

Real-world schemata and scene recognition in adults and children

HOWARD S. HOCK, LORANN ROMANSKI, ANTHONY GALIE,
and CATHY S. WILLIAMS

Florida Atlantic University, Boca Raton, Florida 33431

Recognition memory for previously seen multiobject scenes was examined for different types of contextual arrangements between objects in the scenes. It was found that organized scenes with novel but possible interobject relations were recognized more accurately than either organized scenes with familiar interobject relations or unorganized scenes with impossible interobject relations. This finding was obtained for adults, 8- to 10-year-old children, and 5- to 8-year-old children who indicated concrete-operational ability in Piaget's conservation-of-liquid quantity task. The results were interpreted in conjunction with a two-stage model of scene processing involving the formation of a schema to represent a scene (Stage 1), and the operation of the schema in governing the further processing of detailed information in the scene (Stage 2). It was concluded that preoperational children can form schemata to represent organized scenes (Stage 1), but it is not until the emergence of concrete operations that these schemata become operational with respect to guiding the further processing of information in the scene (Stage 2).

The observation that native speakers of a language can discriminate between grammatical and ungrammatical sequences of words in their language has had an important influence on psycholinguistic research (Chomsky, 1957). Psycholinguists have extensively studied the role of grammatical organization in sentence processing and retention (see Fodor, Bever, & Garrett, 1974). A similar observation can be made regarding native perceivers of the real-world environment. That is, native perceivers can discriminate between organized possible arrangements of real-world objects and unorganized impossible arrangements of real-world objects. It is only recently, however, that the role of organization in the processing and retention of multiobject scenes has come under systematic investigation.

The perceptual processing of organized and unorganized scenes has been investigated by Biederman and his colleagues (Biederman, 1972, 1977; Biederman, Glass, & Stacy, 1973; Biederman, Rabinowitz, Glass, & Stacy, 1974). In one set of experiments, they studied the detection of individual objects in photographs of multiobject real-world scenes. When the photographs

were presented in their original coherent form, they found that target objects in the scenes were detected more accurately and more rapidly than when the photographs were cut into sections and presented in jumbled form. More recently, Biederman (1977) has systematically introduced violations of specific relations (e.g., support, position, size, etc.), into hand-drawn scenes. He found that as the number of violations in a scene was increased, subjects' accuracy in detecting target objects in the scene decreased, as did the time required for subjects to judge that there was "something wrong" with the scene.

The retention of organized and unorganized scenes has been investigated by Mandler and her colleagues (Mandler & Johnson, 1976; Mandler & Parker, 1976; Mandler & Ritchie, 1977; Mandler & Stein, 1974). The general procedure in these studies was to present a sequence of hand-drawn scenes for subjects to remember, and then to follow with a forced-choice recognition task in which each "original" scene was paired with a new distractor scene. The distractor was varied from the original scene in a number of different ways (e.g., changing the location of an object, deleting an object, replacing an object with another object). The purpose of these manipulations was to determine the type of information that was retained for the organized and unorganized scenes that were presented in the original sequence. In general, these studies indicated that differences in recognition accuracy between the organized and unorganized scenes varied with the nature of the distractor in the recognition task. For distractors that assessed the retention of spatial relations between the objects in a scene, organized scenes were recognized better than unorganized scenes.

This paper is an expanded version of a paper read at the meeting of the Psychonomic Society in November 1975. The research reported in this study was supported by a Seed Grant from Florida Atlantic University. The authors wish to thank Barbara Bittner, Director, and the faculty of A. D. Henderson School for their cooperation in conducting this study. Thanks are also due to Rebecca Whalin and Rebecca Hoffman for their help in gathering data and to David Bjorklund for his careful reading of the manuscript. Requests for reprints should be sent to Howard S. Hock, Department of Psychology, Florida Atlantic University, Boca Raton, Florida 33431.

For distractors that assessed the retention of either the identity or the appearance of the objects in a scene, there was no difference in retention between the organized and unorganized scenes. Finally, for distractors that assessed the retention of the spatial composition of the scene (i.e., the positions in space occupied by objects in the scene), the unorganized scenes were recognized better than the organized scenes. Mandler and her colleagues explain their results in terms of underlying schemata (Bartlett, 1932) that guide the selection and organization of information in a scene. They view schemata as abstract representations comprising relational information that is abstracted from a lifetime of experience with real-world scenes.

In conjunction with tasks such as Mandler's that examine the retention of particular scenes, the role of schemata can be described more precisely. That is, schemata derived from real-world experience can be viewed as structures that govern the formation of episodic representations for scenes presented in a particular temporal context (Tulving, 1972). The present study further examined the nature of real-world schemata as they influence the formation of episodic memory representations for scenes. The initial goal of the study was to determine whether real-world schemata are generalizable to all scenes or whether they are limited in their function to familiar arrangements of objects. The second goal was to determine whether there are developmental changes in the effect of real-world schemata on the formation of episodic representations for scenes.

