Learning a Hierarchical Organization of Categories

Steven Verheyen (steven.verheyen@psy.kuleuven.be)
University of Leuven, Department of Psychology
Tiensestraat 102, B-3000, Leuven Belgium

Eef Ameel (eef.ameel@psy.kuleuven.be)
University of Leuven, Department of Psychology
Tiensestraat 102, B-3000, Leuven Belgium

Timothy T. Rogers (ttrogers@wisc.edu)
University of Wisconsin-Madison, Department of Psychology
1202 West Johnson Street, Madison, WI 15206 USA

Gert Storms (gert.storms@psy.kuleuven.be)
University of Leuven, Department of Psychology
Tiensestraat 102, B-3000, Leuven Belgium

Abstract

Although exemplar models of category learning have been successfully applied to a wide range of classification problems, such models have only rarely been tested on their ability to deal with vertical category learning, that is, cases where the same stimuli may be classified at multiple levels of abstraction. We report an experiment in which participants learned to classify artificial stimuli at both levels of a nested hierarchy and displayed more accurate classification of these items at the lower level of the hierarchy than at the more general level. Some authors have suggested that exemplar models would have great difficulty accounting for this phenomenon, but we show that the ALCOVE exemplar model effectively captures the behavioral pattern arising in the experiment. Despite suggestions to the contrary, superior performance at the lower level of a nested hierarchy does not necessarily invalidate the class of exemplar models.

Keywords: context model; exemplar models; vertical category learning; hierarchies; basic-level effect.

Introduction

When developing and testing formal models of category learning, researchers have primarily relied upon experimental paradigms in which artificial stimuli need to be classified into one of several categories at the *same* level of abstraction. While the wealth of classification models currently available testifies to the fruitfulness of emphasizing what Rosch (1978) termed the horizontal dimension of categories, doubts remain about their usefulness when it comes to determining category membership at *different* levels of abstraction. Vertical relationships between categories, of the kind that exist between natural language categories such as *bulldog*, *dog*, *mammal*, and *animal*, have rarely been studied using traditional artificial category learning methods (see Lassaline, Wisniewski, & Medin, 1992; Mervis & Crisafi, 1982; and Murphy & Smith, 1982, for a few notable exceptions), and efforts to address such data with formal

models are even more rare (Estes, 1993; Palmeri, 1999). Thus the adequacy of such models to explain critical aspects of human category learning remains in question.

Palmeri (1999) considered how vertical category learning might challenge the class of exemplar models. To see this, let us follow Palmeri in thinking about the issue within the context of a hierarchically organized category structure and by investigating the manner in which classification probabilities are computed in Medin and Schaffer's context model (1978), the precursor of all current exemplar models. Let us assume that a particular set of stimuli belongs to category A, while a different set of stimuli belongs to a separate category B. Let us further assume that category A is comprised of two subordinate categories C and D. That is, each of the A stimuli also belongs either to C or to D. A similar hierarchical relationship holds between category B and subordinate categories E and F. According to the context model, the probability of assigning a particular stimulus to one of the specific categories (C, D, E, F) can never exceed the probability of classifying the stimulus in the specific category's superordinate (A or B).

More specifically, the context model proposes that evidence E_X for classifying a stimulus in a particular category X is accumulated by summing the similarity of each of the category's exemplars to the stimulus. Classification probability $P(X)$ is then taken to be the ratio of the evidence E_X to the evidence that the stimulus belongs to any of the categories at the abstraction level of the target category. The probability of classifying a stimulus in category A is written accordingly as:

$$P(A) = E_A / (E_A + E_B), \quad (1)$$

while the probability of appointing the stimulus to category C becomes:

$$P(C) = E_C / (E_C + E_D + E_E + E_F). \quad (2)$$

Because of the hierarchical category structure of the example, each of the superordinate categories can be considered the union of its subordinates, so that the evidence for a more general category is equal to the summed evidence of making a decision in favor of the comprising specific categories:

$$E_A = E_C + E_D \text{ and } E_B = E_E + E_F. \quad (3)$$

Thus, the denominators in equations 1 and 2 above are equal, while the numerator for superordinate classification (equation 1) must be equal to or higher than the numerator for subordinate classification (equation 2). Accordingly, the context model predicts that subordinate classification can never exceed superordinate classification. This prediction, however, was violated by a study of human category learning conducted by Lassaline, Wisniewski, and Medin (1992).

