When children fail to learn new categories: the role of irrelevant features

Jean-Pierre Thibaut (JPTHIBAUT@ULG.AC.BE)
University of Liège. Department of Psychology (B32)
5, Boulevard du Rectorat, 4000 Liège BELGIUM

Abstract

When subjects are confronted with new stimuli that they have to learn to categorize, they have to segment them into relevant features for categorization. Two experiments with four to eleven year old children investigated whether certain irrelevant perceptual aspects of the stimuli prevent learning the relevant features for categorization. In the first experiment, it was shown that children used salient holistic aspects of stimuli for categorization despite the fact that they were only partially relevant for categorization, whereas perfect cues for categorization requiring analysis were not discovered by children. In the second experiment, it was shown that children cannot abstract the relevant cues for categorization when irrelevant perceptual characteristics were crossed with the relevant ones. When these irrelevant cues were absent, children could learn the relevant cues for categorization. Children's biases towards locally salient properties can impair, or even prevent learning new categories, when these are defined by comparatively less salient features. Results are discussed in terms of the relation between children's cognitive competences and the abstraction of relevant descriptors for new stimuli.

Introduction

Developmental psychologists have to explain how children discover the features that are relevant for categorization. In conceptual development, for example, one might explain how children come to know that features such as *four-legged*, *barks*, and so forth are features that distinguish the category of *dogs* from other categories. The fact that a set of objects can be described in terms of an infinite number of descriptors (features) and that children are only exposed to a limited set of data contributes to the difficulty of abstracting the relevant features. To reduce the number of possible parsings, it has been hypothesized that young learners have specific biases towards particular aspects of stimuli (see Eimas, 1994; Landau, 1994; Markman, 1995; Quine, 1960). Some authors have suggested that children's parsing are influenced by perceptual constraints as well as by theories about the organization of the world.

According to the theory-view of conceptual development, new features and concepts are driven by the development of naive theories about the world (Gelman, 1988; Keil, 1989); e.g., the concept of "house" contains features such as "has walls", "has a roof". A theory about "houses" must explain the function of these features. However, most likely, theories do not specify all the perceptual aspect of the relevant features: in the "house" example, the theory would not specify all the possible aspects of roofs or of walls. At best, theories suggest the features a subject should search for, not their perceptual characteristics. There are also situations in which subjects have no a priori theory about the new objects that they have to learn to categorize. In such cases, there is no theory available for the segmentation of the stimuli into relevant features. This will be the case in the experiments to be reported.

Some authors have also suggested that children have biases towards perceptual characteristics of objects. For example, experiments have shown that children's generalizations about objects, at a particular stage of their development, are governed by the shape of objects (see Jones & Smith, 1993; Landau, 1994, for reviews). Children neglect large differences on other aspects of objects. To illustrate, suppose that children have learnt to associate a new word with a set of objects that share a particular shape, texture, size, and so forth. In a subsequent test phase, given the choice between test stimuli that have the same shape or the shape texture or the same size as the original stimuli, children generalize the new word to the test stimuli that have the same shape. This result has been referred to as the shape bias. However, one has to explain (a) how children learn to segment a shape into its relevant descriptors, given the large number of possibilities, (b) attentional changes towards particular aspects of a shape during development.

Differences in analytic ability could explain differences between adults and children's abilities to segment stimuli into their features. Smith (1989) argued that young children and adults are capable of isolating stimulus dimensions. However, they differ on their capacities to focus on isolated dimensions when they compare objects. Young children seem to use holistic aspects of stimuli that are analyzed into components by adults (see Thibaut, 1995, for an experimental demonstration). Consequentely, since subjects have to segment new stimuli appropriately in order to abstract the relevant features for categorization, it can be hypothesized that children will have difficulties when they have to discover these relevant cues for categorization. Given the importance of this topic, it is surprising that there are so few data on the development of segmentation skills and the way they interact with concept learning.

In the following experiments, children aged four to eleven years participated in a category learning task. In order to

discover a classification rule, subjects had to segment a set of stimuli into relevant features for categorization. In two experiments, we manipulated the perceptual saliency of features relevant or irrelevant for categorization. We analyse whether these differences in saliency could influence the discovery of the features relevant for categorization.