In the Mandler studies described above, the organized scenes comprised objects in a spatially coherent relationship. In addition, however, the relationship between the objects in these scenes was familiar. It must be emphasized at this point that the term "familiar" denotes that the relationships between the objects depicted in a scene were similar to previously experienced relationships of objects of that type. Our usage of the term "familiar" does not imply that the particular scene in question is a copy of a scene experienced in the past. A scene is familiar in the same sense that an individual object is familiar: It can be assimilated into a previously established category. Such schemata could be abstracted from experience with various kinds of real-world "situations" or events. They might specify sets of objects that tend to co-occur in the real world (e.g., cars, sidewalks, traffic signals) and the relationships in which they are typically experienced (e.g., to compose a street scene). On the basis of this conceptualization, real-world schemata would be expected to influence the formation of episodic representations only for categorizable scenes with familiar interobject relations.

An alternative possibility is that real-world schemata are generalizable to all physically possible scenes, whether the interobject relations in the scene are

familiar or novel. Hock, Gordon, and Whitehurst (1974) conceived of such generalizable schemata as rule systems that specify the possible spatial locations of various objects (e.g., elephant, baseball bat), as well as the possible relations into which objects could enter (e.g., an elephant could hold a baseball bat, but not vice versa).¹ On the basis of this conceptualization, real-world schemata would be expected to influence the formation of episodic representations for all scenes in which the interobject relations are consistent with the rules specifying the physically possible relations between real-world objects.

The experiments reported in this paper were designed to separate the effects of scene organization and scene familiarity on recognition memory. Subjects were presented: (1) spatially organized scenes with familiar interobject relations, (2) spatially organized scenes with novel, but possible, interobject relations, and (3) spatially unorganized scenes with impossible interobject relations (see Figure 1 for examples). The significance of the distinction between scene organization and scene familiarity lies in the information it provides about the generalizability of real-world schemata. If real-world schemata involve categorical representations derived from experience with real-world "situations" or events, then recognition accuracy for scenes with familiar interobject relations should be significantly different from recognition accuracy for either novel possible scenes or unorganized impossible

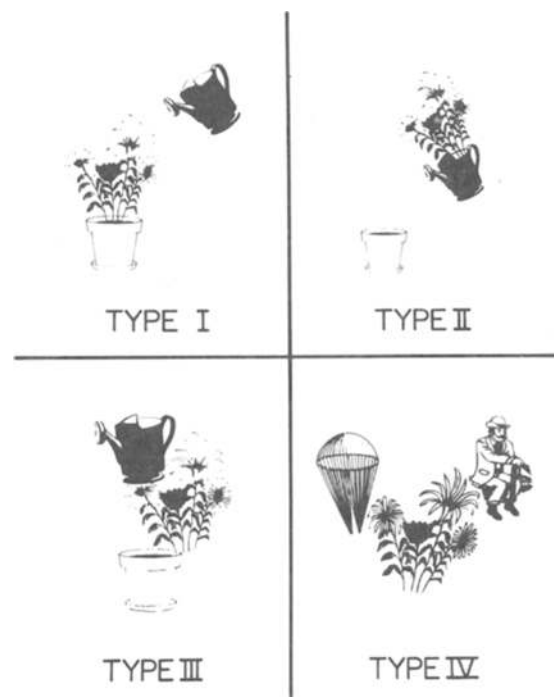


Figure 1. Four scenes involving the picture of the cluster of flowers. Each scene illustrates one of the four "types" of contextual arrangement.

scenes. If, however, real-world schemata involve generalizable rule systems that specify the physically possible relations between objects, then recognition accuracy for all rule-following scenes, whether the interobject relations are familiar or novel, should be significantly different than recognition accuracy for unorganized impossible scenes.

Support for the hypothesis that real-world schemata involve generalizable rules of physical possibility would be consistent with a two-stage model of scene processing suggested by Biederman et al. (1973). The first stage would involve the formation of a schematic representation for the scene. This might involve identifying the objects in the scene at their "basic" level, the most abstract level of categorization at which physically possible relations between real-world objects could be specified (Rosch, Mervis, Gray, Johnson, & Boyes-Bream, 1976). The "basic" category names might then enter into a relational (schematic) description that would constitute the initial interpretation for the scene. The second stage of the model would involve the operation of the schema formed in Stage 1 in governing the selection and organization of detailed information in the scene. Thus, Stage 2 processing might involve attention to the detailed features that identify the critical objects in the scene at a level of categorization more specific than (i.e., subordinate to) the "basic" level of Stage 1. It might also involve a more precise specification of the spatial locations of the interacting objects in the scene.

The two-stage model described above leads to the second goal of this study: the determination of whether there are developmental changes in the effect of real-world schemata on the formation of episodic representations for scenes. The first stage in the model, the generation of a relational schematic representation for a scene, seems within the competence of young children. For example, 2- and 3-year-old children can construct arrangements of objects involving the relations of "containing" and "supporting" (R. Brown, 1973), as well as interactive relations (e.g., kiss, chase) between objects (Bever, 1970). The second stage of the model involves the operation of the schema in governing the selection and organization of detailed information in the scene. In terms of development, Piaget has maintained that schemata are not operational until the emergence of concrete operations (Piaget & Inhelder, 1969). From this point of view, a preoperational child could form a schema to represent a scene (Stage 1), but the schema would not be operational with respect to guiding the selection and organization of detailed information in the scene (Stage 2). It was hypothesized on this basis that schema formation would have a relatively small effect on recognition performance for preoperational compared with concrete-operational children. That is, the difference in recognition accuracy between physically possible and physically impossible

scenes was expected to be smaller for preoperational than concrete-operational children.

