Category learning of schematic faces in children: Relationships between attribute knowledge and modes of processing

CHANTAL PACTEAU, FRANÇOISE BONTHOUX,

PIERRE PERRUCHET, and JACQUES LAUTREY

CNRS & Université René Descartes, Paris

Abstract. *In a pluralistic view of cognitive functioning, it has been proposed that, when having to learn categories, individuals can use two forms of processing: either an analytic mode in which the necessary and sufficient properties are extracted, or a holistic mode in which objects are processed as indivisible webs of attributes and relations. To further investigate this issue, children were observed during a perceptual categorization task based on the Kemler Nelson (1984) paradigm. Besides the usual error pattern, response times and attribute knowledge were used to diagnose what process was elicited to learn family resemblance categories. After a study phase of two sets of schematic faces which varied on six attributes (shape of eyes, nose, etc.), 9-year-old children were administered a category assignment task during which response speed and accuracy were recorded. Subsequently, knowledge of individual attributes was evaluated using a specifically designed task, the attribute identification task. Convergent data from the two tasks show that most of the children were multiple attribute learners, but used either one attribute (about 60% of the subjects) or more attributes (the remaining 40%) to make category decisions. Reasons for the absence of genuine Gestalt-like processes in our study are discussed.*

Key words: Category learning, analytic and holistic processing, faces.

Correspondence should be sent to Chantal Pacteau, Laboratoire de Psychologie Différentielle, Université René Descartes, 28 rue Serpente, 75006 Paris, France.

INTRODUCTION

Classically, theories of categorization (Bruner, Goodnow, & Austin, 1956; Inhelder & Piaget, 1964; Vigotsky, 1962) have emphasized a unique process, abstraction, which identifies the properties that objects have in common and combines these properties through the use of logical operations to form classes. In this view, an object can only be a member of a given class if it has the necessary and sufficient properties that define that class. The reduction of classification to this sole activity has been challenged by findings which suggest that it is triggered by artificial stimuli which lack psychological coherence (Medin, Wattenmaker, & Hampson, 1987) and/or whose properties (shape, color, size, etc.) are often independent (Rosch, 1978). Objects belonging to natural categories are better described by co-occurrences between properties and by their perceptual and functional resemblance relations than by lists of independent properties. Categorization processes could thus reflect the information bundle structure of real-world objects. According to Rosch, Simpson, and Miller (1976), individuals elaborate categories by generalizing on the basis of overall similarity with prototypical members. The more prototypical a category member, the more properties it shares with other members of its category and the less it shares with members of contrasting categories. For Brooks (1978) and Medin and Schaffer (1978), generalization by overall similarity is not built up from a mean prototype (whether real or virtual) but from exemplars that maintain their identities and allow for accumulation of episodic information, which in turn generates rich and flexible knowledge.

For certain authors, the corollary to category structure diversity is the subject's capacity to employ different modes of object grouping. In the field of perceptual category learning which is the focus of the present paper, two main classes of processes are thought to operate (Kemler Nelson, 1984, 1989): an *analytic* process which involves serial comparisons of discrete properties and singles out some of them to make groupings, and a *holistic* process which apprehends objects as indivisible webs of attributes and relations and sorts them on the basis of their overall similarity. Which process is elicited is the result of an interaction between an individual endowed with certain competences and goals, and an environment afforded with specific and meaningful properties (Lautrey, 1988, 1990; Rieben, de Ribaupierre, & Lautrey, 1990). In this vein, several studies (Brooks, 1978; Kemler Nelson, 1984; Medin & al., 1987; Sugimura & Inoue, 1987 a & b, 1988) have been devoted to the issue of the influence of factors such as type of objects, category

structure and learning conditions (which result in intra-individual variations), and subject's age and expectations (which result in inter-individual variations). In particular, it has been hypothesized that family resemblance categories favor holistic responding and that the developmental trend proceeds from holistic to analytic modes of processing (Kemler Nelson, 1984; Kemler & Smith, 1978).

In the field of categorization, as in others, the conception of a plurality of mental processes does not receive unanimous support. For proponents of the classical view, empirical evidence for holistic processes is the result of faulty diagnosis. This is the reason why the current debate is centered on questions of methodology. Since the Kemler Nelson paradigm (1984) is widely used to investigate the issue (Kemler Nelson, 1988, 1989; Sugimura & Inoue, 1987 a & b, 1988; Ward, 1988, 1989; Ward & Scott, 1987; Ward, Vela, & Hass, 1990), we shall only briefly describe the general procedure to make it clear how it can be used to pinpoint different strategies.

Kemler Nelson's method is based on the assumption that the way subjects apprehend the two categories to be learned can be deduced from the combined index of the number and nature of errors made during a category assignment task. An example of category structure – the one used in our study – is presented in Table 1.

The category members are schematic faces made up of a small number of parts (referred to in the literature as *attributes*) such as hair and nose, whose shape (in the literature, *value*) differs across categories. Two principles guide the way the categories are constructed. First, the members of each category are highly similar to one another since they only differ from their prototype by the value of a single attribute. Second, none of the attribute values is defining since each face possesses a different bundle of the prototypical attribute values of its category. The *analytic categorizer* differentiates attribute values and then elaborates rules on the basis of that differentiation, the simplest rule being to selectively attend to the value of a single attribute while ignoring the others. Note that in the category structure used here, this single attribute rule does not yield perfect categorization performance: no information on category membership is given for faces which differ from their prototype on the *focused-upon attribute* (i.e., the attribute which is judged as defining by the subject), which results in random decisions. The *holistic categorizer* groups all the faces with a strong resemblance in the same family, without decomposing them into attributes. This strategy can lead to perfect scores. An ingenious way of diagnosing what process has been operating is to test subjects in a category assignment task where all

faces have one attribute of the opposite family. The purpose is to trap single attribute processors: those who can correctly categorize all faces except the ones for which the focused-upon attribute has the opposite value. For example, a subject focusing on the nose would judge face 3A2 (which has the category B nose) as belonging to category B and inversely, face 3B2 (nose A) as belonging to category A. This sort of inversion (that we shall later call "Focus" errors) causes a very specific pattern of errors; note that such errors are errors with respect to the overall similarity structure of categories, but not to the subject's point of view. Kemler Nelson considers that subjects who do exhibit this typical pattern of responses are analytic; subjects who do not – i.e., those whose errors, if any, are distributed across different attributes – are regarded *by default* as holistic.

