JOURNAL OF EXPERIMENTAL CHILD PSYCHOLOGY 24, 279–298 (1977)

Developmental Trends in Free Classification: Evidence for a New Conceptualization of Perceptual Development

LINDA B. SMITH AND DEBORAH G. KEMLER

University of Pennsylvania

Two studies are reported to explore the hypothesis that young children perceive *integrally* some stimuli that older children perceive *separably*. In both, kindergarteners, second graders, and fifth graders (approximately 5, 8, and 11 years old) are required to classify sets of stimuli that vary in size and brightness. Triads are used in Experiment 1 and tetrads are used in Experiment 2. Also, in Experiment 2, second classifications, judgments of which classification is "best," and verbal justifications for classifications are obtained. The general finding is that the kindergarten data systematically implicate integrality of size and brightness while the fifth-grade data systematically implicate separability of size and brightness. The second-grade data are more ambiguous. Issues related to refining the developmental hypothesis and to extending its supportive data base are considered in a final discussion.

One of the more consistent trends in perceptual development is the greater tendency in the older child to ignore irrelevant information (Gibson, 1969; Hagen & Hale, 1973; Maccoby, 1969; Wohlwill, 1962). The trend has been documented in selective listening tasks (Doyle, 1973; Maccoby, 1969), speeded classification tasks (Shepp & Swartz, 1976; Strutt, Anderson & Well, 1975), memory tasks (Hagen, 1967; Hagen & Hale, 1973), same–different classification tasks (Smith, Kemler, & Aronfreed, 1975), and discrimination learning tasks (Crane & Ross, 1967; Kemler, Shepp, & Foote, 1976).

Is the developmental difference that is indexed in tasks demanding selective attention symptomatic of a more pervasive difference in the way that children of varying ages apprehend complex stimuli? Smith et al. (1975) suggested that one source of young children's deficit in selective attention

Preparation of this article was supported in part by a Spencer Foundation Grant through the University of Pennsylvania to the second author and by National Science Foundation Grant BNS 76-08351 to the second author.

Our thanks go to Jonathan Baron and Christopher Foard for their comments on an earlier draft of this manuscript. We also gratefully acknowledge the cooperation of Michael Kirsch, Principal, and the teachers of the Chestnutwold elementary school in Haverford Township, Pennsylvania.

Requests for reprints should be sent to Deborah G. Kemler, Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19104.

ISSN 0022-0965

tasks may be their difficulty in initially differentiating the input into its relevant and irrelevant parts. Recently, Garner (1970, 1974a,b) has offered an analysis of the perception of multidimensional stimuli in adults that helps to refine such a hypothesis. Garner distinguishes between dimensional combinations that are perceived by the adult as the conjunction of separate components, *separable* dimensions, and those that are perceived by the adult as *integral*, unitary wholes. Within the unitary wholes formed by integral dimensions, purely selective processing of one dimension, as required in a selective attention task, is an impossibility. For example, brightness and saturation act as integral dimensions for adults. College students cannot rapidly sort stimuli on the sole basis of brightness without being disrupted by orthogonal variation of the saturation of the stimuli (and vice versa). Apparently, brightness and saturation are apprehended as a single unit by adults at a primary perceptual level. The contrasting case of separable dimensions, exemplified by stimuli in which the size of a circle and the angle of a radial line are varied, leads to no disruption in speed when sorting is required on the basis of one dimension.

To extend Garner's analysis to developmental issues is to suggest that *some dimensions that are perceived as separable by adults (and older children) may be apprehended as integral by young children*. As applied specifically to the selective attention problem, this extension implies that young children may fail to resist distracting irrelevant information, where older children succeed, not (solely) because they lack attentional control, but because the perceptual material on which the younger child must operate is unitary (where for older children it is the sum of separate components).

Of course, the analysis of such developmental differences in terms of stimulus integrality is underdetermined by the selective attention data. The value of the conceptualization depends on demonstrating its usefulness in ordering the data that are obtained in a wide variety of processing tasks. Conversely, such demonstrations would also argue for considering developmental differences in selective attention within a broader conceptual framework than has been previously applied. The general purpose of the present studies is to seek this type of converging evidence.

Garner has identified several operations that converge on the distinction between integral and separable dimensions in adults' perception (Garner, 1974a; see also Garner & Felfoldy, 1970). For example, in adults, the same pairs of dimensions that produce interference in speeded tasks requiring selective processing (integral dimensions) also (a) produce facilitation in speeded sorts when the dimensions are redundant, (b) produce estimates of interstimulus differences in direct distance scaling that are best fit by a Euclidean metric, and (c) produce classifications that are solely dependent on the similarity relations among the stimuli. On the other hand, separable dimensions that do not produce interference in the selective tasks also (a) do not produce facilitation when redundantly paired in speeded classification, (b) produce direct distance scaling data that are best fit by a city-block metric, and (c) produce classifications that are sensitive to the dimensional relations among the stimuli.

