Thematic and taxonomic relations in preschoolers: The development of flexibility in categorization choices

Agnès Blaye*
Centre Psycle, Université de Provence, France

Françoise Bonthoux
Laboratoire de psychologie Expérimentale, Université Grenoble, France

Two experiments examined the development of flexibility of categorization in children aged from 3 to 5, using a picture-matching task. During a pre-test, spontaneous matches were observed. Children were next presented with a scene aimed at inducing an alternative match. Finally, they were given a post-test identical to the pre-test to assess whether they would maintain their initial choice. Both experiments showed that preschoolers were able to produce different types of categorization choices. In Experiment 1, children had to choose between a thematic and a taxonomic option. Five-year-olds showed flexibility and changed their mode of response more often in the presence of the scenes (experimental group) than in their absence (control group). In contrast, 3-year-olds’ responses seemed to reflect spontaneous variability since their switches were not related to contextual information. In Experiment 2, two more choice options were added. Three-year-olds’ data replicated the findings of Experiment 1. However, a majority of 4-year-olds were consistent on the thematic mode of response and adapted their responses to the scenes only to a limited extent. Overall, the data suggest a developmental path from spontaneous variability to flexibility, via a predominance of one mode of response.

Early theories of conceptual development held that young children’s categorization evolves from a thematic or perceptual organization to a taxonomic organization. Piaget described this development as a shift from figural to non-figural collections and logical classes (Piaget & Inhelder, 1959), while Nelson (1985) viewed it as a shift from thematic to slot-filler and taxonomic categories. Yet empirical studies have shown (generally in between-participants designs) that thematic, perceptual and taxonomic relations are available at an early age (e.g. Bauer & Mandler, 1989; Fenson, Cameron, & Kennedy, 1988).

Although many recent studies provide evidence that children can categorize based on
thematic, taxonomic or perceptual relations, there are few data bearing on whether they can use several categorical relations concerning the same object. Our main interest was to evaluate to what extent children can adapt their categorical choices flexibly as a function of task or situational demands. As pointed out by Deák and Maratsos (1998), young children’s lack of cognitive flexibility has been one of the most robust empirical findings in the study of cognitive development (Piaget & Inhelder, 1959; Deloache, 1987; Flavell, Flavell, & Green, 1983). Karmiloff-Smith (1992) considers increasing flexibility as one of the major achievements of human cognitive development. We believe that intra-individual flexibility is fundamental for adaptive behaviour because such flexibility allows children to switch from one response mode to another and thus to find the most appropriate solution to the problem at hand. The aim of this investigation was to examine, on a within-participants basis, the extent to which children are able to modify their initial categorization decisions to adjust to contextual information, and how this flexibility develops between 3 and 5 years of age.

Flexible categorization involves the ability to categorize an object on one basis (e.g. taxonomically) on one occasion, and on another basis (e.g. thematically) on another occasion (Deák, 1994). Hence, it requires taking into account different categorical relations involving the same object and switching between them. We investigated intra-individual adaptive flexibility between two different modes of categorization (taxonomic and thematic). The question addressed was whether a child can consider the same object (e.g. a fish) from two different points of view, as a function of the pictorial context in which it is presented. For example, a fish can be considered as a member of a taxonomic category (e.g. an animal—hence it can be associated with a dog) and as part of a schema (e.g. fishing—hence it can be associated with a fishing rod).

In a brief overview of the literature, we first review studies showing that categorical responses are constrained by the experimental conditions. Although these studies are based on inter-individual designs, they reveal some intra-individual variability across items. Secondly, we review the few studies that directly address the question of multiple categorizations of the same objects by a given child.

Recent evidence using a matching or forced choice task, in which participants are required to choose the best match to a target among several pictures, has shown that preschoolers’ choices may be predominantly taxonomic, thematic or perceptual, depending on various contextual factors. For example, naming the objects with nonsense nouns increases the proportion of taxonomic choices (Markman & Hutchinson, 1984; Waxman & Kosowski, 1990), even among children as young as 21 months (Waxman & Hall, 1993). However, when a perceptual choice is available, the label appears to increase perceptual and taxonomic choices (Golinkoff, Shuff-Bailey, Olguin, & Ruan, 1995; Imai, Gentner, & Uchida, 1994). Different instructions also appear to elicit different types of matching (choose the one ‘that is most like’ the target vs. ‘that is the same kind of thing’ as the target (Deák & Bauer, 1995); ‘which goes best’ vs. ‘find another one’ (Waxman & Namy, 1997)). In addition, training conditions (reinforcing either perceptual or taxonomic choices) significantly modify the proportions of taxonomic responses (Deák & Bauer, 1995).

These studies have led to the view that preschool children can take into account different types of relations between objects and can adjust their categorization choices as a function of experimental conditions. However, because this conclusion is based on data
collected between participants, it remains unclear whether, and under which conditions, a given child can use several categorical relations (taxonomic, perceptual and thematic) for the same object. Adaptive flexibility, as defined above, requires the ability to switch between two categorical links as a function of the situational context. Hence, it can only directly be assessed using a within-participant design.

Analysis of individual patterns in previous research (Waxman & Namy, 1997) has shown that a majority of 3- and 4-year-old children produced both thematic and taxonomic choices across different items when asked to find the one that ‘goes best with’ the target. This result suggests that several categorical relations can be activated in a given child by different sets of objects. But what about children’s ability to consider the same object as being involved in several types of relations?