TABLE 1. Description of the combination of attribute values of each face in categories A and B.

A face was composed of six attributes: hair (H), eyebrows (B), eyes (E), nose (N), mouth (M), and chin (C). Each face differed from its prototype by a single attribute, which had an intermediate value of 2 for both categories in phase 1 (study phase), and the value of the opposite category in phase 2 (category/ assignment task).

	Family	Attributes H B E N M C	Family	Attributes H B E N M C
Prototypes	A	A A A A A A	B	B B B B B B
Exemplars Phase 1	1A1	A A A A A X	1B1	B B B B B X
	2A1	A A A A X A	2B1	B B B B X B
	3A1	A A X A A A	3B1	B B B X B B
	4A1	A A X A A A	4B1	B B X B B B
	5A1	A X A A A A	5B1	B X B B B B
	6A1	X A A A A A	6B1	X B B B B B
Exemplars Phase 2	1A2	A A A A A B	1B2	B B B B B A
	2A2	A A A A B A	2B2	B B B B A B
	3A2	A A A B A A	3B2	B B B A B B
	4A2	A A B A A A	4B2	B B A B B B
	5A2	A B A A A A	5B2	B A B B B B
	6A2	B A A A A A	6B2	A B B B B B

This way of equating single attribute strategy to analytic process, and absence of errors, or errors distributed across attributes, to holistic responding needs reexamination. First, analytic categorization cannot be reduced to the single attribute strategy; and second, the absence of systematic errors does not testify by itself to holistic categorization. The distributed pattern of errors, as the zero error pattern, can occur every time subjects learn category membership by processing two or more attributes and elaborate a majority rule to make a category decision. Take for instance a subject who applies a disjunctive rule based on the values of three attributes belonging to the same category, for example category A. This rule can be phrased as follows: "To be an exemplar of category A, faces must have mouth A and eyebrows A or mouth A and nose A". The subject could correctly assign all exemplars to the correct family since for all A faces, 5 out of 6 attributes have the value A; the non-A faces are categorized as B by default. A perfect score can also be obtained by discovering the rule used by the experimenter to build both categories, that is: "To be an exemplar of A (or B), a face must have five out of six attributes with an A (or B) value". Between these two extremes, there are different majority rules that can be used, the number of which depends on the number of attributes composing the faces. Whatever these rules, they are analytic and can result in patterns of responding that might be mistaken for holistic processing.

Because of the amount of theoretical work that the Kemler Nelson paradigm has generated and of its empirical relevance – especially in studies with children for whom the task is immediately signifiant and does not raise problems of expertise (all children are "expert" in faces, whatever their age) – it is crucial to clarify whether distributed patterns of errors are underpinned by Gestalt or multiple attribute processes. To do so, different approaches used in the field of categorization may be useful.

One approach which involves taking for granted that categorization in the child evolves from part-undifferentiation to piecemeal modes of processing (Kemler & Smith, 1978; Kemler Nelson, 1984, 1989), consists of charting changes in response patterns at different ages. If patterns of distributed errors are underpinned by holistic operations, there should be an age-related decrease in the proportion of subjects showing such patterns. If this trend is not observed, namely if an increase (or stability) in the proportion of patterns of distributed errors occurs with age, it can be inferred that category decisions result from majority rules.

A second approach is to manipulate processing resources. When these resources are limited, as during incidental learning, it has been hypothesized that category formation proceeds holistically; by contrast, when subjects mobilize their resources as in intentional learning, they are thought to develop rules, even when no simple principles underlie category structure (Brooks, 1978; Kemler Nelson, 1984, 1988). This suggests that during learning there should be Gestalt processing under incidental conditions, and attribute computations under intentional conditions.

A third way to diagnose modes of processing is to study response times when category decisions are being made. It is generally assumed (Cooper, 1982) that the holistic mode is governed by parallel operations, with the percept being "Gestaltly" compared with the holistic representation in memory. It is thus faster than the analytic mode, which requires sequential scanning of different attributes to apply the rule of category membership.

Data on these three approaches are scanty and the findings are inconsistent. The validity of Kemler Nelson's reports that show more holistic responding in children than in adults – and more in incidental than in intentional conditions of learning – has been challenged in a number of studies. For example, Sugimura and Inoue (1987 a & b, 1988) observed that a large proportion of 6-year-old children can clearly perform as single attribute categorizers, and that instruction effects on the elicitation of one mode of processing over another are extremely limited. Similarly, Ward and his collaborators (Ward, 1988, 1989; Ward & Scott, 1987; Ward et al., 1990) have provided strong evidence that category processing is analytic for highly schematic stimuli in both children and adults learning under intentional and incidental conditions, even with family resemblance categories. This evidence is supported by data from response patterns and time measures (Ward, 1989; Ward & Scott, 1987). What apparently differentiates children from adults is the latter's ability to elaborate sophisticated analytic rules and modify hypotheses when faced with disconfirming evidence (Gholson & Beilin, 1979). Another source of difference arises from the low value children place on identity as a special kind of the sameness relation (Evans & Smith, 1988): in attribute identity categorization tasks (which is what Kemler Nelson's task is), defining attribute(s) may be processed as highly similar, instead of identical. This disregard of identities in object grouping may lead to instability in the weighting of attributes, and consequently, to apparent overall similarity categorizations. Regarding the observed contrast in performance between incidental and intentional procedures – which is

not seen in all learning situations (Perruchet & Pacteau, 1990) – the discrepancy may not result from the nature of the elicited processes, but from differences in what is computed. Finally, response times are not sufficient by themselves to infer underlying mental operations (Marquer & Pereira, 1990).