In the first experiment, subjects had to learn to categorize a set of new unfamiliar stimuli into two categories. In the first experimental condition, the stimuli were composed of two parts, the lower and the upper part. The upper part could be used as a characteristic feature (i.e., a feature neither necessary nor sufficient for categorization). The perfect cue for categorization was hypothesized to have a low saliency (a particular spatial layout of the legs, see Figure 1a). In the second experimental condition, there was no salient characteristic feature (see Figure 1b).

In order to categorize the stimuli adequately, subjects had to focus their attention on the relevant feature and filter all the other aspects of the stimuli. It was predicted that because of their limited analytic skills, it would be easier for children (particularly young children) to classify stimuli on the basis of the salient characteristic feature than on the basis of the sufficient, although not salient, feature. In other words, children should not find the feature sufficient for categorization and should not be able to categorize the stimuli that do not have a characteristic feature.

We also designed a second experimental condition (the *legs only* condition) in which we presented the stimuli used in the first condition, except that the characteristic feature was erased. It was hypothesized that children would discover the cues relevant for categorization more easily since there was no salient feature that could draw their attention.

In experiment 2, the aim was to evaluate more thoroughly the role of irrelevant features for categorization on the abstraction of the relevant ones.

Experiment 1

Methods

Subjects Twenty children aged 8;0 to 8;11 and 20 children aged 11;0 to 11;11; 10 children aged 6;0 to 6;11 also participated in the second experimental condition.

Material For the first experimental condition (called *body* condition), two categories (A and B) of eight stimuli were constructed. The stimuli are outlines of unknown shapes composed of two parts, the upper part (the body) and the lower part (four legs). The two categories have the same structure. In five out of the eight stimuli, the upper part has a mushroom shape that is slightly distorted over the 5 stimuli in the case of category A, and an angular shape in the case of category B stimuli. These two shapes were selected for their distinctiveness and perceptual saliency. The three remaining stimuli from the two categories were constructed with three different upper parts (UP1, UP2, UP3). Since UP1, UP2, and UP3 are present in the two categories they cannot be used as cues for categorization (see Figure 1a). For each stimulus, the lower part is composed of four legs that are spatially grouped as one leg on the left and three legs on the right in category A (1-3), and two sets of two legs (2-2) in category B (see Figure 1a). Participants have to learn this distinction between the two categories (1-3 versus 2-2) in order to categorize all the stimuli appropriately. The size of the stimuli is approximately 7 by 7 centimeters.

In the second experimental condition (*legs only* condition), we erased the upper part of the stimuli (see Figure 1b).

Procedure Familiarization phase. The set of stimuli was presented once to the subject. Each stimulus was shown for five seconds. Then, it was removed and followed by a new stimulus. There was no feedback during this phase. At the end of this phase, participants were told that they would have to learn to sort the stimuli into two categories, the name of which was provided.

Learning phase. A first stimulus was presented for approximately five seconds and the subject had to guess its name. The experimenter gave the appropriate feedback and presented the second stimulus in the same way, followed by the other stimuli. Feedback was provided about the accuracy of the answer. The order of presentation of the stimuli was random. Once the entire set of stimuli had been presented to the subject, it was presented a second time. The learning phase ended after two successive presentations of the set of stimuli with no errors or after the ninth presentation. Subjects were tested individually. A session lasted for 10 to 25 minutes, depending on the number of trials necessary to complete the task.

Results and discussion

In the *legs only* condition, none of the six and the eight year olds, and three eleven year olds could reach the criterion. In the *body* condition, one eight year old and no eleven year old could reach the criterion. A Fisher Exact Test revealed that no age group differed significantly from any of the other age groups. In order to assess whether the upper part was used as a partial cue for categorization when it is distinctive (i.e. for stimuli with the mushroom shape or the angular shape) in the *body* condition, we compared the number of errors produced for the five stimuli with the distinctive upper part and the three stimuli with no distinctive upper part (UP1, UP2, UP3). A two-way ANOVA (2 x 2) with Age as a between variable, and Part (*Distinctive* part and *No distinctive* part) as a within variable was performed on the mean number of errors obtained for the 10 stimuli with adistinctive (characteristic) part and the mean number of errors obtained for the 6 stimuli without distinctive upper part (the UP stimuli) was carried out. There was a significant effect of Age: $F(2,27) = 5.33$ ($p < .05$) (8 year olds: mean = 1.54; 11 year olds: mean = 2.19), a significant effect of Part: $F(1,27) = 157.49$ ($p < .0001$) (with distinctive upper part: mean = .8; no distinctive upper part: mean = 3.7), and no significant interaction between Age and Part.