EXPERIMENT 1

Method

Stimuli. The stimuli used in this experiment were identical with those used by Hock et al. (1974). They were hand-drawn scenes, each scene comprising three different familiar objects. There were four types of contextual relationships among the objects in a scene. Type 1 comprised scenes with familiar relations between the objects and Type 2 comprised the same objects arranged in a novel, but physically possible manner. The Type 3 drawings depicted the same objects, but in an unorganized, physically impossible arrangement. Ten sets of objects were used. Each set comprised a group of objects that tended to co-occur in the real-world environment. Each set was used to generate three scenes (Types 1, 2, and 3) by varying the arrangement of the objects. An additional 10 sets of objects composed the Type 4 scenes. The objects in each Type 4 scene did not tend to co-occur in the real-world environment. They were selected from the 10 sets of objects composing the Type 1, 2, and 3 scenes, and then arranged in an unorganized, physically impossible manner. Because there was no contextual information inherent in the Type 4 scenes, they provided a baseline level of recognition performance. In summary, there were a total of 40 scenes, 10 per contextual "type." The "types" of contextual relation for the four scenes that include a cluster of flowers are illustrated in Figure 1.²

The initial phase of the experiment involved the sequential presentation of 40 different scenes. This was immediately followed by the recognition phase, in which pairs of scenes were presented in sequence. One scene came from the previously presented sequence (the original). The other scene (the distractor) was of the same context type, but differed from the original scene by one object (see Figure 2 for example). In this way, there were two versions of each scene, the two versions differing by the same critical object in all four context types. For example, each of the scenes in Figure 1 had a counterpart in which the flowers were replaced by a rubber plant. The critical objects interacted with the other objects in the same manner for both versions of each scene (e.g., the flowers in the watering can; the rubber plant in the watering can).

Design. The experimental design is illustrated, by example, in Table 1. The initial consideration was that the critical objects be assigned equally often to the "original" scenes and their distractors. For example, if the Type 1 and Type 3 scenes with the flowers were in the original list, then the Type 2 and Type 4 scenes in the original list would have the rubber plant instead of the flowers. In the subsequent recognition list, the distractor for the Type 1 flowers (from the original list) would be the Type 1 rubber plant, the distractor for the Type 2 rubber



Figure 2. Example of pair of scenes presented in the forced-choice recognition test.

Table 1
Illustration of Experimental Design

Type	Sequence		
	Original	Recognition	
Group 1			
1	flowers	<i>flowers</i>	rubber plant
2	rubber plant	flowers	<i>rubber plant</i>
3	flowers	<i>rubber plant</i>	<i>flowers</i>
4	rubber plant	<i>rubber plant</i>	flowers
1	table	desk	<i>table</i>
2	table	<i>table</i>	desk
3	desk	table	<i>desk</i>
4	desk	<i>desk</i>	table
Group 2			
1	rubber plant	flowers	<i>rubber plant</i>
2	flowers	<i>flowers</i>	rubber plant
3	rubber plant	<i>rubber plant</i>	flowers
4	flowers	rubber plant	<i>flowers</i>
1	desk	<i>desk</i>	table
2	desk	table	<i>desk</i>
3	table	<i>table</i>	desk
4	table	desk	<i>table</i>

plant (from the original list) would be the Type 2 flowers, and so on. A different pattern of critical-object assignment to the various context types was applied to the other scenes. For example, instead of the 1-3, 2-4 pattern assigned to the flowers/rubber plant stimuli, a 1-2, 3-4 pattern was assigned to scenes in which table and desk were the critical objects. That is, the Type 1 and Type 2 scenes with the table and the Type 3 and Type 4 scenes with the desk were in the original list. In the subsequent recognition list, the distractor for the Type 1 table (from the original list) was the Type 1 desk, and so on. Finally, the objects that were in the original scene and the objects in the distractor scene were reversed for different groups of subjects.

In the initial phase of the experiment, 40 scenes were presented in one of three random orders. In the subsequent recognition phase, pairs of scenes were presented in random order. In the latter sequence, which was the same for each subject, the "original" scene (the recipient of the correct recognition response) appeared equally often in the left and right positions of the stimulus pairs.

Procedure. The stimuli were back-projected onto a translucent screen by random-access Carousel projector. Each scene intercepted a visual angle of about 4 deg, and the overall visual angle when pairs of scenes were presented in the recognition test was about 10 deg. During the initial presentation of the 40 scenes, each scene was shown for 3 sec, the subjects being instructed to "try to remember them." In the subsequent forced-choice recognition task, which followed immediately after the sequence of "originals," each pair of scenes was presented until the subject denoted (by saying "left" or "right") the scene that they recognized from the original list.

Subjects. Sixty-six unpaid undergraduate students at Florida Atlantic University voluntarily participated in this experiment, which lasted about 15 min.