As we argued above, the value of using the integrality–separability distinction to conceptualize developmental differences in perceptual processing will depend analogously on finding converging evidence from diverse task settings. Some relevant data now exist. Within a discrimination transfer paradigm, Tighe and Tighe (1972) have demonstrated that, whereas older children tend to learn and transfer on the basis of the dimensional components of the stimuli, younger children appear to learn about "compounds" or holistic stimulus "objects." This finding is perfectly consistent with the notion that younger children primarily perceive complex stimuli as integral units, whereas older children perceive many of them as separable. Shepp and Swartz (1976) provide further evidence from a paradigm borrowed directly from Garner and Felfoldy (1970). First- and fourth-grade children speed-sorted three types of stimulus decks on the basis of one named dimension. In the Control deck only the named dimension varied. In the Orthogonal deck, an irrelevant dimension varied in a manner orthogonal to the named dimension, and in the Redundant deck, it varied redundantly with the named dimension. When color and shape were paired, fourth-graders' speeded-sorting patterns implicated stimulus separability: no differences in sorting speed across the three types of decks. But first graders showed the typical pattern for stimulus integrality: interference in the Orthogonal deck and facilitation in the Redundant deck. Again, the suggestion is one of increasing stimulus separability with age.

Garner (1974a) has attempted to deal conceptually with the question, "What is dimensional integrality?" He says:

> Psychologically, if dimensions are integral, they are not perceived as dimensions at all. Dimensions exist for the experimenter, and they may even exist in elaborately calculated multi-dimensional solutions for data from similarity scaling experiments. But these are concepts that are highly derived and do not reflect the immediate perceptual experience of the subjects. . . .
>
> On the other hand, if the dimensions are separable, distances do not really exist for the subject. What he perceives are dimensions, and . . . relative distances on each of two dimensions have little to do with how a subject perceives and classifies a set of stimuli (Garner, 1974a, p. 119).

A nonspeeded classification task would seem to be the most direct operation to tap into this conceptually important distinction between the perception of integral and separable stimuli. It is possible to select groups of stimuli that subjects may partition one way by using similarity relations or another way by using dimensional relations as the basis for their sort (see Handel & Imai, 1972). For example, consider a stimulus set consisting of a small black form, a small white form and a slightly larger form which is light gray. A dimensional classification would be realized by putting together the first two stimuli, those that are *identical on the dimension* of size. A similarity classification would be realized by putting together the last two stimuli, those that are *most similar overall*. A dimensional classification implicates perceived stimulus separability. A similarity classification implicates perceived stimulus integrality. Here, we will pursue the developmental hypothesis of increasing stimulus separability with age using just such free classification tasks.

EXPERIMENT 1

Method

Subjects. The subjects were 30 students attending kindergarten, second, or fifth grade at a suburban Philadelphia elementary school. Within each of the grades five female and five male children were chosen at random to participate. The mean age was 5 years, 8 months (range: 5 years, 2 months to 6 years, 1 month), 7 years, 11 months (range: 7 years, 3 months to 8 years, 3 months), and 10 years, 11 months (range: 10 years, 2 months to 11 years, 6 months) for kindergarten, second, and fifth grades, respectively.

Stimuli and design. The stimulus set consisted of varying size and brightness combinations within a constant irregular quadrilateral form. The forms, mounted on 4 × 6-in. white cards, were cut from six sheets of Coloraid paper that ranged from almost white to black (Coloraid No.: 1, 2, 4, 5, 7, black). The six different sizes had the following areas: .42, .72, 1.00, 1.28, 2.59, and 3.88 $in.^2$. Adjacent values on each of the dimensions were discriminably different from one another.

In order to determine approximately equal intervals of psychological distance on each of the dimensions, eight undergraduate students rated the similarity of pairs of stimuli differing on one dimension. The method of magnitude estimation was used. From these data, 16 different triads were prepared for use in the experiment.

These 16 triads were of three different *types* as shown in Fig. 1. Type I and II triads both pit similarity relations against dimensional relations. They contain two stimuli (A and B) that share an identical value on one dimension but differ substantially on the other. One of these two stimuli (A) differs only slightly, but on *both* dimensions, from the third stimulus (C) in the triad. As determined using adults' similarity judgments, the sum of the two differences between A and C is always smaller than the one-dimensional difference between A and B. Type I and Type II triads differ from each other only in the noncritical relation between stimuli B and C.