Palmeri pointed out that the response rule in ALCOVE (Kruschke, 1992) - a connectionist extension of the context model - incorporates nonlinearities that exempt the model from the problem of being unable to produce higher classification accuracy at the subordinate level of classification. In order to verify this assertion Palmeri sought to replicate Lassaline et al's findings, and to determine whether they would invalidate ALCOVE. He had one group of participants classify artificial stimuli in four categories (C, D, E, and F), while another group classified the same stimuli in two superordinates (A and B). As in Lassaline et al., Palmeri found that performance was better at the subordinate level than at the superordinate level. Moreover, he showed that ALCOVE could account for the observed pattern of performance, indicating that, although vertical category learning might challenge some particular kinds of exemplar models, such learning did not pose an in-principle problem for exemplar-based approaches more generally, or for the ALCOVE model in particular.

In an influential textbook on semantic concepts Murphy (2002) questioned these conclusions, pointing out that, in Palmeri's experiments, each participant only ever learned to classify the stimuli at one level of abstraction (either subordinate or superordinate, but not both). Murphy argued that, since participants never learned that a particular stimulus can belong to more than one category, an essential aspect of the vertical organization of natural language concepts was missing among Palmeri's artificial counterparts. As a consequence it remains unclear whether any exemplar model can capture performance in a truly hierarchically organized classification task.

The current work addresses this question. We extended Palmeri's experiment by incorporating Murphy's suggestion. In our experiment, every participant learned that each stimulus belonged to one of four specific categories *and also* to one of two more general categories. Since learning multiple category labels for a single stimulus is quite challenging, we chose to have participants learn the correct classifications at one level of abstraction before having them move on to the other level. A first group of participants learned to classify at the specific level before moving on to the more general level; whereas a second group of participants started by learning the more general classification followed by the specific classification. A similar procedure was previously used by Murphy (1991) and by Murphy and Smith (1982). We assessed whether (i) the pattern of better performance at the subordinate level of abstraction than at the superordinate one was retained when participants learned that a stimulus belonged to more than one category and (ii) whether ALCOVE was able to account for such classification performance.

Method

Participants
The participants were 36 undergraduate students of the Universities of Leuven and Wisconsin-Madison who took part in the study for partial fulfilment of course credit.

Materials
We used the same stimuli Palmeri (1999) did. Participants were shown schematic drawings of spaceships that differed on four dimensions: the shape of their nose, tail, wings, and porthole. Along each of these dimensions every spaceship took one of four possible shapes. These stimuli correspond to the ones previously employed by Hoffman and Ziessler (1983).

Every participant was presented with 12 different spaceships which they needed to classify both at a specific level as a C, D, E, or F and at a general level as either an A or a B. As was true in Palmeri (1999) and Lassaline, Wisniewski, and Medin (1992) Categories C and D were subordinate to A, while E and F were subordinate to B. Table 1 shows how in the experiment each of the values along the first stimulus dimension points towards a specific category. A hierarchical organization is obtained by having values 1 and 2 along this dimension signify membership of Category A, while values 3 and 4 signify that a stimulus belongs to Category B.

At the onset of each experimental session physical dimensions and shapes were randomly assigned to the stimulus structure.

Procedure

Each participant was randomly assigned to one of two conditions. Half of the participants first performed a supervised category learning procedure at the specific level. Upon presentation of a particular spaceship they were required to indicate whether it was of the Diler (C), Frite (D), Jebet (E), or Katone (F) type. Participants were provided with corrective feedback after each incorrect decision, indicating what the answer should have been. The 12 different spaceships were presented in a random order for 25 blocks or until a perfect score was achieved on two consecutive blocks. In this case a perfect score was awarded on the remaining blocks. At this point participants were told that the procedure would be repeated but that they would

have to classify the spaceships as members of two categories instead of four. We stressed that spaceships would be identical to those in the first part, but no further information on the relationship between both parts was provided. In the second part of the session the nonsense names Alpha and Beta were used for categories A and B, respectively. The supervised category learning task again lasted for 25 blocks. When participants were able to achieve a perfect score on two consecutive blocks, the session came to an end earlier. A perfect score was awarded on the blocks that otherwise would have followed.

Table 1: Stimulus and category structure.