Results and Discussion

Percentage recognition errors for each of the four types of contextual relation are presented in Table 2. The effect of context type on recognition accuracy was significant [$F(3,192) = 10.63$, $p < .005$, $MSe = 182$].³ Subsequent Newman-Keuls comparisons indicated that

performance for the Type 2 scenes was significantly more accurate ($p < .01$) than performance on each of the other three types. None of the other comparisons were significant ($p > .05$; the Type 1/Type 3 difference fell just short of significance at the .05 level). Further examination of the data indicated that the differences between Type 2 and the other three types of arrangement did not result from the effect on recognition accuracy obtained for only a few of the 10 sets of scenes having common objects (the four scenes presented in Figure 1 compose 1 of the 10 sets). Recognition accuracy was: (1) greater for Type 2 than Type 1 arrangements for 7 of the 10 sets of scenes, (2) greater for Type 2 than Type 3 arrangements for 8 of the 10 sets of scenes (one tie), and (3) greater for Type 2 than Type 4 arrangements for 8 of the 10 sets of scenes.

The evidence that the novel possible scenes (Type 2) were recognized more accurately than the impossible unorganized scenes (Type 3) supported the hypothesis that the formation of episodic representations for scenes is governed by generalizable rules that specify the physically possible relations between real-world objects. For scenes that conform to the rules of physical possibility, as was the case for the Type 2 scenes, episodic schemata could be generated to represent the scenes. For scenes that do not conform to the rules of physical possibility, as was the case for the Type 3 scenes, episodic schemata could not be generated.

Of further interest was the significant difference in recognition accuracy between the scenes with familiar interobject relations (Type 1) and the scenes with novel, but possible, interobject relations (Type 2). This finding provided evidence that real-world schematic knowledge is not restricted to a generalizable rule system specifying the relational possibilities of objects. The difference in recognition accuracy between the familiar (Type 1) and novel (Type 2) scenes suggested the existence of a second level of real-world schemata that is abstracted from experience with real-world "situations" or events. As indicated in the introduction to this paper, such schemata might specify sets of objects that tend to co-occur in the real-world as well as the relationships in which they are typically experienced. Schemata of this kind could function as categorical structures for the recognition of scenes with familiar interobject relations. Thus, relational

Table 2
Mean Percentage of Recognition Errors for Each Type of Contextual Arrangement (Experiment 1: Adult Subjects)

	Context Type	Percent Errors
1	Familiar	26.8
2	Novel, Possible	20.0
3	Unorganized, Impossible	31.7
4	No Contextual Information	31.2
	Mean	27.4

schemata could be generated to represent either familiar (Type 1) or novel (Type 2) scenes, but the schemata formed for the familiar scenes could be assimilated into categorical representations derived from experience with interobject relations similar to those depicted in the scenes.

The consequences of categorization for recognition performance can be understood in conjunction with Bruner's (1957) theory of perceptual readiness. According to this theory, the categorization of a stimulus is based on inferential processes that focus attention on features that specify the categorical identity of a stimulus. The emphasis in Bruner's theory is clearly on efficiency in perceptual processing. However, the efficiency achieved by selectively attending to category-defining features can have significant costs in subsequent tasks involving the retention of the categorized stimuli. In terms of subsequent recognition tasks, for example, attention to features that do not define category membership can be just as important for performance as attention to features that do specify category membership. Categorization could decrease recognition accuracy for the familiar Type 1 scenes relative to the novel Type 2 scenes because selective attention to category-defining features would reduce attention to other features that could contribute to the discrimination of the "original" Type 1 scenes from similar distractors (in the forced-choice recognition task).

On the basis of the two-stage model proposed in the introduction to this paper, the difference in recognition accuracy between the Type 2 and Type 3 scenes could be attributed to the effect of schemata generated in Stage 1 on the selection and organization of detailed information in Stage 2. Recognition performance for the Type 2 scenes would be relatively good because schemata could be generated for the Type 2 scenes and these schemata could direct the further processing of detailed information in these scenes. Recognition performance for the Type 3 scenes would be relatively poor because schemata could not be generated for the Type 3 scenes and thus would not be available to direct further processing of detailed information in these scenes. Further, the Type 1/Type 2 difference in recognition accuracy could be attributed to the schemata formed to represent the familiar Type 1 scenes (Stage 1) being assimilated into categorical structures that narrow further processing of the scenes to features that would confirm that the scenes were correctly categorized (Stage 2). This focusing of attention on category-defining features could result in recognition accuracy being poorer for the familiar (and thus categorizable) Type 1 scenes than the novel (and thus uncategorizable) Type 2 scenes.⁴

EXPERIMENT 2

The data reported in this experiment provide a replication of the results of Experiment 1 with a

group of children ranging in chronological age from 8 to 10 years. On the basis of their age, it was inferred that concrete-operational ability would be well established for these children (Piaget & Inhelder, 1969). It was expected, therefore, that real-world schemata would influence their formation of episodic representations for the scenes, and the results of the recognition task would resemble those of the adult subjects in Experiment 1.

Method

The stimuli and design were identical to Experiment 1. There was only one change in the procedure. In the forced-choice recognition task, the subjects responded by pointing at the scene they recognized from the original list rather than by saying "left" or "right." The subjects were 44 students at A. D. Henderson University School in Boca Raton, Florida. There were 26 8-year-old subjects ranging in age from 8.1 to 9.0 years and 18 10-year-old subjects ranging in age from 10.1 to 10.9 years. The mean age for the entire group was 9.4 years.