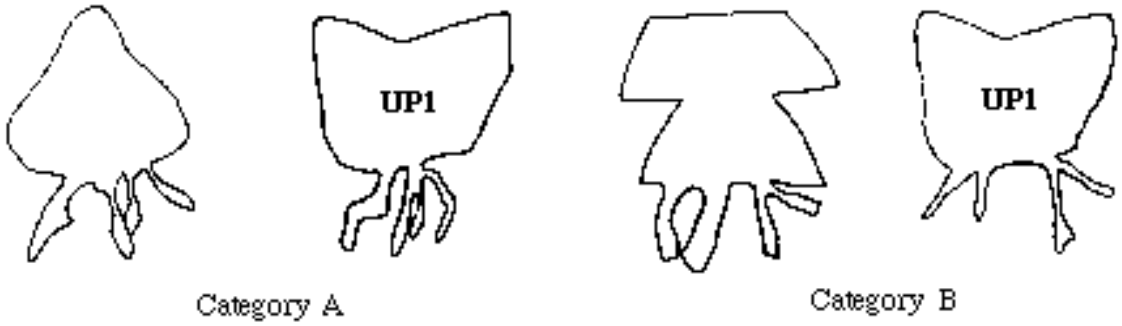

Figure 1a. Four stimuli from categories A and B. The first simulus has the upper part characteristic of category A and the third stimulus has the upper part characteristic of category B. The UP1 stimuli are neutral stimuli.

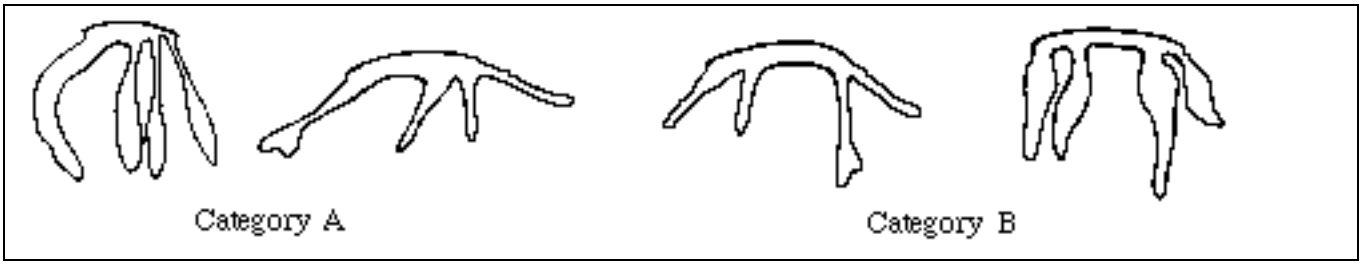

Figure 1b. Four stimuli from categories A and B.

The results indicate that children were unable to abstract the relevant cue for categorization. Surprisingly and contrary to expectations, the *legs only* condition was not easier than the *body* condition. In fact, a subsequent analysis showed that children made more mistakes in the *legs only* condition. In other words, even when the salient upper part, irrelevant for categorization, was removed children could not find the relevant cue. This indicates that children had difficulties when they had to abstract a spatial cue within the legs. When they were able to use the upper part as a distinctive cue, they had virtually no problem to solve the categorization task.

The preceding results can be interpreted in two different ways: 1) the rule is intrinsically too complex for children; 2) aspects of stimuli, irrelevant for categorization (the body part, irrelevant properties of the legs, such as their size, ...), are salient for children and mask the relevant cues for categorization. Separate interviews after the experiment revealed that children focused on irrelevant aspects of stimuli. Surprisingly in the *legs only* condition, they referred to details such as the length of the legs, or their orientation, or their size, or other features irrelevant for categorization. This observation means that children noticed details of the stimuli and tried to use them as categorization cues. Moreover, preliminary observations indicated that young children could use the relevant rule 2-2 versus 1-3 when the experimenter pointed explicitly to this characteristic on the stimuli.