As can be seen in Fig. 1,there are three types of classifications that can be imposed on Type I and Type II triads. The subjects can make (a) a dimensional classification (DIM), i.e., put the two stimuli that share an identical value in the same group; (b) a similarity classification (SIM), i.e., put the two stimuli that differ only slightly on both dimensions in the same group;

Size x Brightness Triads

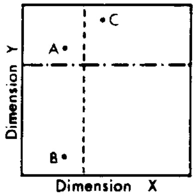
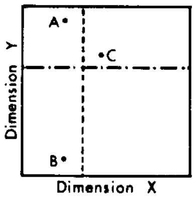
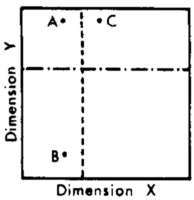
Triad Type	Results	
	Grade	Classification
I 		DIM
	K	.40 (.11)
	2	.57 (.18)
	5	.66 (.25)
II 		DIM
	K	.34 (.07)
	2	.50 (.22)
	5	.68 (.22)
III 		DIM ONLY
	K	.34 (.12)
	2	.41 (.09)
	5	.41 (.17)

FIG. 1. In the left column, schematic representations of the size × brightness triad types, drawn so that values of the two dimensions are represented along the horizontal and vertical axes. Distances along the axes represent perceived psychological differences. The regularly broken line in each schematic indicates a dimensional partition; the irregularly broken line indicates a similarity-maximizing partition. In the right column are mean proportions of systematic classifications that were of the first kind. Standard deviations are in parentheses. See text for further clarification.

or (c) a haphazard classification, i.e., put together two stimuli which differ considerably on both dimensions.

Type III triads are of a different kind in that both the pair A and B and the pair A and C share a value. Thus, there are potentially two dimensional classifications for these triads. However, in one such classification subjects can put together two stimuli which share a value on one dimension and differ substantially on the other (DIM ONLY). In the second dimensional classification, subjects can group together two stimuli which share one value and differ only slightly (as measured by the similarity judgments) on the other dimension (DIM + SIM). There is a stronger expectation for this type of classification (that maximizes intragroup similarity) to predominate when the stimuli are perceived integrally than when they are perceived separably. A third response (haphazard) is also available: The subject can put together two stimuli which differ considerably on both dimensions.

There were six unique Type I and four unique Type II triads. A unique configuration was constructed by selecting a specific value on the shared dimension and appropriately constrained values on the nonshared dimension. For exactly half the triads of each type,the shared dimension was size; for the other half, brightness. There were six unique Type III triads. For half these triads, the two stimuli that shared a value of size were also the most similar pair; for half the opposite was the case.

The triads were presented in a randomly ordered series of 58 trials, with each Type I and Type II triad repeated four times and each Type III triad repeated three times within the series.

Procedure. The experimenter faced the subject across a table and set out three stimuli (either in a row or as if at the vertices of a triangle). The subject was asked to "put together in one group the two that go together." "Which two most go together?" The words "most alike" or "most similar" were not used. Periodically, the experimenter encouraged the subject, but no specific feedback was given. Presentation of the 58 trials took approximately 30–45 min.

Results and Discussion

The interesting data consist of the relative frequencies of the two types of systematic classifications. In fact, haphazard responses occurred infrequently, even within the kindergarteners' classifications. For no triad within any age group did the frequency of haphazard partitions account for more than 10% of the observed responses. For every subject, the number of haphazard responses across triad types was reliably less than would be expected if sorting were simply on a random basis.

The proportion of systematic classifications (SIM or DIM) that were dimensional (DIM) was computed for each subject for Type I and Type II triads. The proportion of systematic classifications (DIM ONLY or DIM + SIM) that were only dimensional (DIM ONLY) was computed for each subject for Type III triads. The mean proportion of such classifications for each grade and triad type are shown in Fig. 1. Type III triads were analyzed separately from Type I and Type II triads, as Type III triads offered two dimensional solutions and therefore did not put a similarity solution in direct conflict with a dimensional solution.

Type I and Type II triads. The proportions of DIM classifications for Type I and Type II traids were submitted to an analysis of variance for a

3 (Age) × 2 (Triad type) × 2 (Dimension that shared a value) mixed design. The analysis revealed a significant main effect of Age [$F(2,27) = 6.93$, $p < .01$]. As can be seen in Fig. 1, the proportion of dimensional responses increased reliably with age. No other main effects or interactions approached significance.