Results

The percentage of recognition errors for each of the context types is presented in Table 3. While overall performance was somewhat lower for the children in this experiment compared with the adults in Experiment 1, the pattern of errors was quite similar in the two experiments. The effect of context type on accuracy was significant [$F(3,126) = 9.21, p < .005, MSe = 190$], and the results of subsequent Newman-Keuls comparisons indicated that recognition was significantly more accurate ($p < .01$) for the Type 2 scenes than for each of the other three types. None of the other comparisons were significant ($p > .05$). Further examination of the data again indicated that the differences between Type 2 and the other three types of arrangement were not restricted to effects on recognition accuracy obtained for only a few of the 10 sets of scenes having common objects. Recognition accuracy was: (1) greater for Type 2 than Type 1 arrangements for 8 of the 10 sets of scenes (one tie), (2) greater for Type 2 than Type 3 arrangements for 8 of the 10 sets of scenes, and (3) greater for Type 2 than Type 4 arrangements for 8 of the 10 sets of scenes.

EXPERIMENT 3

The purpose of this experiment was to investigate developmental differences in the effect of real-world

Table 3
Mean Percentage of Recognition Errors for Each Type
of Contextual Arrangement (Experiment 2:
8- to 10-Year-Old Children)

	Context Type	Percent Errors
1	Familiar	35.0
2	Novel, Possible	24.3
3	Unorganized, Impossible	33.2
4	No Contextual Information	39.3
	Mean	33.0

schemata on the formation of episodic representations for scenes. Following Piaget (Piaget & Inhelder, 1969), it was hypothesized that only when children develop concrete-operational ability would the schemata formed for a scene be expected to influence the selection and organization of detailed information in the scene. On this basis, it was predicted that the difference in recognition accuracy between physically possible and physically impossible scenes would be smaller for preoperational than for concrete-operational children.

Method

The stimuli and design for the scene-recognition task were identical to Experiment 2. Piaget's conservation-of-liquid quantity task was administered to each child prior to the scene-recognition task. The emergence of concrete-operational ability was inferred from successful performance on this task.

Procedure. The two identical beakers used at the start of the conservation task were 9 cm in diameter and 11 cm in height. The third (transfer) beaker was 3 cm in diameter 26 cm in height. After being satisfied that the amount of water in the two identical beakers was the same, the child watched as the experimenter poured the water from one of the beakers into the tall thin beaker. The child was then asked whether the amount of water in the two beakers was the same, or whether one beaker now had more water. The water was then returned to the original beaker and the procedure was repeated. This time the question was put to the child in reverse order. That is, the child was asked whether one beaker now had more water or did they have the same amount. Following his response, the child was asked to explain why he or she had made that response.

Subjects were placed in two groups on the basis of their responses on the conservation task (their explanations were considered afterward). The design of the scene-recognition task was complete within each group. That is, the assignment of scenes as "originals" or "distractors" was counterbalanced between two groups of conservers and two groups of nonconservers. The recognition-memory task was administered as in Experiment 2.

Subjects. The subjects were 134 students at A. D. Henderson University School in Boca Raton, Florida. The children ranged in age from 5.4 to 8.2 years.

Results

Seventy-four children gave incorrect responses on the conservation task. They invariably said that there was more water in the tall thin beaker. These nonconservers ranged in age from 5.4 to 7.5 years (mean = 6.3). Sixty children, ranging in age from 5.8 to 8.2 years (mean = 7.1), correctly indicated that the amount of water was the same in the two beakers. Of these children, 49 gave appropriate explanations for their responses based on either the identity of the water in the two different beakers or the reciprocal relation between height and width (Piaget & Inhelder, 1969). The remaining 11 children gave either no explanation or an explanation that was uninterpretable. An examination of their data for the scene-recognition task provided no justification for considering these 11 children separately from the remaining 49 conservers. Subsequent analyses were therefore based on the total set of 60 conservers.

The percentage of recognition errors for each of the four context types in the scene-recognition task are presented in Table 4. The difference in performance between the conservers and nonconservers was significant [$F(1,130) = 11.06$, $MSe = 263$], as was the effect of context type [$F(3,390) = 4.11$, $p < .01$, $MSe = 227$].⁵ The interaction between groups (conservers and nonconservers) and context type was significant [$F(3,390) = 2.92$, $p < .05$, $MSe = 227$]. Tests of simple effects indicated that the effect of context type was significant for the group of conservers [$F(3,390) = 6.01$, $p < .005$] but was not significant for the group of nonconservers [$F(3,390) = 1.01$, $p > .05$]. Newman-Keuls comparisons for the conservers followed the same pattern as in the previous two experiments. The Type 2 scenes were recognized significantly more accurately than any of the other context types ($p < .05$), and none of the other comparisons were significant ($p > .05$). Further examination of the data for the conservers indicated that recognition accuracy was: (1) greater for Type 2 than Type 1 arrangements for 7 of the 10 sets of scenes (one tie), (2) greater for Type 2 than Type 3 arrangements for 8 of the 10 sets of scenes, and (3) greater for Type 2 than Type 4 arrangements for 8 of the 10 sets of scenes.

An additional analysis examined the extent to which differences in chronological age were responsible for the differences in recognition performance between the conservers and nonconservers. It will be recalled that the conservers averaged 7.1 years of age, while the nonconservers averaged 6.3 years of age. For this analysis, we formed pairs of conservers and nonconservers that were matched within 1 month in age. Because of the age distributions, only 36 matched pairs could be formed. The mean age of these children was 6.7 years. The scores on the recognition task for the match groups of conservers and nonconservers are presented in Table 5.