Although few investigations have examined flexibility directly, a number of studies have assessed it indirectly. For example, recent studies dealing with polyonymy (applying several names to a same entity) showed that children as young as 3 years old can produce several words for the same object (Deák & Maratos, 1998; Waxman & Hatch, 1992). That is, young children can simultaneously represent an object as belonging to different levels of the same taxonomy (basic, superordinate and subordinate) or to different taxonomies according to its function or shape.

Deák and Pick (1994) reported additional evidence on a within-participant basis of multiple categorical representations of objects in 4-year-olds. The authors first showed that children could follow instructions requiring them to match pictures of ‘hybrid’ objects on the basis of either common function or common shape. When later required to label the hybrid object, most participants in both conditions produced function-based labels. This result suggests that some participants in the ‘match according to shape’ condition demonstrated flexibility between the matching and labelling tasks. All of these experiments deal with within-participant categorical flexibility in preschoolers; however, none of these studies has analysed individual patterns of change between different types of categorical relations (but see for one exception Deák, 2000).

A few studies using the matching task have also indirectly assessed flexibility. In those studies, children were presented twice with the same set of objects and were required either to make a second matching choice after their first one or to justify the non-chosen pairing. Melot and Houdé (1998) presented 3-year-old children with a contrasting taxonomic and perceptual choice (same colour). They observed that a majority of the children found the non-chosen option acceptable (however, no justifications were required). In contrast, results from other studies analysing the proportion of correct explanations of the non-chosen pair suggest that most children younger than 4 years of age are unable to justify two choices for the same target (Greenfield & Scott, 1986; Lucariello, Kyratzis, & Nelson, 1992; Smiley & Brown, 1979; Tversky, 1985). A significant improvement in flexibility was observed between 3–4 and 6–7 years. It is difficult to determine whether this rather late achievement is owing to slow development of flexible behaviour or to preschoolers’ difficulty in verbally describing different relations. Moreover, in all of these studies, the children were told explicitly to turn their attention to a different type of relation, instead of investigating how children can spontaneously recategorize a target as a function of new contexts.
The goal of the two studies reported here was to assess the development of the ability to flexibly ‘recategorize’ the same set of objects. Both experiments were run over three sessions. The first session (pre-test) was designed to measure spontaneous choices. The second session (test) addressed the issue of whether preschoolers could shift from one type of relation to another on the same series of targets, as a function of contextual information. Children were introduced to a scene before each choice, the content of which was intended to induce the mode of response that was not produced in the pre-test (e.g. thematic if the pre-test choice was taxonomic). The procedure was designed to determine whether changes in the mode of response from pre-test to test were induced systematically by the scenes or were reflections of instability over time. Only the first kind of change is considered an indicator of adaptive flexibility, while the second reflects spontaneous variability. A third session identical to the pre-test was included to assess whether children would (a) reproduce their original choice as soon as the contextual constraints were removed, or (b) adopt the relation induced by the scene in the test session. Few studies have examined this issue of stability of choices over time. In the absence of any source inducing changes in the mode of response (no contextual scenes presented on test, same drawings on pre-test and post-test), changes in the control groups between the first and third session provide a baseline rate for spontaneous variability. If changes were found to be more frequent in the experimental groups compared to the controls, this would imply that the contexts introduced in test session have an extended effect.

In agreement with recent findings, we predicted that children would produce a variety of categorization choices as early as 3 years of age (Dunham & Dunham, 1995; Waxman & Namy, 1997). However, we expected 3-year-olds to be less consistent in their relational choices, and less adaptive to contextual cues provided by the scenes, than older children (Deák, 2000; Deák, Ray & Pick, 1999). Indeed, these two abilities require a rather explicit representation (sensu Karmiloff-Smith, 1992) of the two (thematic and taxonomic) modes of response.

**EXPERIMENT 1**

Children’s performance was examined in a forced-choice task contrasting a thematic and a superordinate taxonomic matching. Each choice was instantiated through a pair of drawings to increase the probability that young children would identify the taxonomic or thematic relation underlying the proposed associations.

**Method**

**Participants**

Participants were 26 3-year-olds (14 female, 12 male; mean age = 3;9, range = 3;3–4;2) and 43 5-year-olds (20 female, 23 male; mean age = 5;10, range = 5;3–6;2). Both age groups attended preschools (1st and 3rd year). Half of the children of each age group were randomly assigned to a control group, and half to an experimental group.

**Material**

**Stimuli.** The stimuli consisted of sets of line drawings of objects (approximately 5 × 5 cm), presented on cards. Each set comprised a target and two pairs of items. One pair was thematically and the other
taxonomically related to the target (e.g. when the target was a ball, the taxonomic pair consisted of a teddy bear and a rocking horse, and the thematic pair consisted of a football player and sport shoes). Taxonomic relations were superordinate (see Appendix 1). The target was centred on the table. The two pairs were placed closer to the child on either side of the target. The two drawings of a given pair were presented on a single card and hence could not be chosen separately. There were eight targets and two pairs of choices (thematic and taxonomic) for each target. The drawings, presented as choice-pairs, were different between pre-test and test (see the procedure below).

Stimulus selection. In a control experiment, we interviewed six 4-year-old children (mean age = 4;3) to assess their familiarity with the taxonomic and thematic relations used. Two familiarization trials were aimed at showing them that pictures can ‘go together in different ways’: children were presented with one thematic and one taxonomic choice, with order counterbalanced across individuals. They were then shown each target with the corresponding thematically or taxonomically related pair. Triads were presented in a randomized order. For each triad, we asked: ‘Why do they go together?’ An explanation was coded ‘thematic’ if it related the three objects by reference to a common event or scene. It was coded ‘taxonomic’ if it referred in one way or another to the fact that the three objects were of a common sort. Children provided thematically based explanations (e.g. ‘the boy plays football with its sports shoes’) for 83.5% of the thematically related triads, and taxonomically based explanations (e.g. ‘they are all toys’) for 77% of the taxonomically related triads. This suggests that the relations were overall familiar and accessible to preschool children.