From this brief review, it appears that none of the above approaches can provide a clear answer to the question of the existence of holistic processing in perceptual categorization. In our view, this is because they lack an empirical assessment of the essential property of holistic processes – whether in its "weak" or "strong" form – namely, that "attribute information is not accessible" (Kemler Nelson, 1989). To test this point, we specifically designed a task intended to assess subjects' attribute knowledge after the category assignment task. In this new task (the "attribute identification task"), subjects are asked to assign each attribute, presented out of the context of a whole face, to the family to which it belongs. In addition, we used the indirect measure of attribute scanning, that is, we recorded response times during the category as-signment task. Our rationale was that, to discriminate holistic processing from multiple attribute analysis, performance in the category assignment task and the attribute identification task have to be congruent on two points. Holistic categorizers are those subjects (a) whose category deci-sions take less time or the same amount of time, as decisions based on a single attribute, *and* (b) whose responses in the attribute identification task are given randomly. Conversely, multiple attribute categorizers are those subjects (a) who take longer to respond than single attribute categorizers (they need to process more information), *and* (b) whose capacity to identify isolated attributes is better than chance.

The degree of decontextualization of the rule built by single attribute categorizers was also examined. Our rationale stems from Ward and Scott's observation that latencies are longer on conflictual faces than on non-conflictual ones, in all subjects except 5-year olds. They interpreted these response time variations as intrusions of some known relations between the value of the attribute focused upon and the value of the other attributes when assigning category membership: conflictual infor-mation interferes and delays decision making. By contrast, no response time difference between correct and incorrect responses may indicate either learning restricted to the focused upon attribute, or autonomy vis-à-vis known facial attributes which are irrelevant to rule application. In the present study, we used Ward and Scott's method to assess rule decontextualization.

In order to obtain a sample with a comparable number of analytic and holistic subjects, according to Kemler Nelson criteria, we observed 9-year-old children. Kemler Nelson (1984, Experiment 4) reported that at approximately this age, children are able to pick up categories organized either on the basis of a strong family resemblance or defining attributes, which she takes as evidence for the occurrence of holistic processing (in "the family resemblance problem") or analytic processing (in "the defining attribute problem").

METHOD

Subjects

Seventy-one middle-class third graders (36 girls and 35 boys) with a mean age of 9 years 1 month (range 8.0 to 10.5) took part in the experiment.

Material

Phases 1 and 2. The schematic faces of the two categories (six faces per category) were composed of six parts (or attributes): hair (H), eyebrows (EB), eyes (E), nose (N), mouth (M), and chin (C).
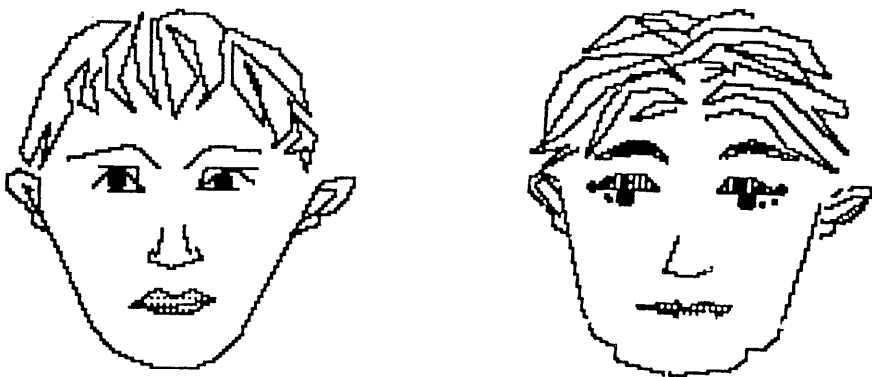

Figure 1. The two prototypes used to derive the exemplars for categories A and B. The faces were generated by Police Artist (Sir-Tech).

Each face measured approximately 9 cm in height and was drawn in black on white cardboard. Each of the six faces in a category differed from the category prototype by a single attribute. An abstract description of these faces is provided in Table 1. As shown, the set of five prototypical attributes differed for each face in a given category, and the attribute which differed from the prototype had an intermediate value of X for both categories in phase 1, and had the value of the opposite category in phase 2 (B for category A, A for category B). The same intermediary values were used for both categories in phase 1, so as not to provide any information on category membership.

Phase 3. The stimuli were the six isolated attributes of each of the two prototypical faces. Each attribute was drawn in black in the center of a white piece of cardboard.

Procedure

Children were tested individually in a room of their school. They were told they would be playing a game which was unrelated to school activities and to which there were several ways of responding. They were then told that they would be simultaneously shown two sets of faces hidden behind the two pieces of cardboard hung on the wall in front of them. "Behind this piece of cardboard" (experimenter pointed to one display) "are the faces of the inhabitants of a remote galaxy called Futura, and behind this one" (experimenter pointed to the other display) "are the faces of their enemies. Look at them carefully because I am going to hide them again. Then, I am going to show you some new faces and I would like you to tell me which are the inhabitants of Futura and which are their enemies. Of course, the inhabitants of Futura look very similar and their enemies all look alike too."

Study phase (phase 1). After the instructions, the experimenter removed the cardboard screens and the children were allowed 45 seconds to watch the two sets of faces (exemplars of phase 1, Table 1) presented side by side; each set was composed of six faces arranged in a circle. For 30 subjects, the faces of the Futura inhabitants (category A) were placed on the left-hand side and the faces of their enemies (category B) on the right. Left-right presentation was reversed for the remaining 31 subjects. Moreover, for both presentations, A faces were classified as "Futuras" and B faces as "Enemies" approximatively half of the time; and vice versa for the other half.