To determine whether, at each age level, similarity or dimensional classifications predominated, the mean proportion of DIM responses for each Age on each Type was compared with the value expected (.50) if subjects were dividing their responses equally between the SIM and DIM classifications. On both Type I and Type II triads, kindergarteners gave DIM responses significantly *less* than 50% of the time [$t(9) = 2.90$, $p < .05$, and $t(9) = 7.33$, $p < .001$, respectively, for Types I and II]. Second graders' responses did not differ reliably from the value expected by chance [$t(9) = .77$ and $t(9) = 0$, for Type I and Type II triads, respectively], and fifth graders gave DIM responses significantly *more* than 50% of the time [$t(9) = 3.16$, $p < .05$, and $t(9) = 2.64$, $p < .05$ for Type I and Type II triads, respectively]. So, whereas kindergarteners responded with the similarity classification more than with the dimensional classification on Type I and Type II triads, second graders' responses were divided about equally between the two classifications, and fifth graders responded more often with the dimensional classification.

Type III triads. In a separate analysis of the Type III data, systematic responses that did not maximize similarity (DIM ONLY responses) were submitted to an analysis of variance for a 3 (Age) × 2 (Dimension shared by most similar pair) mixed design. This analysis revealed no main effects or interactions that approached significance. As may be seen in Fig. 1 the predominant response was the dimensional classification that maximized the similarity of the stimuli on both dimensions [$t(29) = 3.14$, $p < .01$]. Thus, overall similarity is used as an ancillary determinant of classifications even in older subjects as long as it is not in conflict with a dimensional partition.

Conclusions

The data of Experiment I appear to provide straightforward support for the developmental hypothesis of increasing stimulus separability with age. Size and brightness, as indexed by the basis for classifications, are treated more frequently as integral dimensions by kindergarteners and as separable dimensions by fifth graders, with the second-grade data falling regularly in between. These intermediate second-grade data do not appear to result from any identifiable systematic trends at this age level. Six second graders' classifications divide about equally between SIM and DIM classifications on Type I and Type II. Of the remaining second graders, two subjects' performances pattern in the fifth-grade manner; the other two pattern in a kindergarten manner.

Some additional findings should be highlighted. First, systematic classifications, as distinguished from haphazard groupings, are produced even by kindergarteners. Second, and in a related vein, kindergarteners' systematic responses tend to be of a consistent form across stimulus variation. Particularly notable is that the predominant use of the SIM classification by the kindergarteners overrides the variable of which dimension in a set presents the shared value. Although a few subjects at all age levels do show a reliable tendency to group on the basis of one dimension in preference to the other, the group data are clear in indicating that systematic classification on the basis of stimulus *structure* dominates kindergartener and fifth-grade classifications.

Would the same systematicity of classificatory responding in young children occur if the dimensions of color and form were used to construct the stimulus sets? These are dimensions for which strong preferences (that change with age) have been reported in children (e.g., Suchman & Trabasso, 1966). Of course, with such nominal

Color x Form Triads

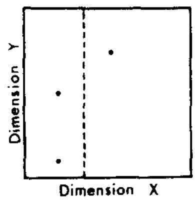
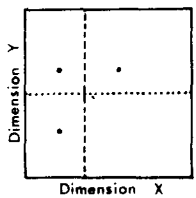
Triad Type		Results: Grade	Results: Classification
A			DIM
		K	.96 (.09)
		2	.96 (.13)
		5	1.00 (.00)
B			FORM
		K	.65 (.34)
		2	.63 (.30)
		5	.93 (.21)

FIG. 2. In the left column are schematic representations of the two color × form triad types. The broken line for Type A indicates the single possible dimensional partition. The two broken lines for Type B indicate the two possible dimensional partitions. In the right column are mean proportions of certain classifications for each type and grade level. Standard deviations are given in parenthesis. See text for further clarification.

dimensions, the triad configurations that pit similarity against dimensions cannot be constructed. Overall similarity is a rather meaningless concept here, or else it is measurable as the number of shared values, and is thus indistinguishable in practice from dimensional relations. However, one can still ask of children's classifications of triads of color and form whether they are systematic or haphazard across a set in which sometimes color or sometimes form is the only consistent basis for a partition.

Figure 2 shows the results we obtained when children were presented with two types of color and form triads. The most interesting type of set (Type A) was constructed to allow only one possible classification based on a shared value (which could be color or form). The other triad (Type B) simply pitted a color match against a form match. All subjects consistently sorted both types of triads by matching on one dimension. When two different dimensional matches were available in the Type B triads, there was a reliable increase with age in the choice of a form-matching partition [$F(2,28) = 9.00$, $p < .01$]. This trend is consistent with the dimensional preference literature in children. However, by far the more important result is the systematic approach to the classification problem evidenced at all ages. "Putting things together that go together" is an instruction that taps into reliable perceptual phenomena at all age levels tested. The main data of Experiment 1 suggest, beyond this, that the nature of what is primarily apprehended changes with age.