Scenes. Line drawings presented on A4 sheets constituted the contexts. These scenes were chosen to induce either a thematic or a taxonomic choice for each set (see Appendix 1). The represented scenes were designed to make the association of one of the two pairs with the target more probable than the other. For example, in a ‘toyshop’ (taxonomic scene), the probability of observing a ball and a teddy bear and a rocking horse is higher than that of observing the same ball and a football player and a pair of sport shoes. None of the stimuli objects were depicted in the scenes.

Scene selection. Nine additional participants (mean age = 4;5) were interviewed to assess whether the scenes activated the intended matching. They were asked 16 questions corresponding to the eight targets presented in two different scenes. For each target, children were first presented with the scene drawings, and the target was then stuck on the picture. The experimenter asked: ‘Imagine you find yourself in this place [pointing to the whole scene] and you see this object [pointing to the target in the scene], what else would you probably see, this [pointing to one choice option, e.g. the thematic pair] or this [pointing to the other choice option, e.g. the taxonomic pair]?’ The order of pointing to the options was counterbalanced, and the order in which targets were presented was randomized. Overall, there were 87% of expected choices (86% for the thematically oriented scenes and 88% for the taxonomic ones). These results confirm that preschool-age children interpret the scenes as we do.

Procedure

Children were tested individually in a quiet room in their preschools. There were three sessions, each separated by a one-week delay. For the experimental group, the first and last sessions were identical (pre- and post-tests), whereas the second session (test) differed. Each set of the forced-choice task was preceded by the presentation of a scene, which was intended to induce the non-produced choice on this specific item in the pre-test. For the control group, the procedure was the same except that in the second session (test phase), no scenes were presented before each item. All sessions involved eight sets of drawings.

The positions of the two pairs of choices were counterbalanced across trials. In a given session, the order of presentation of the sets varied randomly among children. From one session to another, each

1 We used Waxman and Namy’s (1997) procedure. As noted by these authors and many others, 3-year-olds are much less able than 4-year-olds to articulate explanations for pairing that are nevertheless familiar to them.
participant saw the series in a different order. The same targets were presented during the three sessions. The same pairs of choice drawings were presented on the pre- and post-tests, but different pairs of choice objects were presented in the test session to avoid immediate retrieval of a previously activated association between a target and a particular pair.

Pre-test. To ensure that the drawings were correctly identified, the participants were asked to name each of them before being instructed to choose. Most participants could name all the drawings; if not, they were told the name generally used by their peers. They were then asked to choose the pair that went ‘best with’ the target.

Test. In the experimental group, children were presented with a scene drawing before each forced choice: ‘We will play the same game as last week, but this time I will show you a large drawing first. It will help you in making your choice.’ This last sentence was mentioned to the children only once at the beginning of the test session to explain why they were required to look at a scene before being presented with a new set. The scene was presented alone a few seconds before the matching task itself; it was then left on the table without any further mention during the participants’ choices. Children were asked to choose the pair that went ‘best with’ the target (same instructions as in the pre-test). The selection of the scene to be presented (either thematic or taxonomic) was determined by the pre-test performance: the selected scene was supposed to activate the categorical relation not chosen on the pre-test. The children in the control group were presented with the same series of targets and choices as experimental group, but no scenes were presented before the choices. The instructions were the same as in the pre-test.

Post-test. The post-test was identical to the pre-test for all participants, except that children were not asked to name the drawings.

Results

Children’s choices were examined set by set. The mean percentage of participants making a taxonomic choice when all sets were collapsed was 52%. The mean for each set was within two standard deviations of this overall mean. This suggests that each selected set could induce both choices.

Statistical analyses were performed on the pre-test session to examine the diversity of choices and individual preferences. The mean proportion of choices revealed no bias (either thematic or taxonomic) in either age group: the proportions of taxonomic (or thematic) choices did not differ significantly (t-tests) from chance (50%). The results of the two following sessions indicated the same lack of predominance of one choice over the other.

Turning now to individual patterns, the proportion of participants exhibiting a preference for one mode of response (seven or more responses of the same type out of eight, binomial test, \( p < .04 \)) was very low in both age groups: 3.8% (1/26) at 3 years and 16% (7/43) at 5 years old. It is noteworthy that both taxonomic (5) and thematic (3) preferences were observed.

To test the effect of the introduction of scenes in the experimental group, a two-way ANOVA was conducted on the number of changes\(^2\) between pre-test and test (Fig. 1a), with condition (2) and age (2) as between-participant’s factors. As hypothesized, a main effect of condition was observed (\( F(1,65) = 21.29, p < .0001 \)). Participants shown the

\(^2\) As the pairs of choices were different in the two sessions, the specific drawings chosen were necessarily different. What we were interested in was the number of changes that corresponded to shifts between modes of responses (from taxonomic to thematic or vice versa), not between specific associations for a given target. Hence, the number of changes could vary between 0 and 8 for a given individual.
Scenes switched to the other relational choice on a majority of sets (57%: $M = 4.57$, $SD = 1.44$), while control participants tended to maintain their pre-test choice type (changes on only 36% of the sets: $M = 2.85$, $SD = 1.30$). There was no main effect of age but a significant interaction between condition and age ($F(1,65) = 8.67, p < .005$); the difference between the experimental and the control group was significant only for the older participants (60%: $M = 4.77$, $SD = 1.38$; and 29%: $M = 2.33$, $SD = 1.11$ of changes, respectively, in experimental and control group; post-hoc Scheffé test, $p < .002$). Both groups of 3-year-olds changed their mode of response on about half of the sets (53%: $M = 4.23$, $SD = 1.53$ in the experimental group; and 46%: $M = 3.69$, $SD = 1.18$ in the control group). In contrast, 5-year-old control participants maintained their past response on a majority of the sets, although choices were instantiated with different drawings in pre-test and test. Thus, the scenes had systematic effects at 5 years of age.