As can be seen in Table 5, the pattern of data for the recognition task is essentially unchanged compared with Table 4, in which the conservers and nonconservers were not matched in age. Although the smaller number of subjects in the matched-groups analysis reduced the

Table 4
Mean Percentage of Recognition Errors for Each Type of Contextual Arrangement (Experiment 3: Children Placed in One of Two Groups on the Basis of Their Performance in Conservation Task)

	Context Type	Percent Errors	
		C	NC
1	Familiar	40.0	43.5
2	Novel, Possible	30.0	40.3
3	Unorganized, Impossible	40.2	39.7
4	No Contextual Information	36.8	39.7
	Mean	36.8	41.5

Note—C = conservers; NC = nonconservers.

Table 5
Mean Percentage of Recognition Errors for Each Type of
Contextual Arrangement (Experiment 3: Conservers and
Nonconservers Matched in Chronological Age)

	Context Type	Percent Errors	
		C	NC
1	Familiar	43.1	38.6
2	Novel, Possible	31.7	40.6
3	Unorganized, Impossible	40.9	40.3
4	No Contextual Information	36.9	40.0
	Mean	38.2	39.9

Note—C = conservers; NC = nonconservers.

statistical power of the analysis of variance for these data, the results suggest that the differences in performance between the conservers and nonconservers in the scene-recognition task were not due to the small difference in age between the two groups. The interaction between groups (conservers vs. nonconservers) and context type, which was significant when the data for all the subjects were analyzed, fell just short of significance at the $p = .05$ level [$F(3,102) = 2.56$, $MSe = 228$] for the matched-groups analysis.⁶ However, tests of simple effects for the matched groups again indicated that the effect of context type was significant for the conservers [$F(3,102) = 3.96$, $p < .05$], but not for the nonconservers [$F(3,102) < 1.0$]. Finally, Newman-Keuls comparisons for the group of conservers indicated that the Type 2 scenes were again recognized better than the Type 1 and Type 3 scenes ($p < .05$). The difference between Type 2 and Type 4, as well as all other comparisons, was not significant ($p > .05$).

Discussion

The results of the experiment supported the predicted relation between performance in the conservation-of-liquid quantity task and performance in the scene-recognition task. The emergence of concrete-operational ability was inferred for children who responded correctly in the conservation task. These children exhibited the same pattern of recognition errors over the four context types that was obtained for adults and 8- to 10-year-old children in the previous two experiments. Children who responded incorrectly in the conservation-of-liquid task were inferred to be preoperational. For these children, the effect of context type on recognition accuracy was not significant.

The difference in recognition performance between the conservers and nonconservers supported the hypothesis that real-world schemata would have relatively little effect on scene recognition prior to the development of concrete operationality. It is at this stage of development, according to Piaget, that schemata become operational. With regard to our two-stage model, operationality implies that the schemata formed to represent a scene (Stage 1) can govern the further processing of information in the scene (Stage 2).

Thus, the difference in recognition performance between the conservers and nonconservers can be attributed to the operation of the schema in governing Stage 2 processing for the conservers, but not for the nonconservers.

The results for the nonconservers were, at least on the surface, inconsistent with the results of previous research. Typical of the latter is a study by Horowitz, Lampel, and Takanishi (1969), who examined 3- and 4-year-old children's ability to recall objects deleted from previously seen pictures. They found that objects from spatially organized pictures were recalled more accurately than objects from unorganized pictures. A similar result has been obtained in studies of memory for narrative sequences. A. Brown (1975) and A. Brown and Murphy (1975) presented children ranging in age from 3 to 7 years with sequences of interactive pictures. In one condition the sequences of pictures were logically ordered (i.e., they were sequentially organized). In the other, the sequences of pictures were arbitrary. They found that the children could verbally recall as well as reconstruct the order of pictures more accurately for the logical sequences than the arbitrary sequences. However, when a similar group of children were given a recognition test in which they were required to discriminate the original sequences from distractor sequences, the difference in performance between the logical and arbitrary sequences disappeared (A. Brown, 1975). This last finding converges with the results obtained for the nonconservers in Experiment 3. That is, when retention for preoperational children is tested by recognition instead of recall, real-world schemata do not seem to affect performance for logical sequences relative to illogical sequences (A. Brown, 1975), nor do they affect performance for possible scenes relative to impossible scenes (the present Experiment 3).

Why, for preoperational children, would real-world schemata influence the recall and reconstruction of scenes and sequences of pictures, but not their recognition? One approach to answering this question is the two-stage model of scene processing. Based on studies by Bever (1970) and R. Brown (1973), it was assumed that preoperational children were capable of forming schemata to represent scenes (Stage 1). The results of studies involving the recall of objects from scenes (Horowitz et al., 1969) and studies involving the recall and reconstruction of sequences of pictures (A. Brown, 1975; A. Brown & Murphy, 1975) are consistent with this assumption. A relational schema formed for a scene could facilitate recall by providing an organizational framework for the storage and retrieval of objects from the scene.