Stimulus	Stimulus structure				Category structure	
	D1	D2	D3	D4	General	Specific
1	1	1	3	4	A	C
2	1	2	4	1	A	C
3	1	3	1	2	A	C
4	2	4	2	1	A	D
5	2	1	3	2	A	D
6	2	2	4	3	A	D
7	3	3	1	3	B	E
8	3	4	2	4	B	E
9	3	1	3	1	B	E
10	4	2	4	2	B	F
11	4	3	2	3	B	F
12	4	4	1	4	B	F

The other half of the participants first distinguished categories at the general level before moving on to the specific level. In all other respects they were subject to the same procedure as the participants completing the experiment in the reverse order.

Results and Model Fitting

Figure 1 holds the average probability of committing an error with regard to level of abstraction, learning order, and learning block. The general level of abstraction is indicated by circles, while the specific level is indicated by squares. The two experimental conditions are separated by colour. Black symbols indicate a start at the general level, while white symbols indicate a start at the specific level. The coding of the learning curves is of the format *abstraction level (learning order)*. We first report on the statistical analysis before discussing the ALCOVE simulations.

Statistical Analysis

Using a between-subjects design Palmeri (1999) found that classification of items at the lower level of the nested hierarchy was superior to classification of these items at the more general level. For our purposes, of major interest is the question whether the performance difference remains when a within-subjects design is employed. To address this question the experimental results were first subjected to a 2 (general level first vs. specific level first) x 2 (general level

vs. specific level) x 25 (learning block) analysis of variance with learning order as a between-subjects variable and level of abstraction and learning block as within-subjects factors. All statistical tests were performed at the .05 level of significance.

A significant main effect of block [$F(24,816) = 48.42$, $MS_e = .70$] was found, indicating that participants improved as the experiment progressed. A significant two-way interaction between level and block [$F(24,816) = 4.25$, $MS_e = .07$] was also observed, indicating faster learning of specific relative to general categories in the earliest blocks.

Most interesting to our current purposes, however, was a significant order x level interaction [$F(1,34) = 8.16$, $MS_e = 5.54$] which we tested by contrasting the two levels of abstraction for both learning orders. The effect of level did not reach significance among participants who had learned to differentiate between specific categories before moving on to the general ones [$F(1,34) = 3.12$, $MS_e = 2.12$, $p = .08$]. For this particular learning order, the effect of level observed by Palmeri was not retained. However, the participants who completed the experiment in the reverse order did perform better at the specific level than at the general level [$F(1,34) = 5.17$, $MS_e = 3.51$], thus replicating Palmeri's observation using a within-subjects manipulation. Note that this is exactly the pattern of results that, according to reasoning explained above, challenges some kinds of exemplar models. Unlike the undergraduates participating in Palmeri's experiment, participants in this experiment were aware of the fact that stimuli could be classified at multiple levels of abstraction.

Finally, the analysis of variance revealed that the observed difference between levels was more pronounced among the first learning blocks, as indicated by the three-way interaction between order, level, and block [$F(24,816) = 2.02$, $MS_e = .03$].
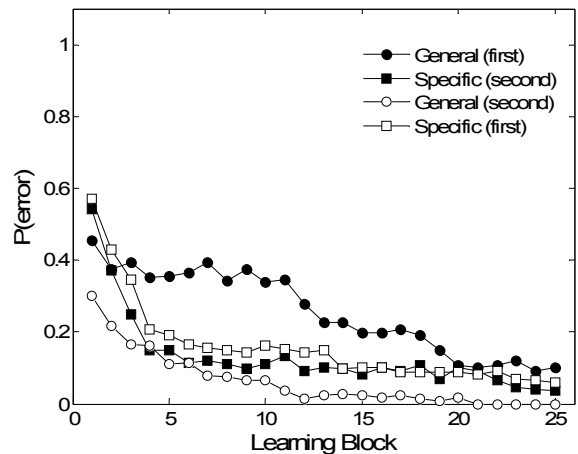

Figure 1: Averaged probability of error with regard to level of abstraction, learning order, and learning block.

In the above analysis we followed Palmeri (1999) in treating learning block as a 25-level factor in the ANOVA to enable comparability with his results. One might object to

this since the block levels follow a specified order and cannot be considered independent. However, the critical observation of the current work – better performance at the specific level of abstraction than at the general level when general categories are mastered first – does not rely on the inclusion of learning block. The observed effect of level was corroborated by a Bayesian analysis in which it was demonstrated that the learning curves, corresponding to the two levels of abstraction, were more likely to be fit by two separate exponentially decaying functions than by a single one. (See Dry, Lee, Vickers, & Hughes, 2003 for a similar use of function-fitting to test experimental hypotheses.)