Moreover, although the small number of children exhibiting a preference for one mode of response does not allow statistical analyses, it is worth noting that among 5-year-olds, three out of four of the children of the experimental group showing a preference on pre-test changed their mode of response on average on six out of eight sets from pre-test to test, while those with no preference changed on average on 4.72 sets. In the 5-year-old control group, the mean number of changes was .66 among those who had shown a preference on pre-test and 2.33 among those with no preference. Preferences for one mode of response corresponds to an unequivocal interpretation of the instructions on the whole series of sets. Such an interpretation on the part of these children suggests that they have a rather explicit representation of the common sorting criterion behind their choices, and hence, they show adaptive flexibility, sticking with their initial choice in the control group and, in the experimental group, switching to the choice induced by the pictorial context on most of the sets.

The delayed post-test assessed whether children returned to their initial choice when placed in the same conditions. A 2 (condition) × 2 (age) ANOVA was performed on the number of changes (max. = 8) from pre-test to post-test (Fig. 1b). Overall, younger participants reproduced their initial choices less frequently than older participants (52.4% vs. 71.7% in 3- and 5-year-olds, respectively; $F(1,65) = 14.5, p < .0003$); they changed their choices on about half of the sets. There was no main effect of condition, but there was a significant interaction between age and condition ($F(1,65) = 5.08, p < .03$). The difference between the experimental and control groups for the older participants fell short from significance (post-hoc Scheffé test, $p < .09$). Hence, despite a two-week delay between pre-test and post-test, 5-year-old control participants reproduced their initial choice on about 80% ($M = 6.38$, $SD = 1.36$) of the sets vs. 64% ($M = 5.09$, $SD = 1.74$) in the experimental group. This suggests a delayed effect of the introduction of the scenes. In the experimental group, children who had a preference for a relational choice on the pre-test tended to go back to their initial choices more often than others (72% of the sets vs. 62% of the sets).

**Discussion**

Results support both predictions. First, taxonomic and thematic matchings on the pre-test were observed in similar proportions at both age levels. This strengthens the claim
that both types of relational link are available at an early age. Such a conclusion is supported by individual patterns: the majority of children produced both modes of response over a series of targets in the same session. Although our stringent definition of consistency on the pre-test (seven or eight sets out of eight, with the same mode of response) does not enable us statistically to assess an increase of consistency with age, the data were congruent with the expected difference.

Regarding flexibility of categorization, the results suggest that 5-year-old children took into account the information provided by the scenes in making their categorization decisions. They tended to maintain the same mode of response on a given target when no extra information was provided (control group), whereas they changed from one mode of response to another on a majority of sets when presented with the scenes (experimental group). Moreover, when confronted twice (pre-test and post-test) with the same sets of stimuli, 5-year-olds were significantly more consistent with their choices than 3-year-olds. Furthermore, older participants’ sensitivity to the contextual cues appears to last over a one-week delay: experimental participants tended to stick to the mode of response induced by the scenes and hence to reproduce their pre-test mode of response less often that control participants.

Three-year-olds’ fluctuations of choices reveal a kind of variability that does not appear to be related to the introduction of the scenes. Given that choices were instantiated by different pictures in pre-test and test, this suggests that these children make their decisions on the basis of the salience of specific associations of pictures.
Moreover, as previously mentioned, their frequent changes between pre-test and post-test suggest that the degree of salience of specific associations is rather unstable (see Barsalou, 1987, 1993; Yeh & Barsalou, 2000, for similar examples of the instability of adults’ categorical judgments).

In summary, changes in matching from one session to the next appear to be of a different kind in 3- and 5-year-olds. The fluctuations of choices of the younger participants could be interpreted as spontaneous variability of choices. On the contrary, older participants’ changes varied as a function of the presence/absence of the scenes and hence could reflect adaptive flexibility. Strategic behaviour of 5-year-olds is also supported by results of the control group.

However, as there were only two types of choices, any observed change was necessarily congruent with the relation induced by the scene. Thus, changes that seemed to follow from the scenes might result either from an effective processing of the content of the scenes or from a particular understanding of the task demands (i.e. the presence of scenes requires children to do something different than in the pre-test). The following experiment was designed to distinguish between these two interpretations.

**EXPERIMENT 2**

To examine the developmental process more precisely, the age range was reduced to one year; 3- and 4-year-olds were tested. To distinguish between spontaneous and adaptive switching, children were offered four choices instead of two. One of the new choices was related perceptually to the target, and the other was unrelated to the target (irrelevant choice). Because children were presented with four choices, it was possible to measure whether shifts were congruent with the scenes (e.g. a shift from a thematic to a taxonomic response when the scene induced a taxonomic choice) or not (e.g. a shift from a thematic to a perceptual response when the scene induced a taxonomic choice).

In Experiment 1, the tendency of 3-year-old control participants to modify their responses even in the absence of any contextual scene suggested that offering them different choices between pre-test and post-test may have induced some variability of responses that does not reflect the kind of flexible behaviour we are studying. Hence, in Experiment 2, we used exactly the same sets of pictures in the three sessions.