Recognition tasks, however, can be more demanding than recall tasks when the "originals" are sufficiently similar to the "distractors." For example, in a recall task it could suffice to remember that an object in a scene was a plant. For correct recognition, however,

it might be necessary to discriminate between one type of plant (original) and another (distractor). As a result, the recognition of a scene would be facilitated if a schematic representation for the scene (Stage 1) directed attention to the detailed features (Stage 2) that would result in the critical objects in the scene being identified at a more specific level than was the case for Stage 1 processing. Thus, the schema formed for an organized scene (Stage 1) by a preoperational child could be sufficient to facilitate the recall of objects from the scene, but it would not be until the onset of concrete operations that the schema could govern the further analysis of detailed information in the scene (Stage 2) and thereby improve recognition accuracy for organized scenes relative to recognition accuracy for unorganized scenes.

GENERAL DISCUSSION

The design of the experiments reported in this paper differed somewhat from the design of several similar experiments by Mandler and her colleagues. The main advantage of Mandler's design was that the type of distractor in the recognition test was varied systematically. In the present research, only one type of distractor was used (the identity of one object was changed). As a result, our conclusions must be confined to this one type of distractor manipulation. However, the present design had an advantage relative to Mandler's in that the effects of scene organization and scene familiarity were separated. In Mandler's research the spatial relations between the objects in the organized scenes were familiar. As a result, it is difficult to determine which of Mandler's results were due to scene organization and which were due to scene familiarity. Such a determination is particularly important in light of the results of the present research, which indicated that the effects of scene organization on recognition accuracy are significantly reduced when the scenes involve familiar interobject relations.

The experimental findings reported in the present paper were interpreted in terms of a two-stage model of scene processing. According to this model, which was proposed by Biederman et al. (1973), the first stage involves the formation of a relational schematic representation for a scene. The second stage in the model involves the operation of the schema in governing the further processing of detailed information in the scene. It was noted in the discussion of Experiment 3 that schema formation (Stage 1) might be sufficient to explain the superior recall of objects from organized compared with unorganized scenes.⁷ However, differences in recognition accuracy between organized (Type 2) and unorganized (Type 3) scenes were attributed to the operation of schemata in governing the further processing of detailed information in the scene (Stage 2).

Evidence that schema formation is itself insufficient to explain the difference in accuracy between the possible (Type 2) and impossible (Type 3) scenes came from a comparison between the familiar possible (Type 1) and novel possible (Type 2) scenes. If differences in recognition accuracy depended only on whether or not a schematic representation was formed for a scene, then recognition accuracy for the Type 1 scenes should be as good as, if not better than, recognition accuracy for the Type 2 scenes. It is difficult to conceive of why schema formation should be more difficult for the familiar Type 1 scenes than the novel, possible Type 2 scenes. Yet, recognition accuracy was consistently poorer for the Type 1 scenes than the Type 2 scenes. It seemed most sensible to attribute this difference to the effect of the schema on further processing of scene information (Stage 2) rather than attributing it to differences in schema formation per se (Stage 1).

The significant difference in recognition accuracy between the Type 1 and Type 2 scenes was also important because of the evidence it provided against the possibility that the results of this study depended on stimulus generalization. That is, it might be argued that the novel possible scenes (Type 2) were recognized more accurately than the unorganized impossible scenes (Type 3) because the Type 2 scenes were more similar than the Type 3 scenes to scenes previously experienced in the real world. For this stimulus generalization argument to be acceptable, however, it would be necessary for the familiar Type 1 scenes to be recognized as accurately as, if not more accurately than, the novel Type 2 scenes, since the Type 1 scenes would be still more similar to previously experienced scenes. The evidence that Type 2 scenes were recognized more accurately than Type 1 scenes therefore ruled out stimulus generalization as an explanation for the obtained results.

In the course of this study, there was occasion to refer to three levels of real-world schemata in scene recognition. At the "deepest," most abstract level were schemata comprising generalizable rules that specify the relational possibilities of objects. It was asserted that these abstract schemata were the basis for the formation of episodic schemata to represent particular scenes, and thus were responsible for the significant difference in recognition accuracy between the physically possible (Type 2) and physically impossible (Type 3) scenes.

Also proposed were categorical schemata. These seem to occupy a level of abstractness intermediate to deep-level rule systems and surface-level episodic schemata. It was suggested that categorical schemata specify sets of objects that tend to co-occur as well as the relationships in which they are typically experienced. Thus, the formation of an episodic schematic representation for a scene with familiar interobject relations could result in the assimilation of that scene into a

categorical schema. On the basis of Bruner's (1957) theory of perceptual readiness, it was argued that such categorical structures would focus further attention on features that would confirm that the scene was correctly categorized. The finding that familiar Type 1 scenes were recognized significantly less accurately than novel, and therefore uncategorizable, Type 2 scenes, was attributed to the narrowing of attention to category-defining features of the familiar scenes.

REFERENCE NOTE

1. Hock, H. S., Whitehurst, R., & Throckmorton, R. *Alternative modes of processing and the retention of visual stimuli*. In preparation.

