removed. The efficacy of this procedure was verified by computing medians on the raw data which closely paralleled the results of the pruning process.

The logic of the experiment was to evaluate response latency under an expectation of high accuracy of responding. There were several image-descriptive pairs that were removed from the dataset for mean percent correct accuracy below a threshold set at 70%. As it turned out, there was not high consensus (under speeded conditions and for these particular photographic representations) about glasses being fragile, toothbrushes being plastic, forks having prongs, tennis rackets having strings, clocks having hands, hammers being heavy, and telephones having buttons. These items were evenly distributed (2,2,3) across the three counterbalancing forms. While the overall results were not impacted either way, the analysis was conducted on the remaining 59 of the original 66 image-description pairings. For the critical items, mean accuracy was above 90% for all three item types. The filler items were successfully employed in that they were overwhelmingly rejected by participants at a rate of 98%.

The mean latency data for correctly answered items are shown in Figure 2. A repeated measures ANOVA was conducted on response latency revealing a significant main effect of item type, $F(2, 81) = 128.25$, $MSe = 1186891$, $p < .001$). Paired sample t-tests showed all pairwise differences to be reliable ($p < .001$). Category descriptives were significantly faster to verify than either Physical or Function descriptives. A smaller effect also showed Function descriptive to be verified faster than Physical. Therefore, the expected ordering from the Features-first view was found to be lacking on all counts. It is worth noting that the large size of these observed differences is on a different scale than variation attributable to lower-level processes such as reading time. These effects are at the level of semantic processing.

In a secondary analysis, a significant interaction was found between counterbalancing form and item type. There was no main effect of forms. In order to interpret the interaction, an ANOVA was conducted separately for each form. The Physical vs. Function difference was significant in one form ($p < .001$), marginal in the second form ($p < .1$) and non-significant in the third form ($p > .3$). Accordingly, there is some question about the generality of the Function vs. Physical difference. However it is clear that the fastest responding occurs for the Category descriptives and that this effect is robust across forms.

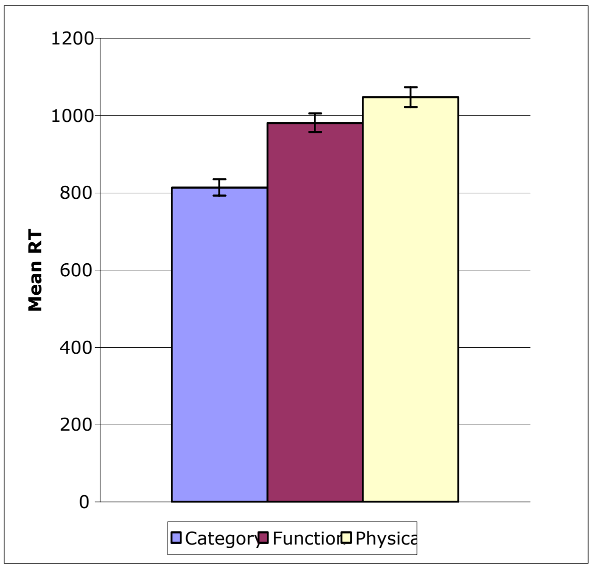

Figure 2: Response latency for correct items by condition.

Additional analyses were conducted to evaluate potential differences among possible subtypes within each item type. The physical features were selected with awareness of the following subtypes: characteristics (e.g. heavy); parts (e.g., lens); and substances (e.g., wood). No reliable differences were found comparing mean latency for these subtypes. The Function items can also be interpreted as subtypes: what the object does (e.g., scissors cut); what is done by a person to an object (e.g. bananas are eaten); and what is done by a person with an object (e.g., glasses allow seeing). Once again, no reliable differences in performance were observed across subtypes.

Experiment 2

The second experiment was designed to replicate the basic finding under slightly different task conditions. In the first experiment, participants did not know what type of description they would be asked to evaluate on any given trial. It is possible that participants developed a strategy such as assuming a particular question type and then compensating when wrong. It should be noted that the three types of questions were equally frequent, so there was no obvious basis for establishing such a hierarchy. If the two types of feature descriptives are considered together, there are in fact twice as many feature verifications to make as category verifications. Therefore, if there were to be a frequency-based bias, it ought to be toward preparing for a feature verification rather than a category verification. Another possibility is that the task was tapping into some unusual form of cognitive processing because of the multifaceted, uncertain nature of the task from trial to trial.

In the current experiment, the items from the three within-subjects conditions were presented in blocks rather than randomly distributed. Therefore, participants received three blocks of trials and within each block all of the descriptives were of the same type (Category, Functional or Physical). The order of the three blocks was randomized by subject.

The same pattern of results was predicted: even when the participant knows what type of descriptive they will be asked to evaluate, it should take longer to verify a feature then a category. If this pattern is observed, it provides even stronger evidence that the semantic encoding of the features of an object is slower than the encoding of its category membership.

Method

Subjects A total of 79 undergraduates at Binghamton University participated in the experiment in order to receive course credit.

Materials The same stimuli were used as in Experiment 1.

Procedure The same procedure was used as in Experiment 1 except that item order was randomized by blocks of item type rather than by item. Participants were instructed that they would encounter all three types of items, but that each of the types would be grouped together.

Results and Discussion

The same procedure was applied to remove outliers from the response time distribution resulting in the removal of 20 individual trial RT's across the entire data set. The filler items were again successful in that they were overwhelmingly rejected by participants at a rate of 99%. After the pruning process, mean accuracy on the critical items was above 90% for all three item types. The mean latency data for correctly answered items are shown in Figure 4. A repeated measures ANOVA was conducted on response latency revealing a significant effect of item type, $F(2, 88) = 157.57.25$, $MSe = 1945930$, $p < .001$). Paired sample t-tests showed all pairwise differences to be significant ($p < .001$).