**Method**

**Participants**

Participants were 30 3-year-olds (15 female, 15 male; mean age = 3;6, range = 3;2–3;11) and 28 4-year-olds (14 female, 14 male; mean age = 4;8, range = 4;1–5;3) from two preschools (1st and 2nd year).

**Material**

*Stimuli.* The stimuli were 10 sets of black and white photographs of highly familiar objects. Each photograph (5.5 × 4.0cm) contained a single object centred on a light grey background. A set of pictures consisted of five photographs (see Appendix 2): a target (e.g. an apple), a taxonomic choice (e.g. a banana), a thematic choice (e.g. a knife), a perceptual choice (e.g. a ball) and an unrelated picture (e.g. boots). Photographs were used, rather than drawings, in order to increase the realistic aspect of stimuli. Because the greater number of options (as compared to Experiment 1) increased the information processing requirements of the task, each option was instantiated by only one object (instead of two as in
The four choices were arranged horizontally on white A4 cardboard under the target. Their positions were counterbalanced across sets and participants.

**Similarity ratings.** Twelve students enrolled in psychology courses rated the similarity between the target and each of the four options to ensure that the perceptual choices were more similar to the target than the other three options. Participants were instructed to rate each pair using a scale from 1 = extremely dissimilar to 9 = extremely similar. They were told that they should focus on the appearance of the objects themselves and should ignore their categorical membership.

The mean similarity rating across the 10 sets was 6.5 for the perceptual choices; it was significantly higher ($p < .01$) than for each of the three other alternatives (2.7 for the taxonomic choices, 1.8 for the thematic choice, and 1.7 for the unrelated objects). These differences were also reliable in each set.

**Stimulus selection.** We interviewed six 4-year-old participants (mean age = 4;6), with the same procedure as in Experiment 1, to assess children’s familiarity with the taxonomic and thematic relations depicted above. Children provided thematically based explanations for 97% of the thematically related pairs\(^3\) and taxonomically based explanations for 76.5% of the taxonomically related pairs. This suggests that the relations were generally familiar and accessible to them.

**Scenes.** Black and white 8 $\times$ 12 cm photographs of familiar scenes served as contexts (see Appendix 2). These scenes were intended to induce either a taxonomic or a thematic choice for each of the 10 targets.

**Scene selection.** The same group of nine participants as in Experiment 1 was presented with the 20 scenes and targets from this second experiment. The same procedure as in Experiment 1 was followed to assess children’s interpretation of the scenes. In all, 82% of expected choices were obtained (86% for the thematically oriented scenes and 78% for the taxonomic scenes).

**Procedure**

Children were tested individually in a quiet room in their preschools over three sessions, each separated by a one-week delay. The same 10 sets of pictures were used in all three sessions. The instructions and procedure were the same as those of the experimental group in Experiment 1. As there were only two kinds of scenes (thematic and taxonomic) and four choice options, the presented scene in the test was alternatively taxonomic or thematic for perceptual or irrelevant matching in the pre-test. The four choice options presented with each target made the task rather demanding for 3- and 4-year-olds. Thus, we decided to use the pre-test as a selection criterion to discard participants who could not process the pictures in a sensible way, that is they would select the irrelevant choice as the best associate.

**Results**

A first analysis of children’s pre-test responses revealed that 10 participants of each age group produced at least one irrelevant choice. These participants did not participate in the next sessions, and their performance on pre-test were not included in further data analyses. Children’s pre-test choices were then examined set by set to ensure that the nine sets (see footnote 3) were roughly equivalent regarding the mode of response they elicited. For all sets, the number of each type of matching deviated less than two standard deviations from the mean (collapsed across sets).

\(^3\) An additional 3-year-old control group ($N = 12$) assessed recognition of the target relations. They were presented with each target twice (once with the taxonomic and once with the thematic associate) and had to point to the choice (thematic or taxonomic vs. two unrelated choices) that ‘went well’ with the target. They were not asked for verbal explanations. All the relevant pairs were recognized by at least 75% of the children, except a thematic pair relating a cherry with a ladder that was recognized by only 40% of the sample. Hence, the ‘cherry’ set was removed from the data analysis.
An ANOVA was performed on pre-test choices, with age (2) as a between-participant factor and type of pertinent choice (3) as a within-participant factor. There was a significant effect of the mode of response ($F(2,72) = 7.64, p < .001$). The interaction between age and mode of response approached significance ($F(2,72) = 2.72, p < .08$). In 3-year-olds, proportions of each mode of response did not differ (Scheffé test), despite a tendency to produce more thematic responses (39% for thematic: $M = 3.5$, SD = 2.7; 30% for taxonomic: $M = 2.7$, SD = 2.0; and 28% for perceptual matchings: $M = 2.6$, SD = 2.1). In contrast, proportions of choices differed significantly in 4-year-olds, with thematic choices being more frequent (58%: $M = 5.2$, SD = 2.9) than taxonomic (20%: $M = 1.8$, SD = 2.2; Scheffé test, $p < .03$) and perceptual choices (19%: $M = 1.7$, SD = 2.4; Scheffé test, $p < .03$). A similar pattern of results was obtained in the next two sessions. This is consistent with the predominance of thematic choices found in previous experiments (Greenfield & Scott, 1986; Smiley & Brown, 1979; Waxman & Namy, 1997).