All exemplars were presented simultaneously to allow subjects to develop a strategy of their own. Simultaneously seeing all the members of a category might prompt subjects to attend to the strong resemblance between members of a category and then to process them holistically. However, it could also lead to direct intra- or inter-category attribute-by-attribute comparisons, resulting in hypothesis testing of defining attributes.

Category assignment task (phase 2). After the cardboard screens had been placed over the study faces, the experimenter repeated the instructions as follows: "Now, as I told you before, I will be asking you to guess whether the faces I am going to show you belong to the Futura inhabitants or to their enemies. To respond, just say: 'Futura' or 'Enemy', as soon as you have found the answer". Children were then presented with the twelve new faces for categories A and B (exemplars of phase 2, Table 1), one at a time. As soon as each card had been turned up so that children could see the face drawn on it, the experimenter asked: "Who is it?" No feedback was given as to response correctness. All twelve faces were presented in a single random order. This procedure was repeated twice using a different random order each time.

Attribute identification task (phase 3). The experimenter explained that this part of the game was a little more complicated. She went on: "Now, I am going to show you different parts of faces. Try to remember if they are parts of the faces you have just seen. If you think they are, tell me if they belong to the faces of the inhabitants of Futura or to their enemies". Children were then presented with the twelve isolated facial attributes, one at a time, in random order. For each attribute, the experimenter asked: "Here is a ... (nose, chin, etc.). Have you already seen it?" When children responded "I don't know", the next attribute was presented. When children responded in the affirmative, they were asked: "Is this the ... (nose, chin, etc.) of an inhabitant of Futura or an enemy?" Response accuracy was scored. Whatever the type of response, no feedback was provided.

The three phases of the experiment were conducted in immediate succession.

Finally, for the first time the two prototypes were shown and the children were asked to classify them as the Futura inhabitants or enemies. This was done in order to control for category memory after the interfering attribute identification task. Prototypes were chosen since they were the only variants for which every attribute could perfectly

predict category membership. Children were also asked some informal questions about the way they had handled with the different tasks in the experiment.

Subject classification

Learning was operationalized as the degree of accuracy in transfer to new exemplars. In spite of the changes introduced in the Kemler Nelson paradigm[1], we followed her method of subject classification (1984: Experiment 1; note of Experiment 2). *Learners* were those subjects who made more correct categorizations than would be expected on the basis of chance (binomial probability $<.05$), which in our study was more than 17 correct out of 24 responses. The remaining subjects were labelled *Non-learners*. Then the learners who exhibited a single attribute pattern of errors (the "analytic" learners of Kemler Nelson and followers) were classified as *Type I categorizers* and the other ones (the "holistic" learners) as *Type II categorizers*. We opted for the labels Type I and Type II, which make no reference to presumed mode of processing, since one of the issues of this study was to test whether the so-called "holistic" subjects were in fact multiple attribute categorizers. Furthermore, we refer to subjects as "categorizers" – instead of "processors" as is usually done – to stress the fact that children are sensitive to differences between tasks, and use these differences to determine which types of information to process. The corollary is that the knowledge used during categorization decisions may not reflect *all* the knowledge acquired during category acquisition.

Type I responders, who relied on the information given by a single attribute, the "Focus" attribute, could be pinpointed because they categorized two faces (the conflictual faces) in the opposite family in each series of twelve transfer faces: the family A face for which the Focus attribute value was B, and the family B face for which the Focus attribute value was A. Since there were two successive presentations of the series

---

1. The two major changes we introduced vis-à-vis the Kemler Nelson paradigm are (1) the number of attributes composing faces, which is higher here than it is in general practice (six instead of four), and (2) the classification of subjects as Learners on the basis of their capacity to categorize new rather than previously studied exemplars. By virtue of these aspects, our procedure is more stringent than the usual ones.

of transfer faces, and since we allowed no more than one inconsistent response per attribute, whether or not it was the Focus attribute, the criterion for being judged as a Type I categorizer was 3 or 4 errors on the focused-upon attribute (Focus errors) and 0 or 1 error on the other attributes (Non-focus errors). The probability of obtaining this pattern of response by chance is $(n/N)^a$, where n is the number of patterns with at most one error per attribute (one correct pattern plus four patterns differing from the correct pattern by one and only one attribute), N is the number of possible patterns for one attribute (2^4) and a is the number of attributes (six). This probability is .0009. Hence the Type I pattern of responding cannot reflect a guessing strategy.

Among the Type I subjects, we differentiated a group of "Strict" categorizers who exhibited a perfect pattern of single attribute processing (i.e., 4 errors on the conflictual faces and no errors on the other faces). The remaining subjects were labelled "Loose" categorizers. Similarly, we divided Type II subjects into those who made no more than 2 errors during the category assignment task, the Strict Type II categorizers, and those who made more than 2 errors, the Loose Type II categorizers. Since the Strict and Loose levels were not equivalent in the two Types, the factor "Strictness" was not included in the statistical analysis.

Accuracy in response time measures

Response times (RT) were obtained from tape recordings (speed deviation of tape recorder = .3%). The rater was unaware of the subjects' classification and response accuracy. RT was calculated as the interval between the time the experimenter asked the question "Who?" and the time the subject initiated the reply. In order to control time measure reliability, RTs of a subgroup of subjects (n = 25) chosen at random were decoded by a second rater who was unaware of the objective of the study. The inter-rater correlation computed on the mean RTs for these 25 subjects was .97. A measure of both decoding fidelity of the rater and consistency of responses across trials was obtained by computing the correlation between the mean response latency on the first presentation of the 12 faces and the mean response latency on the second presentation. The weighted mean for the correlation coefficients calculated for Learners (Type I and Type II) and Non-learners was .785, indicating satisfactory fidelity in RT measures.