Model Fitting

In our simulations we employed a version of the exemplar model ALCOVE which is detailed in Palmeri (1999). It differs from the original version (Kruschke, 1992) in that it deals with discrete values along the input dimensions. When a stimulus is presented along these dimensions it activates a layer of exemplar nodes through the similarity of its representation to theirs. The extent of this activation is determined by the strength of attention the model learns to allocate to each stimulus dimension. Exemplar node activation is then passed on to all category nodes through learned association weights. Here activation strengths and association weights are linearly combined to determine classification evidence. Classification evidence is finally converted into classification probability.

For each learning order maximum likelihood fits to the empirical data were obtained using a Nelder-Mead simplex algorithm. We employed a six (four specific and two general) category node ALCOVE model that was initialized by setting attention strengths and association weights to zero and awarding random values to the model's four parameters. The model was then trained on a first sequence of 25 x 12 randomly ordered stimuli. During this training sequence attention strengths and association weights were updated after every presentation of a stimulus to the model. Depending on which learning order was being simulated, either the association weights leading to the four specific category nodes or those leading to the two general ones were updated.

After all stimuli from the first sequence had been processed, the model was further trained on a second sequence of 25 x 12 stimuli. The association weights that had been updated during the first sequence were left unchanged, while the others were now updated through backpropagation after each trial. This manner of adapting the ALCOVE model to the procedure of our experiment assumes that there is information taken from one part of the experiment into the other and that this is located at the level of the attention strengths: They are updated across all 600 stimulus presentations.

From 1000 repetitions of this procedure for every learning order, the simulations by the model with the highest likelihood were withheld. Figure 2 shows the simulated probability of error with regard to level of abstraction,

learning order, and learning block. As was true for Figure 1, the general level of abstraction is indicated by circles, while the specific level is indicated by squares. Each learning order received a different colour code. Black symbols indicate a start at the general level. White symbols indicate a start at the specific level. The coding of the simulated learning curves is of the format *abstraction level (learning order)*.
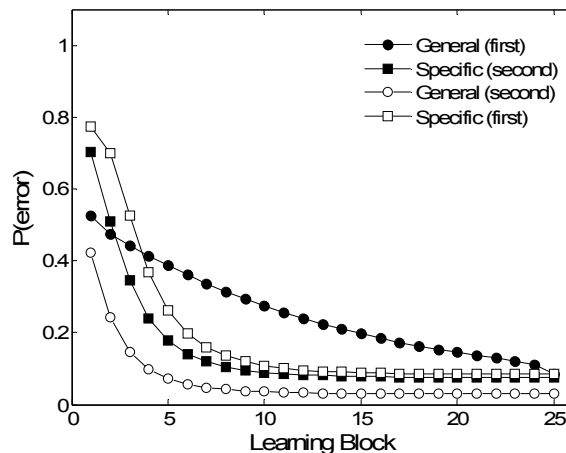

Figure 2: Probability of error with regard to level of abstraction, learning order, and learning block as simulated by ALCOVE.

Of major interest was ALCOVE's ability to display superior performance on the more specific level of a truly hierarchically organized category structure. More accurate classification of items at the lower level of a nested hierarchy than at the more general level had been deemed challenging for the class of exemplar models. Analysis of the empirical results revealed that this pattern arose in our experiment when participants started classification at the general level. As is shown by the black curves in Figure 2, the ALCOVE model had no difficulty accounting for this phenomenon.

No such difference in performance was observed when participants completed the experiment in the reverse order. In fact, although it did not reach significance, the effect of level among participants starting classification at the specific level was in the opposite direction. Participants were somewhat more accurate classifying items at the general than at the lower level of the hierarchy. There is no profound reason why exemplar models would have difficulties accounting for the latter result, and such is evidenced by the white simulated learning curves in Figure 2: They display the desired pattern of performance.

Discussion and Summary

Formal models of category learning hardly ever treat classification at various levels of abstraction. For them to remain valid accounts of human classification performance they need to be explicitly tested on their ability to deal with

such a vertical organization of categories. Indeed, it has long been established that many of the entities we encounter in everyday life are referred to with multiple names, often indicating a hierarchical organization of categories. Some authors have suggested that exemplar models would have great difficulty accounting for experimental results in which superior performance on the more specific level of such a hierarchically structured organization of categories is shown - a phenomenon reminiscent of the basic level effect in natural categories (Rogers & Patterson, 2007; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). In the current work, we took up this challenge to exemplar theory by putting ALCOVE, a widely used exemplar model, to test on the results of an experiment in which participants dealt with a hierarchically organized category structure.