Regarding individual preferences for one categorical relation in the pre-test, participants were scored as having a preference if they consistently produced the same mode of response on different sets (six sets or more out of nine, multinomial test, probability calculated for one among three relevant matchings, $p < .04$). Most 4-year-olds showed a preference (14 children among 18, 78%), and more 4- than 3-year-olds (five children among 20 (25%), $\chi^2 (1) = 8.55, p < .005$, Yates' correction) showed a preference (see also Waxman & Namy, 1997). In congruence with previous research (Golinkoff et al., 1995), thematic preferences were more frequent than taxonomic or perceptual preferences (13 among 19 children showed a thematic preference, age groups combined). In the older group, only one child showed a taxonomic preference and two a perceptual preference, whereas in the younger group, two showed a taxonomic preference and one a perceptual preference.

We now turn to our main point of interest by analysing first whether children changed their mode of matching from pre-test to test and, second, to what extent these changes were congruent with the scenes (Fig. 2). Younger children changed their mode of response on 41% of the sets ($M = 3.7$, SD = 2.05; not different from chance), whereas older children changed on only 31% of the sets ($M = 2.8$, SD = 2.4), which is significantly below chance ($t(17) = 3.03, p < .01$). This suggests that 4-year-olds tended to stick to their original response more strongly than 3-year-olds (note, however, that the difference in proportion of changes between 3- and 4-year-olds did not reach significance). Moreover, 46% (SD = 18.5) of the changes were congruent with the scenes in 4-year-olds as opposed to 37% (SD = 21.3) in 3-year-olds. This proportion was above chance for the older participants ($t(13)^4 = 2.62, p < .03$), but not for the younger participants. Thus, the effect of the scenes seems more systematic in older children.

Three-year-olds who showed a preference on pre-test switched in a scene-consistent way in the same proportion (37%) as those who did not show a preference. However, 4-year-olds who had a preference tended to change more often as a function of context

---

4 As there were four choices, the probability to observe a change congruent with the scene, by chance, was 33%. Four children were not included in this comparison since they made no changes at all between pre-test and post-test.
(49% vs. 39%) than those who had no preference. The small number of participants did not allow statistical analysis.

The analysis of the stability of matchings between the two identical sessions, pre-test and post-test, showed that older children reproduced their initial choice more often than younger children (73%: $M = 6.6$, $SD = 2.2$; and 55%: $M = 5.0$, $SD = 2.0$, of the nine sets at 4 and 3 years of age, respectively; $t(36) = 2.36$, $p < .03$). Moreover, at both ages, children who showed a preference on the pre-test were more likely to maintain their responses on the post-test (79.6% of the sets for those who had a preference on the pre-test: $M = 7.2$, $SD = 2.0$ vs. 47.3% for those with no preference: $M = 4.3$, $SD = 1.4$; $t(36) = 2.56$, $p < .02$).

**Discussion**

The data of 3-year-olds replicate and extend those obtained in Experiment 1. Children’s choices were distributed in a roughly equivalent manner across the three relevant options, whereas 4-year-olds produced a majority of thematic responses. Such an increase of thematic matching from 3 to 4 years is consistent with Waxman and Namy’s findings using the same instruction (Waxman & Namy, 1997). Furthermore, and consistently with Waxman and Namy’s data, there was an increase with age in the number of children exhibiting a preference for one mode of response.

A major difference in the procedure of our two experiments was that pictures offered as choices remained the same between pre-test and test in Experiment 2 while they differed in Experiment 1. This could explain the lower proportions of changes from pre-
test to test in Experiment 2 compared to Experiment 1 (experimental groups). More interestingly, this rules out the hypothesis that the introduction of scenes *per se* is understood by preschoolers as a requirement to do something different than in the pre-test. Hence, it favours an interpretation of changes as resulting from an effective processing of the content of the scenes.

Three-year-olds’ performance strengthens our interpretation of Experiment 1: changes at this age level are likely caused by spontaneous variability of behaviour; they do not reflect an ability to adaptively respond to the new information contained in the scenes. Conversely, one of the most striking characteristics of 4-year-olds’ categorization responses is the internal coherence of their choices (mainly thematic). Among the different types of available relations, children of this age select one and consistently apply it both within the same session and from one session to the next. However, these children proved able to adjust their categorical choices to contextual cues, if only to a limited extent.

**GENERAL DISCUSSION**

These two experiments challenge hierarchical theories of categorization development that view development as progressing from schematic to taxonomic categories (Nelson, 1985; Piaget & Inhelder, 1959). They provide further evidence that preschool children can take into account various relations between objects (see also Deák & Bauer, 1995; Greenfield & Scott, 1986; Smiley & Brown, 1979; Waxman & Namy, 1997).

The present research was designed to explore the development between 3 and 5 years of age of categorical flexibility, which is the ability to switch from one categorical relation to another for a same object, as a function of contextual cues. While much previous empirical work has shown that children of this age can be induced to use one type of relation or the other (e.g. Deák & Bauer, 1995; Markman & Hutchinson, 1984; Waxman & Namy, 1997), there is less direct evidence of within-participant flexibility. The main contribution of the present experiments then is to sketch the developmental evolution of categorical flexibility.

The results of both experiments reveal that 3-year-old children have a different response to that of older participants to task requirements. Their categorical choices were evenly distributed among the different options. These data are congruent with results obtained by Deák *et al.* (1999) showing that when instructions are non-specific ('find the one that goes best with'), 3-year-olds do not sort consistently on one criterion (in their study, either shape or function). In the task presented here, the sorting rules were not stated explicitly. Hence, to have a unique interpretation of the instructions on the whole series of sets, children must represent the common nature of the relation underlying such diverse pairings as ‘snowman/ski sticks’ and ‘bowl/spoon’. Three-year-olds probably do not reason about the type of relational link. They are more likely to produce specific responses to each forced choice depending on the salience of one pairing over the others at the moment of testing.