Experiment 2

The preceding experiment showed that subjects could not abstract the relevant features for categorization. In the following experiment, we made the relevant feature for categorization more salient, e.g., by separating the two subsets of legs (the subset of 2 legs from the other subset of 2 legs, or the isolated -1- leg from the subset of 3 legs) more clearly.

We also added a systematic dimension that could also be used as a feature for categorization: half of the stimuli had large and vertical legs whereas the rest of them had thin legs with one leg pointing to the right. Given our interviews of children or their spontaneous comments, it can be hypothesized that a rule for categorization defined in terms of size or orientation would be easily learnt by children.

We therefore designed a set of stimuli that could be categorized according to two sets of different relevant cues, each defining an experimental condition. In the first condition (A), the two categories of stimuli had to be categorized according to the cues "thin legs, one leg pointing to the right" for the first category versus "large and vertical legs" for the second category. In the second condition (B), the stimuli had to be categorized into two categories (orthogonal to the categories of condition A) based on the cues relevant for categorization used in the preceding experiment (1-3 versus 2-2). We also designed a third condition (condition C) in which all irrelevant features for categorization were removed : the relevant dimension was 2-2 versus 1-3 and the legs had the same orientation and size.

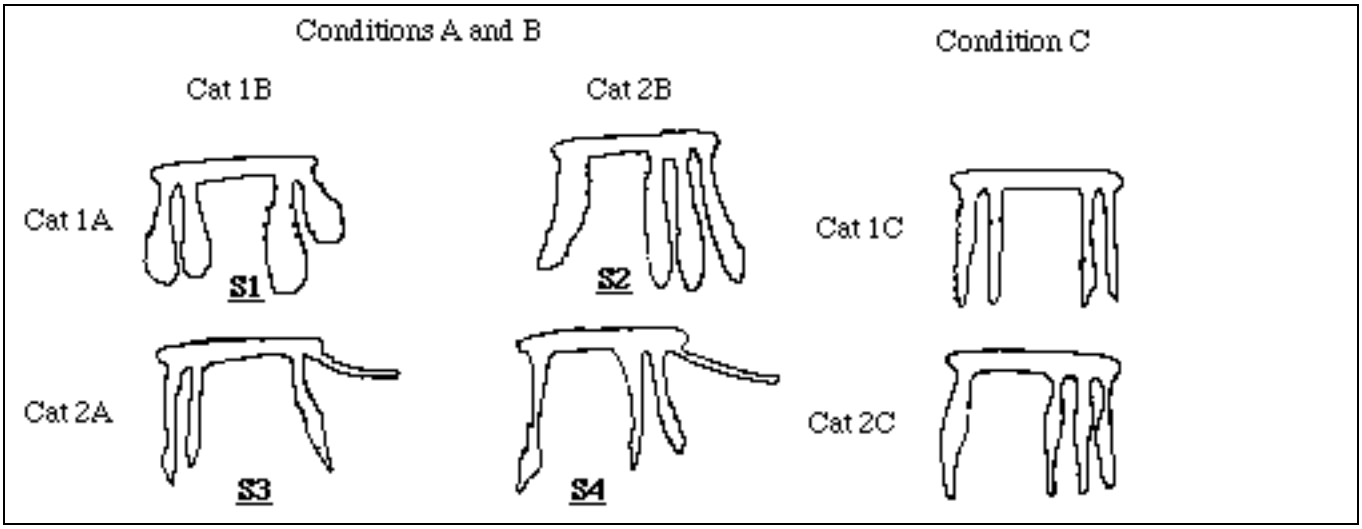

Figure 2. Stimuli from condtions A, B, and C. Note that Cat 1A contains **S1** and **S2,** Cat 2A contains **S3** and **S4** (Condition A). Cat 1B contains **S1** and **S3**, and Cat 2B contains **S2** and **S4** (Condition B).

	Four year olds		Six year olds		Eight year olds	
	Correct	Failure	Correct	Failure	Correct	Failure
Condition A	10	0	9	1		
Condition B	0	10	3	7	9	1
Condition C	8	2	10	0		

Table 1: Number of subjects who reached the criterion in the three groups of age (out of 10) and the three experimental conditions. (Note: empty cells were not run)

We predict that subjects will learn the relevant dimension 1-3 versus 2-2 more easily in the present situation than in the preceding experiment. We also predict that subjects will be able to learn the relevant dimension "thin-one leg pointing to the right" versus "large and vertical legs", since it is based on an obviously salient feature.