Once again, a significant interaction was found between counterbalancing form and item type. This time, all of the follow-up comparisons showed reliable pairwise differences. The interaction is likely due to one of the forms showing somewhat faster mean latencies on only some of the item types. An additional issue in this design is whether the order of the blocking of the item types influenced performance. A mixed-design ANOVA testing the repeated measures factor of item type and the between-subjects factor of blocking order (six possible orderings of the three blocks) showed no significant interaction ($p > .1$).

Additional data collection was conducted in order to evaluate the image-descriptive pairings used in Experiments 1 and 2. While every effort was made to choose the most obvious and appropriate descriptives of each type, it is important to evaluate these selections. In a separate mini-experiment, the set of images used for the critical items in the previous experiments was presented to participants in a random order. The task was to type into a response area on the computer screen the first descriptive that came to mind. Unlike the previous experiments, a between-subjects design was used, so each participant was asked to produce only one type of descriptive throughout the task. For example, in the Category condition, the participant was asked to type in the first Category label that came to mind for each image.

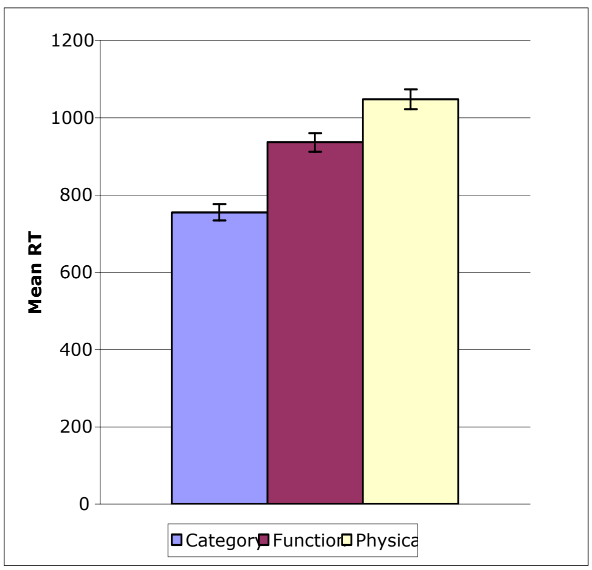

Figure 3: Response latency for correct items by condition under blocked presentation (E2)

The resulting data allowed us to determine which of the descriptives used in Experiments 1 and 2 were of high or low dominance in a generation task. By considering only the subset of high-dominance descriptives, an evaluation was possible of the speed of verification differences beteween item types for only the most salient, accessible, and agreed-upon descriptives. A preliminary version of such an analysis showed all of the observed pairwise differences remaining intact for the subset of the experimenter-selected descriptives which matched the descriptives most frequently generated by participants presented with these images.

General Discussion

The surprising result of these studies is that people know the category of a familiar everyday object before they know its features. The result is actually consistent with introspective experience: when you look up and see an object in the room, what do you know first about it: that it has *legs*, that it is for *sitting*, or that it is a *chair*? The results of two studies offer concrete evidence contradicting the Features-First view.

There are possible counterarguments, but none that are especially compelling. It is possible that the features are encoded, but somehow unavailable for purposes such as a verification task. Features may be encoded in some notation that is sufficient as input to the categorization system, but not sufficient to allow a fast verification judgment. If this is the case, it is an important issue to begin to understand. The evidence is clear that, at the very least, people verify categorical information more quickly than featural information for highly familiar object categories. The further conclusion that the mental encoding of category membership precedes the mental encoding of compositional semantic elements of the stimulus also seems hard to escape.

Additional work is underway to evaluate whether features are verified more quickly than categories for newly acquired or weakly understood categories. Such a reversal would provide further insight into the machinery of concept formation and use.

Acknowledgments

Many thanks to Aliza Nelson, Leora Schanfield, Soon Park, Ai Koizumi, Huned Rangwala, Eric Dietrich, and the members of the Learning and Representation in Cognition (LaRC) Laboratory at Binghamton University.

References

Goldstone, R.L. (2003). Learning to perceive while perceiving to learn. in R. Kimchi, M. Behrmann, and C. Olson (Eds.) *Perceptual Organization in Vision: Behavioral and Neural Perspectives.* Mahwah, New Jersey. Lawrence Erlbaum Associates. (pp. 233-278)

Kurtz, K.J (1997). The Influence of Category Learning on Similarity. *Unpublished doctoral dissertation*.

Kurtz, K.J. & Dietrich, E. (in preparation). The construal view of categorization.

Kruschke, J.K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99,* 22-44.

Love, B.C., Medin, D.L, & Gureckis, T.M (2004). SUSTAIN: A Network Model of Category Learning. *Psychological Review*, 111, 309-332.

Medin, D.L. (1989). Concepts and conceptual structure. *American Psychologist, 44,* 1469-1481.

Murphy, G.L. (2002). The big book of concepts. Cambidge, MA: MIT Press.

Murphy, G.L. & Medin, D.L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92(3)* 289-316.

Schyns, P. G., Goldstone, R. L., and Thibaut, J. (1998). The development of features in object concepts. *Behavioral and Brain Sciences, 21,* 1-54.

Wisniewski, E. & Medin, D.L. (1994). On the interaction of theory and data in concept learning. *Cognitive Science, 18*, 221-281.