RESULTS

Learner children

1. Category assignment task

1.1. Subject classification. Fifty-two subjects (27 girls and 25 boys) were classified as Learners, and 19 (9 girls and 10 boys) as Non-learners. Learner classification and the repartition of Type I subjects according to which attribute they focused on are given in Table 2. As can be seen, the attributes which were used the most to make category decisions were the eyes and the chin, the least used being the nose.

**TABLE 2. Learner classification by type and level
of categorization responses.**

For type I responding, subjects are classified on the basis of the attribute they focused upon: hair (H), eyebrows (B), eyes (E), nose (N), mouth (M), and chin (C).

Class of subjects		Number of subjects						
Type	Level	Attribute						Total
		H	B	E	N	M	C	
	Strict	0	5	3	0	1	3	12
I	Loose	5	0	6	1	2	6	20
	Total	5	5	9	1	3	9	32
	Strict							8
II	Loose							12
	Total							20

1.2. Mean error scores. Out of 24 responses, there was a mean of 4.9 errors ($sd = .2$) in the Type I group and of 3.8 errors ($sd = .5$) in the Type II group. Note that the error score of Type I subjects corresponds to the mean of the Focus errors made by both the Strict subjects (4 by definition) and the Loose subjects, which was 5.5 ($sd = .3$): 3.8 Focus errors ($sd = .1$) because 6 subjects misclassified 3 out of 4 conflictual faces plus 1.7 Non-focus errors ($sd = .2$). No statistical

computation was made on these scores since the meaning of "errors" is not equivalent in the two Types.

The proportion of Non-focus errors for Loose Type I subjects and Strict and Loose Type II subjects, as a function of attribute type, is given in Table 3. As can be seen, the hair, the eye area, and the chin generated the greatest number of errors, and the nose, the fewest.

1.3. Mean reaction times. Table 4 (Total column) presents the mean RT calculated over the 24 category decisions. Inter- and intra-group comparisons were made after normalization of the latency distributions by the usual logarithmic transformation. They showed (1) that Type I subjects responded faster than Type II subjects, $F(1, 48) = 3.50$, $p = .67$, and (2) that in the Type I group, Strict subjects were faster than Loose subjects, $F(1, 30) = 4.93$, $p < .05$. Loose Type I and all Type II (whether Strict or Loose) subjects responded at the same rate.

2. Attribute identification task

The mean number of responses (out of 12) which fell into the "I don't know" category was not significantly different between the Type I (mean $= 3.15$, $sd = .43$) and Type II (mean $= 2.44$, $sd = .66$) groups; nor it was between the levels in each group: in the Type I group, it was 2.73 ($sd = .89$) for the Strict and 3.40 ($sd = .44$) for the Loose subjects, and in the Type II group, it was 1.62 ($sd = .84$) for the Strict and 2.98 ($sd = .93$) for the Loose subjects. To test whether these responses represented an absence of attribute value knowledge, we assumed that during the category assignment task, the unidentified attributes did not play any informative role and thus did not generate errors. To test this hypothesis, Pearson product correlations over attributes were computed between the mean score for the "I don't know" responses and the mean error score for the categorization responses (Table 3). The coefficient calculated for all learners pooled was -.73 (-.56 for Type I and -.90 for Type II subjects). In other words, the less an attribute created errors during phase 2, the more it elicited "I don't know" responses during phase 3, for Type I as well as for Type II responding.

Because the number of "I don't know" responses varied from child to child, we chose to test for attribute knowledge by calculating a difference score, the D-score (as in Perruchet & Pacteau, 1990), which was defined as the number of correct identifications minus the number of false identifications. The results are given in Table 5. A score close to zero corresponds to random responding and a score above zero means

that subjects had encoded one or more attributes. Both Type I (average D-score, when computed over Strict and Loose subjects was 2.94, *sd* = 3.17) and Type II (D-score = 3.55, *sd* = 3.80) children produced scores different from chance (Type I: *t*(31) = 5.25, *p* < .001; Type II: *t*(19) = 4.29, *p* < .001). Thus Type II subjects did not differ significantly from Type I subjects in their attribute knowledge.

TABLE 3. Proportion of errors in the category assignment task (Err.) and of unidentified attributes (NI, corresponding to "I don't know" responses) in the attribute identification task, according to attribute type.

For the type I subjects, the Focus errors, which are consistent with their strategy, are not included.

Attribute	Type I				Type II			
	Strict		Loose		Strict		Loose	
	Err.	NI	Err.	NI	Err.	NI	Err.	NI
Hair	-	.21	.18	.10	.22	-	.25	.14
Eyebrows	-	.24	.35	.25	.22	.15	.21	.11
Eyes	-	.03	.12	.18	.28	.15	.20	.08
Nose	-	.30	-	.21	-	.31	.06	.34
Mouth	-	.16	.15	.16	.14	.31	.15	.14
Chin	-	.06	.20	.10	.14	.08	.13	.19

TABLE 4. Mean response times (ms) as a function of response accuracy in the category assignment task.

Group of subjects	Correct responses		Errors		Total	
	M	SE	M	SE	M	SE
Type I						
Strict	1416	151	1725	278	1468	155
Loose	2330	276	2689	433	2404	352
Type II						
Strict	2092	445	3471	997	2225	491
Loose	2720	373	2383	478	2637	394

TABLE 5. D-scores by attribute type in the attribute identification task.

For the Type I group, D-scores for the Focus attributes are not included.