REFERENCES

- BARTLETT, F. C. *Remembering: A study in experimental and social psychology*. Cambridge, England: Cambridge University Press, 1932.
- BEVAN, W., & FEUER, J. N. The role of context in episodic memory. *Bulletin of the Psychonomic Society*, 1977, **10**, 76-78.
- BEVER, T. G. The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language*. New York: Wiley, 1970.
- BIEDERMAN, I. Perceiving real-world scenes. *Science*, 1972, **177**, 77-80.
- BIEDERMAN, I. On processing information from a glance at a scene: Some implications for a syntax and semantics of visual processing. In S. Treu (Ed.), *User-oriented design of interactive graphic systems*. New York: ACM, 1977.
- BIEDERMAN, I., GLASS, A. L., & STACY, E. W., JR. Searching for objects in real-world scenes. *Journal of Experimental Psychology*, 1973, **97**, 22-27.
- BIEDERMAN, I., RABINOWITZ, J. C., GLASS, A. L., & STACY, E. W., JR. On the information extracted from a glance at a scene. *Journal of Experimental Psychology*, 1974, **103**, 597-600.
- BROWN, A. L. Recognition, reconstruction and recall of narrative sequences by preoperational children. *Child Development*, 1975, **46**, 156-166.
- BROWN, A. L., & MURPHY, M. D. Reconstruction of arbitrary vs. logical sequences by preschool children. *Journal of Experimental Child Psychology*, 1975, **20**, 307-326.
- BROWN, R. *A first language: The early stages*. Cambridge: Harvard University Press, 1973.
- BRUNER, J. S. On perceptual readiness. *Psychological Review*, 1957, **64**, 123-152.
- CHOMSKY, N. *Syntactic structures*. The Hague: Mouton, 1957.
- FODOR, J. A., BEVER, T. G., & GARRETT, M. F. *The psychology of language: An introduction to psycholinguistics and generative grammar*. New York: McGraw-Hill, 1974.
- HOCK, H. S., GORDON, G. P., & CORCORAN, S. K. Alternative processes in the identification of familiar pictures. *Memory & Cognition*, 1976, **4**, 265-271.
- HOCK, H. S., GORDON, G. P., & GOLD, L. Individual differences in the verbal coding of familiar visual stimuli. *Memory & Cognition*, 1975, **3**, 257-262.
- HOCK, H. S., GORDON, G. P., & WHITEHURST, R. Contextual relations: The influence of familiarity, physical plausibility, and belongingness. *Perception & Psychophysics*, 1974, **16**, 4-8.
- HOROWITZ, L. M., LAMPEL, A. K., & TAKANISHI, R. The child's memory for unitized scenes. *Journal of Experimental Child Psychology*, 1969, **8**, 375-388.
- MANDLER, J. M., & JOHNSON, N. S. Some of the thousand words a picture is worth. *Journal of Experimental Psychology: Human Learning and Memory*, 1976, **2**, 529-540.
- MANDLER, J. M., & PARKER, R. E. Memory for descriptive and spatial information in complex pictures. *Journal of Experimental Psychology: Human Learning and Memory*, 1976, **2**, 38-48.
- MANDLER, J. M., & RITCHEY, G. H. Long term memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, 1977, **3**, 386-396.
- MANDLER, J. M., & STEIN, N. L. Recall and recognition of pictures by children as a function of organization and distractor similarity. *Journal of Experimental Psychology*, 1974, **102**, 657-669.
- PIAGET, J., & INHELDER, B. *The psychology of the child*. New York: Basic Books, 1969.
- ROSCHE, E., MERVIS, C., GRAY, W., JOHNSON, D., & BOYES-BREAM, P. Basic objects in natural categories. *Cognitive Psychology*, 1976, **8**, 382-439.
- TULVING, E. Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization and memory*. New York: Academic Press, 1972.
- WINER, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1971.

NOTES

1. Hock et al. (1974) referred to rules of physical plausibility rather than rules of physical possibility. We thank an anonymous reviewer for pointing out that "possible" is more appropriate than "plausible."

2. Other examples of the scenes used in this study are presented in the study by Hock et al. (1974).

3. Also significant in this experiment was the interaction between context type and groups [$F(3,192) = 8.38, p < .005$]. The groups variable involved the counterbalanced assignment of a particular scene as an "original" stimulus for one group of subjects, or as a "distractor" for a second group of subjects. The interaction between context type and groups is apparently the result of response bias. For example, subjects were biased toward choosing the scene with the desk in all pairs where they had a choice between scenes with a desk and scenes with a table. As can be seen in Table 1, which illustrates the experimental design, this response bias meant that subjects in Group 1 would tend to make more errors for Types 1 and 2 than Types 3 and 4, while subjects in Group 2 would tend to make more errors for Types 3 and 4 than Types 1 and 2. A significant interaction between context type and groups was also obtained in Experiments 2 and 3 of the present study. Since the effects of the response bias were controlled by counterbalancing, and the interaction is of little interest, it will not be given any further consideration.

4. Hock, Whitehurst, and Throckmorton (Note 1) have obtained evidence consistent with this explanation. They have found that real words were recognized less accurately than pronounceable nonwords and familiar dot patterns were reproduced less accurately than unfamiliar dot patterns. In both cases, this effect was linked with an analytic-categorical mode of processing (Hock, Gordon, & Corcoran, 1976; Hock, Gordon, & Gold, 1975).

5. Since there was an unequal number of subjects in the two groups, the analysis of variance followed the least-squares procedure (Winer, 1971).

6. Since subjects were matched in age on a 1:1 basis, conservers vs. nonconservers was treated as a repeated-measures variable in the analysis of variance.

7. Bevan and Feuer (1977), using a recall task with adult subjects, have obtained results that are very similar to the results of this study.

(Received for publication August 24, 1977;
revision accepted May 5, 1978.)