We extended earlier work by Lassaline, Wisniewski, and Medin (1992) and by Palmeri (1999) who devised and used a hierarchical category structure that yielded a basic-level-like advantage: Participants assigning stimuli to categories at the specific level of the hierarchy outperformed participants classifying the items in categories at a more general level of the hierarchy. Murphy (2002) argued against such an interpretation of the results since participants were only required to classify stimuli in categories at a single level of abstraction. Hence, they never became aware of the hierarchical organization present among the various categories. Murphy's concerns about the employed procedure proved warranted when we had participants classify stimuli at both levels of abstraction. When participants started classification at the specific level of abstraction and then moved to classification at the general level no significant difference between the two levels of abstraction was observed. However, when participants completed classification in the reverse order, the effect did come about. However, the supposedly challenging effect did not pose a problem for the exemplar model ALCOVE. It demonstrated more accurate classification of items at the lower level of the nested hierarchy than classification of these items at the more general level when the general level was mastered first. Thus, better performance at the subordinate level of a hierarchical category structure than at the superordinate one does not necessarily invalidate exemplar theory.

Clearly, a full explanation of why an effect of abstraction level was observed for one learning order and not the other is of further interest. For such an explanation to correctly inform us about the manner in which people acquire vertical category structure, the current work would have to be extended by control conditions in which participants learn about categories at multiple levels of abstraction that are not hierarchically organized. Such vertical-but-not-hierarchical organizations abound in natural language. The set inclusion relation, essential to the hierarchical organization of categories (Murphy, 2002) is violated in the case of the categories *dog*, *pet*, and *animal*, for instance. Although all *pets* are *animals*, it is not true that all *dogs* are *pets*. Some *dogs* are *stray dogs*. Inclusion of control conditions of this kind would allow for the disentanglement of effects arising from the hierarchical nature of category organization and effects due to stimuli belonging to more than one category.

Such a study would obviously benefit from complementary modeling endeavours to test alternative hypotheses about the mechanisms involved in vertical category learning. For instance, the nonlinearity of the response rule implemented in ALCOVE clearly contributes to the model's ability to display a difference in performance between two levels of abstraction (Palmeri, 1999), but it remains unclear to what extent the interesting behaviour of the model (and the participants) arises from the development of different attention strengths for different levels of abstraction. Moreover, applied to vertical category learning results different from the ones reported here, such modeling efforts might highlight the need for additional mechanisms to deal with the classification of stimuli at different levels of abstraction.

The strength of the current endeavour lies precisely in the fact that the ALCOVE model captures the important qualitative pattern in the presented data with little task-specific tailoring. That is, with very minor adjustments, a model formulated to address horizontal category learning also captured a challenging qualitative pattern arising in an experiment assessing vertical category learning.

Acknowledgments

References

Dry, M. J., Lee, M. D., Vickers, D., & Hughes, P. (2006). Human performance on visually presented travelling salesperson problems with varying numbers of nodes. *Journal of Problem Solving, 1*, 20-32.

Estes, W. K. (1993). Models of categorization and category learning. *Psychology of Learning & Motivation, 29*, 15-56.

Hoffman, J., & Ziessler, C. (1983). Objectidentifikation in kunstlichen Begriffshierarchien [Object identification in artificial concept hierarchies]. *Zeitscrift für Psychologie, 16*, 243-275.

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

Lassaline, M. E., Wisniewski, E. J., & Medin, D. L. (1992). Basic levels in artificial and natural categories: Are all basic levels created equal? In B. Burns (Ed.), *Percepts, concepts, and categories: The representation and processing of information*. Amsterdam: North-Holland.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85*, 207-238.

Mervis, C. B., & Crisafi, M. A. (1982). Order of acquisition of subordinate-, basic-, and superordinate-level categories. *Child Development, 53*, 258-266.

Murphy, G. L. (1991). Parts in object concepts: Experiments with artificial categories. *Memory & Cognition, 19*, 423-438.

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

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

Palmeri, T. J. (1999). Learning categories at different hierarchical levels: A comparison of category learning models. *Psychonomic Bulletin & Review, 6*, 495-503.

Rogers, T. T., & Patterson, K. (2007). Object categorization: Reversals and explanations of the basic-level advantage. *Journal of Experimental Psychology: General, 136*, 451-469.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization*. Hillsdale, NJ: Lawrence Erlbaum.

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