Moreover, 3-year-old children give up their initial response as often as they maintain it, independently of the presence or absence of a contextual scene. This suggests a lack of adaptive flexibility that cannot be accounted for only in terms of inefficient inhibition of previous responses. Similarly, Zelazo, Frye, and Rapus (1996) found that most 3-year-
olds failed to switch between two simple colour and shape sorting rules even when explicitly told to do so. Deák et al. (1999) and Deák (2000) also found that 3-year-olds had difficulties in switching between form and function sorting criteria or in generalizing new words referring either to shape, material or parts of the targets. Together, the instability of responses from one set to the other, and the lack of correspondence between switches on a given target and the introduction of scenes, suggest that 3-year-olds’ pattern of responses correspond to spontaneous variability rather than to flexibility.

Performance of 4-year-olds contrasted with 3-year-olds’ variability of choices in a number of respects. Four-year-olds displayed rather stable behaviour even when presented with new information designed to induce them to change their answer. This result was shown by a preference (within-session) for one mode of response—namely thematic—in a majority of children and by a majority of identical responses from one session to the next. Instead of categorizing on a case-by-case basis, as 3-year-olds appear to do, 4-year-olds appear to rely consistently on a particular type of relational link. Moreover, although they manifested a clear preference for thematic relations, when they switched their responses, they did so in a more principled way than 3-year-olds, taking into account the information provided by the scenes. Consistency of sorting choices was also observed by Deák et al. (1999) in a majority of 4-year-olds. Moreover, Deák et al.’s results suggest that both consistency when confronted with a given rule and adaptive switching when the rules are changed emerge around 4 years of age. On an individual basis, we observed a similar tendency in 4- and 5-year-olds: consistent responding on pre-test seems to be associated with adaptive flexibility. Children showing a preference for one mode of response during pre-test tended to change more often and in a more principled way when the pictorial contexts were introduced and to return more often to their initial choices when confronted to the same stimuli during post-test.

Five-year-old children’s performance offers yet a different picture. While they maintained their original categorization choice when the conditions remained equivalent, most of them modified their initial choice in response to the scenes. This responsiveness to the task demands reveals two skills: these children can maintain consistently one mode of response (even after a time delay), and they can consider another relation and adjust their response to the presented context.

Although we explicitly told the children that the scenes should help them make a decision, it might not have been equally effective at directing 3-year-olds’ and older children’s attention toward the scenes. The experimental manipulation was designed to correspond to naturalistic situations in which children are required to categorize objects in various contexts; it was not designed to elicit optimal competence.

Nevertheless, this research sheds some light on the development of categorical flexibility in the preschool years. In the age range we studied, younger children’s responses appear to vary in time and as a function of specific items presented as choice options. Despite a common instruction, most 3-year-olds do not maintain a single mode of response throughout a session (either thematic or taxonomic); moreover, the introduction of scenes does not significantly influence their choices. The next step in the development of categorical flexibility (observed in 4-year-olds) seems to involve consistency of a single response mode, while only partially adapting to contextual cues.
Finally, older children make categorical choices in a coherent fashion, adjusting their choices when contextual changes are introduced, otherwise maintaining them.

A remaining question of interest is why the consistency of responses co-occurs with the beginning of adaptive flexibility. Our interpretation is that consistency reflects an increasing explicit representation of the underlying relational links (Karmiloff-Smith, 1992), a level of explicitness necessary for adaptive flexibility to occur. As already suggested, the non-specific instructions used in this study do not make clear that a coherent sorting rule has to be followed (as such, our task differs from Zelazo et al.’s (1996) Dimensional Change Card Sort task and Deák’s (2000) inductive generalization tasks). When children, at 4 years of age, begin to respond consistently, they so use the thematic relations, relations that are commonly used by young children.

Further investigations are also needed to test the generality of this developmental path from spontaneous variability to adaptive flexibility, via a predominance of one type of relation. It would be particularly worthwhile to investigate flexibility in a larger variety of categorization tasks. Moreover, although this path might describe a general developmental progression, each stage might occur at different ages as a function of tasks, types of relation and familiarity with the domain at hand. For instance, using a verbal recall task, Plumert (1994) assessed the organizational strategies used by 10- to 12-year-olds. She found that only 12-year-olds clearly adapt their clustering strategy as a function of the changing demands of the tasks, thus suggesting rather late development of flexibility in this situation.

The developmental evolution of categorical flexibility that we observed, from spontaneous variability to categorical flexibility adjusted to contextual information, is compatible with theoretical models of the development of cognitive flexibility (Karmiloff-Smith, 1992; Zelazo & Frye, 1997). Our data suggest that the study of categorical flexibility might be particularly useful for gaining a better understanding of the development of cognitive flexibility in general. A first step could be to explore to what extent the development of flexibility depends on the development of explicit representations. For instance, the explicitness of the representations of categorical relations could be assessed by requiring children to classify sets of objects that have already been grouped either on a thematic, taxonomic or perceptual basis. One could then test whether performance on this task correlates with measures of adaptive flexibility.

To conclude, the present study demonstrates that, although several types of relational links are available as early as 3 years of age, the ability to switch from one relation to another as a function of contextual cues is not mature at this stage, but undergoes further development throughout the preschool years.