Attribute	Type I				Type II			
	Strict		Loose		Strict		Loose	
	M	SE	M	SE	M	SE	M	SE
Hair	.42	.39	.47	.30	.25	.69	.75	.39
Eyebrows	.00	.50	.05	.17	.50	.38	1.17	.34
Eyes	1.00	.40	.29	.20	1.25	.53	1.42	.26
Nose	.50	.33	-.32	.29	-.25	.45	.50	.23
Mouth	-.09	.41	.17	.27	-.50	.53	.08	.39
Chin	.44	.44	.36	.20	1.38	.42	.25	.38
Total	2.27	1.26	1.02	.63	2.63	1.12	4.17	1.22

Two comments should be made at this point about Type I responding. First, all but 6 Loose subjects perfectly identified the attribute upon which they focused during the category assignment task. Second, Type I subjects were able to identify other attributes than the one focused upon, as shown in Table 5. When the D-scores for the focused-upon attributes were dropped, their performance still differed from chance, $t(31) = 2.02, p = .052$.

When D-scores were calculated as a function of attribute type, they were found to differ from zero with a probability below .05 for the eyes (Strict Type I, and Strict and Loose Type II subjects), the eyebrows (Type II Loose subjects) and the chin (Strict Type II subjects). This result was obtained despite the differential selective weighting of attributes by subjects, which results in a small sample size. The high performance in identifying the eyes and the eyebrows, whatever the mode of processing, together with the finding that 14 out of 32 Type I subjects focused upon the eyes or the eyebrows during the category assignment task, suggests that the eye area was particularly salient in our study.

To sum up, as hypothesized, Type I and Type II subjects (whether Strict or Loose) had more attribute knowledge than suggested by their assignment to a mode of categorization.

3. Relations between face categorization strategies and attribute knowledge

3.1. Attribute identification performance as a function of Non-focus Type I and Type II errors. The observation of attribute knowledge in both types of categorizers led us to wonder whether this knowledge was used for making category judgments. One way to tackle this question is to assume that categorization errors stem from the use of attribute knowledge which is not congruent with pure Type I or Type II strategies. Therefore, it was hypothesized that in the Type I group, Loose subjects could temporarily rely on another attribute than the Focus one, and that in the Type II group, Strict and Loose subjects' category errors could arise from the use of a single attribute value, rather than from the use of a majority rule or a holistic strategy. To clarify this point, correlations over attributes were computed between the mean error score on the category assignment task and the D-score on the attribute identification task. The correlation coefficient was .39 for the Loose Type I group, and .58 for the Type II group (Strict: .54; Loose: .56). In other words, the more an attribute created errors during phase 2, the more it was likely to be correctly identified during phase 3. This relation between face categorization errors and attribute knowledge was even stronger for the Strict Type II subjects, since all the attribute values accountable for categorization errors were correctly identified.

3.2. Attribute knowledge and RTs in phase 2. All subjects, whatever their strategies, pinpointed some attribute values which guided their judgment on category membership during phase 2. What differentiates Type I and Type II strategies is that for the former, category decisions were almost totally guided by the same single attribute during the entire task, whereas for the latter, decisions were only occasionally guided by a single attribute, which moreover differed across responses. Does this mean that children were insensitive to other information than that used to base their decisions on? One way of testing this is to determine whether subjects' RTs varied as a function of response accuracy. Following Ward and Scott (1987), we reasoned that an increase in RT on wrong responses implies that subjects are sensitive to some information which conflicts with that given by the attribute upon which they responded, whether temporarily (Type II) or not (Type I). Because of doubt, subjects might show longer response latencies. To test this assumption, we conducted a 2 (Type: Type I vs Type II) by 2 (response accuracy: False vs Correct responses) ANOVA, with type as a between-

subjects variable and response accuracy as a within-subject variable. There was a significant effect for both factors (Type: $F(1, 50) = 4.20$, $p < .05$; response accuracy: $F(1, 50) = 3.81$, $p = .057$) and no interaction.

Further comparisons showed that in the Type II group, there was a significant interaction between the level (Strict vs Loose) and response accuracy factors, $F(1, 18) = 11.18$, $p < .005$. There was no such interaction in the Type I group. Inspection of Table 4 (error and accuracy columns) reveals that all but Loose Type II subjects tended to take longer when giving false than when giving correct responses.

What differentiates Strict from Loose subjects in the Type I group might be the power of the focused-upon attribute to overshadow the other encoded attributes. For Strict subjects, the weight of this attribute was likely to be strong enough to govern all decisions, in spite of multiple attribute knowledge. For Loose subjects, the multiple attribute knowledge could hang on the focused-upon attribute on some occasions (when making Non-focus errors). Our prediction was that the conflicting information arising from both Focus and Non-focus attributes could have produced response hesitation, and thus longer RTs for Non-focus errors (1.8 per subject) than for Focus errors (3.7 errors per subject). Our prediction proved to be correct, in that Loose subjects took longer to respond, $F(1, 17) = 7.13$, $p < .02$, when using a Non-focus attribute (RT = 2368 ms) than a Focus attribute (RT = 1587 ms). When using the focused-upon attribute, their RT no longer differed from the Strict Type I RTs.

4. Prototype classification. All but one subject (a Loose Type I subject) correctly classified the two prototypes which they never saw before.

5. Verbal reports. In the Type I group, most of the subjects (both Strict and Loose) were able to name which attribute they focused upon during the categorization task; some even mentioned a second one. In the Type II group, most of the children cited one or more attributes to explain their categorization strategy.

Non-learner children

For the category assignment task, non-learners obtained a mean error score of 12.2 ($sd = .5$) out of 24 responses, which corresponds to random responding. They took longer to respond than learners (2617

ms). Furthermore, on the attribute identification task, 25.4% of their responses were "I don't know" responses. Their D-score, equal to .1, was not different from chance. Finally, these children were not able to give any information about how they made their category decisions.