EXPERIMENT 2

Experiment 2 was undertaken in order to evaluate the generality of the classification results obtained in the first study. To what degree are the observed developmental patterns independent of the specific task administered to the subject? Consider that the systematic partitioning of triads can be accomplished by an extremely rapid appraisal of the stimulus set. In fact, it was our strong impression in Experiment 1 that the younger children, more than the older children, relied on "first impressions" to generate their sorts. The general issue of whether findings of stimulus integrality and separability depend upon the nature of the subject's mental set has received less attention than it deserves. The issue may be a matter of particular importance in interpreting developmental data. In Experiment 2, therefore, we asked whether the younger children will continue to produce a predominance of similarity-based classifications (and older children a predominance of dimensionally-based classifications) under conditions that force them to be more reflective.

Two types of modifications of the earlier procedure were used in addressing the general question. First, in Experiment 2, the stimulus sets were expanded to include four stimuli (tetrads), to be sorted into two groups of two, under the assumption that such increased task demands should increase reflective responding. If one slows down the responding of younger children by making the task more difficult, will they maintain their preference for similarity sorts? Second, in Experiment 2, some additional questions were posed to each subject. After generating a first sort, the subject was requested to attempt a different one; if two sorts were thus elicited, the subject was requested to select the "best" one. Given that a subject has generated both a similarity and a dimensional classification, will younger children favor the first and older children favor the second? Also, verbalizations of the subjects' criteria for classifying were tapped. By each of these manipulations, the context for interpreting the original sorting data of Experiment 1 is enriched.

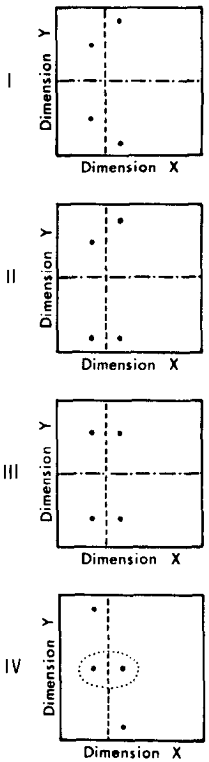
Method

Subjects. The subjects were 30 students attending the same elementary school as those of Experiment 1. The mean ages of those participating was 5 years, 9 months (range: 5 years, 6 months to 6 years, 4 months), 8 years, 6 months (range: 7 years, 8 months to 8 years, 7 months), and 10 years, 10 months (range: 10 years, 6 months to 11 years 7 months), respectively, for kindergarten, second-grade, and fifth-grade children. Five male and five female children were chosen randomly from each grade to participate.

Stimuli and design. The individual stimuli were identical to those of Experiment 1, but they were here presented in sets of tetrads rather than triads. Fourteen unique tetrads of four different types were arranged. The four types are shown in the leftmost column of Fig. 3. Unique instances of the tetrad types were devised in a manner analogous to that for the formation of Experiment 1 triads.

Tetrad Type I and Type II offer conflicting similarity and dimensional classifications. For Type I tetrads, a similarity response consists of forming two groups such that the two members of each group differ slightly on both dimensions. For Type II tetrads, a similarity response consists of forming one group in which the two members differ slightly on both dimensions and one group in which the two members share a value on one dimension and differ slightly on the other. For both of these types, the similarity classification minimizes intergroup similarity while maximizing intragroup similarity. There were four unique Type I tetrads and four unique Type II tetrads. For each type, half of the potential dimensional classifications require partitioning on the basis of shared sizes, and half require partitioning on the basis of shared brightness values.

Size x Brightness Tetrads



Tetrad Type	Grade	Classification	
I		DIM	SIM
	K	.34(.16)	.58(.15)
	2	.36(.24)	.58(.23)
	5	.80(.19)	.18(.16)
II		DIM	SIM
	K	.28(.20)	.61(.26)
	2	.35(.21)	.57(.24)
	5	.68(.16)	.31(.18)
III		DIM ONLY	DIM+SIM
	K	.20(.20)	.70(.20)
	2	.19(.11)	.75(.16)
	5	.28(.18)	.72(.18)
IV		DIM	ONE-SIDED
	K	.53(.14)	.29(.17)
	2	.70(.21)	.18(.15)
	5	.90(.17)	.05(.10)

FIG. 3. In the left column are schematic representations of the four size × brightness tetrad types. Representation is as in Fig. 1. For each type, two partitions are shown. Those given by the regularly broken straight line indicate dimensional partitions. Those given by the irregularly broken line indicate similarity-maximizing partitions. That given by the dotted line indicates a one-sided partition. In the right columns are mean proportions of various classifications for each type and grade level. Standard deviations are given in parenthesis. See text for further clarification.