Acknowledgements

This work was supported by grants from the University of Provence (Aix), the University Pierre Mendes France of Grenoble and from the Centre National de la Recherche Scientifique. Both authors thank the preschoolers and their teachers who participated in these studies. We are grateful to Stephanie Bourgoin, Christele Mosca and Hadda Rahmoune for their assistance in most phases of this project. We thank Yvette Hatwell and Sandra Waxman for helpful comments on earlier versions of this article.
References


Received 20 August 1998; revised version received 28 September 2000

Appendix 1

Material used in Experiment 1

<table>
<thead>
<tr>
<th>Target</th>
<th>Taxonomic pairs</th>
<th>Thematic pairs</th>
<th>Taxonomic scene</th>
<th>Thematic scene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ball</td>
<td>Teddy bear &amp; rocking horse</td>
<td>Football player &amp; sport shoes</td>
<td>Toyshop</td>
<td>Adventure playground</td>
</tr>
<tr>
<td></td>
<td>Building blocks &amp; puppet</td>
<td>Girl &amp; dog</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boots</td>
<td>Trousers &amp; blouson-style jacket</td>
<td>Hat &amp; horse saddle</td>
<td>Open cupboard in a bedroom</td>
<td>Country-side with a horse</td>
</tr>
<tr>
<td></td>
<td>Scarf &amp; shirt</td>
<td>Riding cap &amp; spur</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Axe</td>
<td>Shovel &amp; hammer</td>
<td>Log &amp; stump</td>
<td>Tool shed</td>
<td>Forest</td>
</tr>
<tr>
<td></td>
<td>Screwdriver &amp; pliers</td>
<td>Leafless tree &amp; bundle of wood</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lion</td>
<td>Bird &amp; giraffe</td>
<td>Tamer &amp; whip</td>
<td>Zoo</td>
<td>Circus</td>
</tr>
<tr>
<td></td>
<td>Dog &amp; ape</td>
<td>Tamer (woman) &amp; cage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orange</td>
<td>Carrots &amp; grapes</td>
<td>Bowl &amp; spoon</td>
<td>Open refrigerator in a kitchen</td>
<td>Table set for breakfast</td>
</tr>
<tr>
<td></td>
<td>Water &amp; meat</td>
<td>Teapot &amp; bread</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apple</td>
<td>Carrots &amp; pineapple</td>
<td>Ladder &amp; basket</td>
<td>Grocery shop</td>
<td>Countryside with a tree</td>
</tr>
<tr>
<td></td>
<td>Banana &amp; cheese</td>
<td>Knife &amp; basket</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rake</td>
<td>Wrench &amp; saw</td>
<td>Sun hat &amp; beach umbrella</td>
<td>Garden</td>
<td>Beach</td>
</tr>
<tr>
<td></td>
<td>Hammer &amp; scissors</td>
<td>Sand castle &amp; swim suit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gloves</td>
<td>Pants &amp; hat</td>
<td>Snowman &amp; skis</td>
<td>Open wardrobe in a bedroom</td>
<td>Ski station</td>
</tr>
<tr>
<td></td>
<td>Shirt &amp; sweater</td>
<td>Scarf &amp; ski sticks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Towel</td>
<td>Jumper &amp; socks</td>
<td>Tooth brush &amp; comb</td>
<td>Clothes line</td>
<td>Bathroom</td>
</tr>
<tr>
<td></td>
<td>Dress &amp; shirt</td>
<td>Bath robe &amp; hairbrush</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Appendix 2

#### Material used in Experiment 2

<table>
<thead>
<tr>
<th>Target</th>
<th>Taxonomic choice</th>
<th>Thematic choice</th>
<th>Perceptual choice</th>
<th>Unrelated choice</th>
<th>Taxonomic scene</th>
<th>Thematic scene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apple</td>
<td>Banana</td>
<td>Knife</td>
<td>Ball</td>
<td>Boots</td>
<td>Vegetable counter</td>
<td>Kitchen</td>
</tr>
<tr>
<td>Egg</td>
<td>Sausage</td>
<td>Hen</td>
<td>Necklace</td>
<td>Key</td>
<td>Grocery store</td>
<td>Farm</td>
</tr>
<tr>
<td>Carrot</td>
<td>Cauliflower</td>
<td>Rabbit</td>
<td>Rocket</td>
<td>Scissors</td>
<td>Fruit counter</td>
<td>Hutch</td>
</tr>
<tr>
<td>Fish</td>
<td>Dog</td>
<td>Fishing rod</td>
<td>Rugby ball</td>
<td>Child’s drawing</td>
<td>Animal book</td>
<td>Riverside</td>
</tr>
<tr>
<td>Football</td>
<td>Teddy bear</td>
<td>Football player</td>
<td>Orange</td>
<td>Toothbrush</td>
<td>Toyshop</td>
<td>Football field</td>
</tr>
<tr>
<td>Daisy</td>
<td>Mushroom</td>
<td>Bee</td>
<td>Star</td>
<td>Doll</td>
<td>Flower book</td>
<td>Beehives</td>
</tr>
<tr>
<td>Mouse</td>
<td>Bird</td>
<td>Cheese</td>
<td>chill</td>
<td>Chair</td>
<td>Animal book</td>
<td>Mouse trap</td>
</tr>
<tr>
<td>Cherry</td>
<td>Pineapple</td>
<td>Ladder</td>
<td>Spinning top</td>
<td>Giraffe</td>
<td>Vegetable stand</td>
<td>Orchard</td>
</tr>
<tr>
<td>Bird</td>
<td>Zebra</td>
<td>Nest</td>
<td>Plane</td>
<td>Bottle</td>
<td>Zoo</td>
<td>Tree</td>
</tr>
<tr>
<td>Writing desk</td>
<td>Bed</td>
<td>Pencil</td>
<td>Flag</td>
<td>Violin</td>
<td>Furniture shop</td>
<td>Classroom</td>
</tr>
</tbody>
</table>