CONCLUSION

Two types of categorizers were distinguished on the basis of the number and patterns of errors they produced during the category assignment task: categorizers who never or rarely departed from a single attribute strategy (referred to as Type I subjects) and categorizers who did not rely on a single attribute (Type II subjects). All these children except one (a Loose Type I) were able to correctly classify the prototypical faces that had not been shown before and were only presented after the interfering attribute identification task. This high level of performance shows that they had built consistent categorical representations. The prevalence of subjects classified as Type I suggests that in 9-year-old children, as in adults (Kemler Nelson, 1984; Ward & Scott, 1987), grouping on a single attribute is predominant when schematic faces are being categorized under conditions of intentional learning, even with a family resemblance category structure.

In addition to the usual analysis of the error index, we measured category decision speed, and accuracy in attribute identification, in order to evaluate knowledge of facial units and understand what strategies underpin Type I and Type II response patterns. Our main concern was how Type II subjects really dealt with the category assignment task. We suspected that in the Kemler Nelson studies on cartoon faces (as in other similar studies), the patterns which are usually thought to result from undifferentiated, gestalt-like processes could be the end-product of a partial recombination of knowledge units. The observation that Type II subjects were able to identify a fair number of single attributes, as many as Type I subjects, supports this hypothesis.

A plausible counterargument is that faces may have been processed as Gestalts during phases 1 and 2 of our experiment and that the attribute knowledge evidenced during phase 3 reflects the outcome of redintegrative processes (Horowitz & Prytulak, 1969). In these processes, some elements of an episode (objects or events) have the power to instantiate the representation of the entire episode in memory. In other words, Type II subjects may not differentiate any attribute in the course of learning and testing, and their performance on the attribute identifi-

cation task may have been mediated by direct access to Gestalt representations of prototypes or exemplars. Data analysis, however, ruled out the redintegrative hypothesis here. First of all, Type II subjects were slower to classify than Type I subjects in the category assignment task, which suggests the sequential partial recollection of attributes when responding. Secondly, their false responses were prompted by attributes correctly identified during phase 3. Importantly, the data showed that the eight Type II learners who made no more than two errors during the category assignment task behaved like the other Type II learners: they had long response times and good knowledge of individual attributes. More significantly, all the errors made during the category assignment task were due to known attribute values, and latencies were longer with incorrect than with correct responding. The latter finding indicates that when information arising from the attributes they had differentiated was conflictual, the Strict subjects eventually made a decision dominated by a single attribute, in spite of their reluctance to do so. In sum, the Strict Type II categorizers, like the Loose ones, were using a multiple attribute strategy, but more efficiently since the number of errors they made was lower and their response times differed as a function of the correctness of their decision; they were more confident when giving correct responses than the Loose subjects, who were equally slow for correct and for incorrect decisions. In summary, in spite of the small sample size, the analysis of Strict Type II children's performance runs counter to the possibility that they processed other units of knowledge than Loose children and consequently, that the differentiation of attributes observed in the Type II strategy can be linked to a low level of learning.

The second concern of the present paper was to evaluate the amount of attribute knowledge of Type I subjects. Our assumption was that they can learn more than is necessary for a single-attribute strategy. The effect of response correctness on decision latencies and the fair number of well-identified attributes lends weight to this assumption. Extra knowledge lengthened response times when subjects were confronted with conflicting information that led them to incorrect category decisions. This was true for all Type I subjects. Among this group, what differentiated Strict from Loose children was that the former were not deterred from their single attribute strategy in spite of their multiple attribute knowledge. On the other hand, Loose categorizers were reluctant to rely exclusively on a single attribute, as shown by their longer response latencies for incorrect as well as correct decisions. The further analysis of latencies for wrong categorization responses showed first, that response times arising from Focus errors did not differ between

Loose and Strict categorizers, who were equally troubled when there was non-congruent information between the focus and non-focus attribute values; second, Loose categorizers were even more troubled when basing a category decision on a non-focus attribute.

Overall, these findings suggest that subjects who had been classified as different on the basis of their error patterns in the Kemler Nelson paradigm may rely on qualitatively similar modes of processing. In our study, the strategies that we labelled Type I and Type II are likely to be the end points of a continuum expressing the number of attributes subjects use. Note that all groups of learners, whatever their error pattern, acquired the same sort of attribute knowledge. What differentiated them was the number of attributes they took into account to make their category decision. The between-group order for mean response times fits nicely with the observation that response times were the longest when the number of attended-to attributes increased: one attribute for Strict Type I, one principal plus some secondary for Loose Type I, or a few of equal importance for Strict and Loose Type II. The long response times of the non-learner subjects may indicate that they tried to find and apply some logical category principles. Perhaps, the complexity of the stimuli and the category construction rule may have resulted in a high proportion of subjects failing to learn the category.

At a methodological level, our data point to the necessity of using different indices – such as the ones used in this study – to characterize modes of processing. At a theoretical level, they raise the issue of whether object processing always involves integration of component information; in other words whether, in the final analysis, what is identified as holistic responding is in fact underpinned by analytical operations.

Indeed, our findings are congruent with the data from Martin and Caramazza (1980), Ward and Scott (1987), and Ward et al. (1990) showing that children can use analytical strategies when forming ill-defined categories. These authors posited that the primary mode of categorization is the formulating and testing of deterministic rules to decide upon category membership. Martin and Caramazza (1980) made the assumption that these rules differ across subjects and that typicality or probabilistic processing strategies can be falsely diagnosed by the averaging of responses over subjects. The Ward studies (Ward & Scott, 1987; Ward et al., 1989) indicate that both children and adults tend to adopt an analytic approach in category learning, an approach which varies in complexity according to the subjects' level of development. They reported that children only develop single attribute hypothesis at