Type III tetrads offer two dimensional solutions. Each is constructed as the set of all combinations of two specific values on each dimension. As shown in Fig. 3, one dimensional classification maximizes both similarity within groups and dissimilarity between groups relative to the other. There were two unique Type III tetrads. In one, size was the dimension that maximized intragroup similarity; in the other, brightness was the dimension.

Type IV tetrads offer one dimensional classification and no clear similarity classification. This type was included as a check on subjects' abilities to follow the instructions. Of specific concern was whether subjects' responses would indicate the formation of two groups of two stimuli or whether, contrary to instruction, subjects might simply select the two most similar stimuli for one group, leaving the other two as "left over." The latter "one-sided" response is put in conflict with the dimensional partition in the design of Type IV tetrads, as shown in Fig. 3. There were four unique Type IV tetrads, on exactly half of which a dimensional partition would involve size and half of which, brightness.

The tetrads were arranged in a series of 42 trials, with each unique tetrad repeated three times. The series was randomly ordered except that the 14 unique instances occurred exactly once during the last third of the trials.

Procedure. The experimenter faced the subject across a table top that was bisected by a red line. On each trial, the subject was handed a stack of four stimulus cards, haphazardly ordered, and told to "make two groups": "Put two over here (on one side of the red line) so that they go together, and put two over here (on the other side) so that they go together." On the first two-thirds of the trials, the experimenter periodically encouraged the subject, but gave no specific feedback.

On the last third of the trials, however, the subject was asked to justify his first (primary) classification and to give another (second) classification. The following series of questions was used. Indicated in parentheses are the types of follow-up questions that were added if the subject failed to answer or comprehend the first form of the question.

1. *Is this a good way* (relative to the primary response)? *Why? How come you did it this way? Why do these two go together* (indicating one group)? *Why do these go together* (indicating other group)? *Is there anything alike about these two? Anything alike about these two?*

2. *Could you do it another way? How?* (*How might somebody else do it? Try it another way.*) (If a subject strongly asserted that there was only one way to do it or offered a justification for that assertion, he was not prodded further to produce a second classification.)

3. Repeat of Question 1 above in relation to the subject's second classification, if produced.

4. *Which way is the best way? (This way or the one before? Are you sure?) Why? Show me the best way again* (as the experimenter restacks the stimuli and presents them again for classification).

In most cases, all data from a single subject were collected within one experimental session of 30–60 min in duration. For three kindergarteners, however, two sessions were required.

Results and Discussion

We had the strong impression in collecting the data of Experiment 2 that the use of tetrads, as opposed to triads, did prompt more reflective classifying in the youngest subjects. Specifically, the kindergarten children took a good deal more time to partition tetrads than triads. Whereas the kindergarteners responsed in Experiment 1 by rapidly indicating their partitions, they responded cautiously and sometimes laboriously in Experiment 2. Frequently, their final classifications were produced after several preliminary and tentative attempts. Sometimes, kindergarteners commented spontaneously on the difficulty of the task.

Type IV tetrads. In light of these observations, it is important to examine the results for the Type IV tetrads, preliminary to addressing the main data on integrality–separability. The Type IV results provide a specific check on whether the children's classifications respond to the instruction to *partition* the tetrads (into two groups of two) as opposed to the lesser demand to select *one* pair of stimuli that go together (leaving the remainder as the second pair). For the Type IV tetrads, producing a *partition* of the set, such that consistent relations hold within both groups (a DIM response, in this case), is put into conflict with the strategy of choosing for one group the one pair of stimuli that is most similar overall (a ONE-SIDED response).

Primary classificatory responses for the Type IV tetrads are given in the bottom rows of Fig. 3. The proportions of DIM responses were submitted to an analysis of variance for a 3 (Age) × 2 (Dimension that had two values) mixed design. The analysis did yield a significant main effect of Age [$F(2,27) = 11.69$, $p < .01$], but that alone. Application of Tukey's method for pairwise contrasts ($\alpha = .05$) revealed that at each higher age level, the proportion of DIM responses increases reliably. Still, as can be seen in Fig. 3, the predominant response at all age levels was the DIM classification. Thus, even under the demanding conditions established by the Type IV tetrad, subjects tended to follow the instruction to produce *two* well-formed groups. As even stronger supporting evidence, the dimensional classification, if given, was judged as the "best" classification on 70, 74, and 100% of the trials by kindergarten, second-grade, and fifth-grade subjects, respectively.[1]

Type I and Type II tetrads: Similarity vs dimensional responses.

[1] For all tetrad types, it is notable that the "best" judgment data provide information that is almost independent of the information in the primary classificatory data. Subjects chose their first classification as the "best" classification on .56, .59, and .69 of the trials for kindergarten, second grade, and fifth grade, respectively.