The results obtained for conditions A and B will be compared. If children from condition A learn the criterion more easily than children from condition B, and if children from condition C can learn the relevant feature 1-3 versus 2-2 more easily than children in condition B, this will mean that the rule for categorization 1-3 versus 2-2 was masked by the irrelevant cues "thin-one leg pointing to the right versus large and vertical". On the other hand, if condition A is as easy as condition C, this will mean that the dimension 'thin-one leg pointing to the right versus large and vertical", when it is relevant for categorization, is not masked by the dimension 2-2 versus 1-3, when this latter is irrelevant for categorization. This asymmetry would mean that, despite the fact that new relevant features for categorization can be learnt at a given age, they can be masked by other dimensions irrelevant for categorization whereas other relevant features are not or less masked by irrelevant dimensions. In other words, a failure to learn a new concept does not necessarily result from an intrinsic conceptual complexity but can be driven by the presence of irrelevant salient features.

Method

Subjects Twenty children aged 4;0 to 4;11, 30 children aged 6;0 to 6;11, and 10 children aged 8;0 to 8;11.

Stimuli Three experimental conditions called A, B, and C were designed. In the experimental conditions A and B the same set of stimuli was used. A set of 16 stimuli that could be categorized according to two orthogonal rules was constructed. All the stimuli were composed of four connected legs. The cues "thin" "vertical", "the rightmost leg pointing to the right", and "large" were crossed with the cues "one leg and three legs" (1-3) and "two sets of two legs" (2-2) according to four types of stimuli. There were four 1-3 stimuli and four 2-2 stimuli with "thin" legs and "the rightmost leg pointing to the right", and four 1-3 stimuli and four 2-2 stimuli composed of "broad and vertical legs" (see Figure 2 for examples of the 4 types of stimuli). Cat 1A and Cat 2A were categorized according to the cues "large and vertical" versus "thin and one leg pointing to the right" and defined condition A. Cat 1B and Cat 2B were categorized according to the cues 1-3 versus 2-2 and defined condition B.

In the third experimental condition C, 16 new stimuli were constructed, eight 1-3 stimuli and eight 2-2. The difference with the stimuli used in conditions A and B is

that all the legs were thin and vertical. In other words, we tried to remove salient irrelevant features for categorization.

Procedure Children were randomly assigned to one of the three conditions. The instructions were the same as in the first experiment. The children completed the task in 5 to 25 minutes. All subjects were tested individually.

Results

We hypothesized that conditions A and C would be easier than condition B because of the presence of salient irrelevant features for categorization in condition B. Table 1 summarizes the results in the different conditions according to age group. A Fisher Exact Test revealed that there was no significant difference between Condition A and C at four years. It also revealed that the three conditions differed significantly at six years ($p < .05$). Importantly, Conditions B and C also differed significantly ($p < .05$). Children failed to abstract the relevant cue in Condition B until the age of 8. Results suggest that Condition C was as easy as Condition A. However, the difference between Condition B and Condition C indicates that the presence of irrelevant perceptually characteristics in Condition B rendered the abstraction of the relevant feature more difficult.

General discussion

In Condition C of the second experiment, children aged 4 or 6 years could learn the 1-3 versus 2-2 rule appropriately. This result indicates that children aged 4 years have the necessary analytic abilities to abstract the rule. Why did they fail to find the rule in condition B and in the first experiment? It could be argued that, in the first experiment, they focused on the salient upper part. However, when this part was removed (*legs only* condition) or even when the stimuli were simplified (condition B experiment 2), they still failed to abstract the relevant cues.

These results are consistent with the hypothesis that certain characteristics of the stimuli are more salient (such as length, size, and so forth) than others. If one of these aspects is a cue relevant for categorization, subjects discover it easily. It could be argued that subjects in condition C, experiment 2, did not abstract the rule but learnt to recognize two unanalyzed patterns. However, in a test phase not described here, children could correctly generalize the rule to new stimuli that had a different global shape.