age five, whereas more elaborate forms (such as multiple attribute analysis) appear at seven. Their position in favor of the exclusiveness of analytic responding in categorization tasks is challenged by data from other experimental paradigms such as restricted classification (e.g., Smith & Shapiro, 1989), and from other fields of research such as face recognition (e.g., Bruce, 1990). Our viewpoint is that in the present study, as in others, several factors could have oriented children towards construction of their category representations out of facial elements, rather than through apprehension of the family resemblance between members of the same category. First, the intentional nature of the learning conditions could have favored analytic processing (Kemler Nelson, 1984, 1988). This interpretation can be ruled out since the effect of learner intent has been shown to be negligible in face categorization tasks (Ward & Scott, 1987; Sigumura & Inoue, 1988). Second, the preference for an analytic mode of processing could be related to the use of a combination of binary attribute values with two equal sized groups. According to Medin et al. (1987), this sort of constraint may induce a dimensional type of categorization. Further work is needed to test this interpretation. A third and perhaps more crucial factor is the format of representation, which in our experiment, as in other related studies, is the schematic mode of face representation. There is some evidence that photographs lead to more global processing than schematic line drawings, especially in the field of face recognition (e.g., Hagen & Perkins, 1983; Tversky & Baratz, 1985). An experiment is in progress to test the hypothesis that genuine gestalt-like processes satisfying the different criteria of the present study may appear when photographs of faces are used instead of line drawings.

ACKNOWLEDGEMENTS

REFERENCES

Brooks, L. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 169-211). Hillsdale, NJ: Erlbaum.

Bruce, V., Doyle, T., Dench, N., & Burton, M. (1990). Remembering facial configurations. *Cognition, 38,* 109-144.

Bruner, J. S., Goodnow, J. J., & Austin, A. (1956). *A study of thinking.* New York: Wiley.

Cooper, L. A. (1982). Strategies for visual comparison and representation: individual differences. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (vol. 1, pp. 77-124). Hillsdale, NJ: Erlbaum.

Evans, P. M., & Smith, L. B. (1988). The development of identity as a privileged relation in classification: When very similar is just not similar enough. *Cognitive Development, 3,* 265-284.

Gholson, B., & Beilin, H. (1979). A developmental model of human learning. *Advances in child development and behavior, 13,* 47-81.

Hagen, M. A., & Perkins, D. (1983). A refutation of the hypothesis of the superfidelity of caricatures relative to photographs. *Perception, 12,* 55-61.

Horowitz, L. M., & Prytulak, L. S. (1969). Redintegrative memory. *Psychological Review, 76,* 519-531.

Inhelder, B., & Piaget, J. (1964). *The early growth of logic in the child.* New York: Norton.

Kemler Nelson, D. G. (1984). The effect of intention on what concepts are acquired. *Journal of Verbal Learning and Verbal Behavior, 23,* 734-759.

Kemler Nelson, D. G. (1988). When category learning is holistic: A reply to Ward and Scott. *Memory and Cognition, 16,* 79-84.

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

Kemler, D. G., & Smith, L. B. (1978). Is there a developmental trend from integrality to separability in perception? *Journal of Experimental Child Psychology, 26,* 498-507.

Lautrey, J. (1988). Inter- and intra-individual variability: What can it teach us about cognitive development? Symposium Behavioral variability in learning and cognitive processes. XXIVe International Congress of Psychology, Sydney.

Lautrey, J. (1990). Esquisse d'un modèle pluraliste du développement cognitif. In M. Reuchlin, J. Lautrey, C. Marendaz, & T.Ohlmann (Eds.), *Cognition: l'individuel et l'universel* (pp. 185-216). Paris: P.U.F.

Marquer, J., & Pereira, M. (1990). Reaction times in the study of strategies in sentence-picture verification: A reconsideration. *Quarterly Journal of Experimental Psychology, 42A,* 147-168.

Martin, R. C., & Caramazza, A. (1980). Classification in well-defined and ill-defined categories: evidence for common processing strategies. *Journal of Experimental Psychology: General, 109,* 320-353.

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

Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness and category construction. *Cognitive Psychology, 19,* 242-279.

Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General, 120,* 112-116.

Rieben, L., de Ribaupierre, A., & Lautrey, J. (1990). Structural invariants and individual modes of processing: On the necessity of a minimally structuralist approach of development for education. *Archives de Psychologie, 58,* 29-48.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27-48). Hillsdale, NJ: Erlbaum.

Rosch, E., Simpson, C., & Miller, R. S. (1976). Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 491-502.

Smith, J. D., & Shapiro, J. H. (1989). The occurrence of holistic categorization. *Journal of Memory and Language, 28,* 836-399.

Sugimura, T., & Inoue, T. (1987a). Effects of category structure on children's categorization. *Japanese Psychological Research, 29,* 120-130.

Sugimura, T., & Inoue, T. (1987b). Role of classification training on children's categorization. *Psychologia, 30,* 61-69.

Sugimura, T., & Inoue, T. (1988). Factors affecting analytic and holistic modes in children's classification learning. *Perceptual and Motor Skills, 66,* 723-736.

Tversky, B., & Baratz, D. (1985). Memory for faces: Are caricatures better than photographs? *Memory and Cognition, 13,* 45-49.

Vigotsky, L. S. (1962). *Thought and language.* Cambridge, MA: MIT Press.

Ward, T. B. (1988). When is category learning holistic? A reply to Kemler Nelson. *Memory and Cognition, 16,* 85-89.

Ward, T. B. (1989). Analytic and holistic modes of processing in category learning. In B. E. Shepp & S. Ballesteros (Eds.), *Object perception: Structure and process* (pp. 387-419). Hillsdale, NJ: Erlbaum.

Ward, T. B., & Scott, J. (1987). Analytic and holistic modes of learning family resemblance concepts. *Memory and Cognition, 15,* 42-54.

Ward, T. B., Vela, E., & Hass, S. D. (1990). Children and adults learn family-resemblance categories analytically. *Child Development, 61,* 593-605.