Tetrad Type I and Type II offer conflicting similarity and dimensional classifications and thus provide the most information about stimulus integrality or separability. The proportions of total primary responses to these tetrads that were DIM classifications are given in Fig. 3.[2] These proportions were submitted to an analysis of variance for a 3 (Age) × 2 (Type) × 2 (Dimension that shared two values) mixed design. The analysis yielded only a significant main effect of Age [$F(2,27) = 6.93, p < .01$].

As shown in Fig. 3, the age trends in classificatory responses on these tetrads are quite similar to those observed in Experiment 1. Kindergarteners gave SIM responses to these tetrad types reliably more often than DIM responses [$t(9) = 2.25, p < .05$]. Fifth graders, on the other hand, gave DIM responses reliably more often than SIM responses [$t(9) = 4.25, p < .01$]. Although the mean proportion of SIM responses is greater than the mean proportion of DIM responses in second graders, this difference is not reliable [$t(9) = 1.60, p > .05$]. (As in Experiment 1, the mean proportion of dimensional responses in the group of second graders is representative of individual subject's performances.)

On the last third of the trials, when asked to give two classifications, subjects tended to produce both SIM and DIM responses. The mean proportions of relevant trials on which both responses are given are .67, .71, and .56 for kindergarten, second-, and fifth-grade subjects, respectively. A one-factor (Age) analysis of variance did not reveal a reliable effect [$F(2,27) = .62$]. Table 1 shows the conditional probabilities of "best" judgments on Type I and Type II tetrads. (These probabilities sum to less than 1 due to a few cases in which subjects judged the two classifications equally good.) Four separate one-factor (Age) analyses of variance were computed on the data. There are unequal n's due to the nonoccurrence of some solutions in some subjects (e.g., SIM responses by some fifth graders). Therefore, the analyses were computed via the least-squares method for unequal group sizes. All four analyses yielded significant effects of Age. As shown in Table 1, the probability that a DIM response is judged "best" increases with age, while the probability that a SIM response is judged "best" declines with age.

Thus, the observations of Experiment 1 appear to be stable and generalizable. Younger children produce the SIM response predominantly even in the more difficult tasks of Experiment 2. Moreover, when asked whether a self-generated DIM or SIM classification

[2] Inspection of the data failed to reveal any differences in the types of classifications offered on the last third of the trials, when verbalizations were elicited, and those offered on the earlier trials. Therefore, for all analyses of the primary responses, the data were simply collapsed over trial blocks.

TABLE 1

CONDITIONAL PROBABILITIES FOR THE CLASSIFICATION THAT WAS CONSIDERED THE "BEST" ONE[a]

	Grade	Type I	Type II
DIM Best | DIM and SIM	K	.31 (.19)[b]	.33 (.27)
	2	.47 (.35)	.56 (.35)
	5	.86 (.26)	.71 (.32)
SIM Best | DIM and SIM	K	.64 (.23)	.60 (.38)
	2	.40 (.35)	.30 (.34)
	5	.13 (.26)	.03 (.17)

[a] For example, the first conditional probability reads: the probability that a dimensional classification was judged "best" given there were both dimensional and similarity classifications on that trial.

[b] Standard deviations are given in parentheses.

is the better one, younger children maintain the SIM classification and older children maintain the DIM classification as the one of choice.

Type III tetrads. Type III tetrads offer two dimensional solutions. One (DIM + SIM) maximizes overall similarity within each group more than does the second dimensional solution (DIM ONLY). As shown in the third block of rows in Fig. 3, all subjects tended to classify these tetrads such that overall similarity was maximized. The proportions of DIM + SIM responses were submitted to a one-factor (Age) analysis of variance from which no significant effect emerged. So it appears that all age groups are sensitive to the values on both dimensions. As in Experiment 1, even older children will maximize overall similarity as long as they still can produce a dimensional classification.

When asked to give a second response on the last third of the trials, all subjects tended to give the alternative dimensional classification. The mean proportions of such trials on which both dimensional responses (DIM + SIM and DIM ONLY) are given are: .65, .75, and .90 for kindergarteners, second graders, and fifth graders, respectively. The increase with age in the production of both classifications is consistent with the hypothesis of increasing separability of dimensions with age. If dimensions are separable and the rule is to group objects such that they share a value, then the two dimensional classifications are equally acceptable. If, on the other hand, dimensions are integral and the rule is to group on the basis of overall similarity, then one classification (DIM + SIM) is clearly better than the other (DIM ONLY). Table 2 shows the data for "best" judgments on Type III tetrads. As expected, judgments of DIM + SIM responses as "best" decline with age, and judgments that both solutions are equally good increase with age. Once again, the primary and secondary classification data

TABLE 2

MEAN PROPORTIONS OF "BEST" JUDGMENTS FOR TYPE III TETRADS[a]

Grade	DIM + SIM	DIM ONLY	Both
K	.65 (.24)[b]	.30 (.26)	.00 (.00)
2	.55 (.37)	.10 (.21)	.30 (.41)
5	.40 (.39)	.10 (.21)	.50 (.41)

[a] Showing preferences for the DIM + SIM classification, the DIM ONLY classification, or a judgment that both were equally good.
[b] Standard deviations are given in parentheses.

and the judgment data together suggest that the results of the first experiment are persistent and stable.