One has to explain why children cannot discover particular relevant features for categorization in the presence of others. It could be argued that children focus on perceptual aspects and that the rule 1-3 versus 2-2 is a formal rule with no perceptual interpretation. Complementarily, according to this explanation, children would have found the rule in the experiment 2, condition C because the rule was perceptually salient. However, the relevant spatial configurations 1-3 and 2-2 also have several perceptual interpretations in the other conditions: e.g., 1-3 is an asymmetric pattern whereas the cue 2-2 is symmetric; or, the left part of a 1-3 stimulus is a singleton, whereas for 2-2 stimuli, the left part has two components, a difference that is grounded in perceptual aspects. To summarize, our results cannot be explained on the sole basis of the perceptual saliency of the cues relevant for categorization.

One further possibility is that children cannot ignore perceptually salient irrelevant cues when they have to focus on less salient aspects. The difference between conditions B and C is that, in condition B, children failed to ignore salient aspects of the stimuli, despite the fact that children have the analytic abilities required to solve the task (see condition C). Note, however, that condition B was easier than the two conditions in experiment 1. Most likely, this results from the fact that there was more variability in the stimuli in the first experiment (e.g., the legs were of unequal size, shape, orientation, etc.) than in stimuli in condition B. To summarize, those results support the hypothesis that children's difficulties arise when children have to filter out aspects of stimuli irrelevant for categorization. Since they have the analytic abilities to abstract the relevant features for categorization, their difficulties rest probably in the fact that they do not compare stimuli systematically or that they forget the hypotheses they have already tested.

The shape bias described in the introduction cannot account for the difference between conditions A and B. In order to explain our results, an explanation in terms of a shape bias should mention why subjects has a bias towards particular aspects of a shape. This would amount to explaining why different features have not the same saliency and why salient features prevent the abstraction of less salient features. In a similar vein, the notion of *a priori* naive theories (see introduction) cannot account for the differences between the conditions. Indeed, if an *a priori* theory was available, there was no reason to obtain a difference between conditions B and C in experiment 2.

The results presented here have important implications for the early extraction of features and category learning. Children's biases towards locally salient properties could impair, or even prevent, their learning of new categories, when these are defined by comparatively less salient features. Explanations of conceptual development have to consider the interaction of different sources of information, especially the possibility that certain information might mask the presence of relevant cues (Schyns, Goldstone, & Thibaut, in press). This suggests that when children cannot acquire a new concept, the problem could be due to irrelevant aspects that interfere with the relevant ones, as much as difficulties with the rule.

Acknowledgements

The author would like to thank Bob French, Jean Rondal for valuable comments on this manuscript, Rob Goldstone and Philippe Schyns for many stimulating discussions. Thanks also to Christine Comblain, Myriam Dupont, Juliette Quadri, Tiziana Russo for their help in running the experiments.

References

Gelman, S.A. (1988). The development of induction within natural kind and artifact categories. *Cognition*, *20*, 65-95

Keil, F.C. (1989). *Concepts, kinds and cognitive development*. Cambridge, MA: MIT

Kemler, D. G. (1989). The nature and occurrence of holistic processing. In B. E Shepp & S. Ballesteros (Eds.), *Object perception* (pp. 357-386). Hillsdale, NJ: Lawrence Erlbaum.

Landau, B. (1994). Object shape, object name, and object kind: representation and developement. In Medin (Ed.). *The Psychology of Learning and Motivation*, *31*, 253-304. Academic Press: San Diego, CA.

Markman, E. (1995). Constraints on word meaning in early language acquisition. In L. Gleitman & B. Landau (Eds). *The acquisition of the lexicon*. (pp. 199-227). Cambridge, MA: MIT Press.

Schyns, P.G., Goldstone, R. & Thibaut, J.-P. (in press). The development of conceptual features. *Behavioral and Brain Sciences*.

Smith, L. B. (1989). A model of perceptual classification in children and adults. *Psychological Review*, *96*, 125-144.

Thibaut, J.-P. (1991). Récurrence et variation des attributs dans la découverte des traits conceptuels. Unpublished doctoral thesis, Université de Liège, Liège.

Thibaut, J.P. (1995) The development of features in children and adults: the case of visual stimuli. *Proceedings of the Seventeenth Annual Meeting of the Cognitive Science Society,* 194-199, Hillsdale, NJ: Lawrence Erlbaum.

Thibaut, J.P., & Schyns, P.G. (1995). The development of feature spaces for similarity and categorization. *Psychologica Belgica*, *35*, 167-185.