Justifications of classifications. Whether justifying a dimensional classification or a similarity classification, all subjects tended to verbalize some criterion that is, in fact, true of the partition. The predominant verbal responses in justification of DIM and DIM + SIM responses mention either the dimension that is shared (e.g., size) or the specific values that are shared (e.g., little and big). The mean proportions of such verbalizations, given a dimensional classification, are .78, .88, and .88 for kindergarteners, second graders, and fifth graders, respectively. Analogously, the predominant verbal responses in justification of SIM responses mention the *one* dimension that maximizes intergroup differences: .67, .70, and .60 of the verbalizations are of this kind for kindergarteners, second graders, and fifth graders, respectively. Of some further interest is that no subject ever implies identity on a dimension when such identity does not obtain; thus, "They are *almost* the same size," in justification of one group in a similarity partition.

In contrast to the classification data themselves, developmental differences are minimal in the justification data. There is some tendency for younger children, more frequently than older children, to mention *both* dimensions in justification of similarity responses: .22, .15, and .00 of the responses at the kindergarten, second-grade, and fifth-grade levels. But perhaps more notable is the fact that even the youngest subjects do use *dimensional* terms. One might have expected them to use more global descriptions given their preference for similarity partitions. We explore this issue in the next section.

GENERAL DISCUSSION

The patterns in children's classifications in Experiments 1 and 2 are consistent with the developmental hypothesis of increasing stimulus separability with age. These data in conjuction with the speeded-sorting results of Shepp and Swartz (1976) provide evidence by three (of Garner's four) converging operations that young children apprehend in an integral manner stimuli that are perceived separably by adults. The conceptual implications of this convergence are explored below.

The verbalization data, troublesome as they may appear initially for the developmental hypothesis, provide an excellent springboard for the relevant discussion. The kindergarten children produce classificatory responses in which the dimensional structure of the stimuli is not given primacy, as one would expect for integral perception; yet, at the same time, they use dimensional terms to describe their classifications. Is the essence of the concept of stimulus integrality violated by these last findings?

An uninteresting, but nevertheless plausible way to accommodate the kindergarteners' verbalization data is to suppose that kindergarteners, like second graders, are in a state of transition between integral and separable perception of size and brightness. The implication is that still younger children would fail to produce any dimensional criteria in justifications of classifications as well as in the classifications themselves.

However, two different possibilities can be elaborated in a manner that is of additional interest. First, the attribution of integral perception does not imply that the subject cannot access the underlying dimensions under any and all conditions. In a discussion of how adults process integral dimensions, Lockhead (1972) has pointed out that the distinction between integrality and separability should not be confused with the distinction between analyzability and nonanalyzability [Garner (1974a) concurs]. Whereas integrality–separability describes the *primary* mode of perception, analyzability–nonanalyzability describes the *possibility of analysis* into dimensions, if necessary at a more derived or higher level mode of processing. Saturation and brightness are both integral and analyzable for adults. Analogously, we may say that size and brightness are integral and analyzable for 6-year olds. By this line of reasoning, the classification data tap their primary mode of perception, but the verbal justifications are accessing a more derived perceptual mode.

In a still different vein, the incongruent findings can be reconciled by a second clarification of the concept of stimulus integrality. As Garner (1974a) has suggested, the distinctive feature of perception in the integral mode is that the axes that mark the dimensions in multidimensional space have no special status to the perceiver relative to all other possible axes in the space. Consistent with this view is the notion that subjects, perceiving in an integral mode, still may be able to distinguish axes in this space, so that changes in size and changes in brightness are not apprehended as changes of the same sort. In clear contrast to when they are separable, changes in size alone have no more special status than simultaneous variation in size and brightness. By this analysis, kindergarteners, then, as shown by their tendency to group stimuli on the basis of overall similarity, do perceive these stimuli integrally. Their correct mention of the dimensions in their verbalizations shows that they are sensitive to direction as well as to proximity in the multidimensional space formed by size and brightness. But their use of terms referring to the two *dimensional* axes is exclusive only because the language restricts them to such terms. As a corollary, if English contained a term meaning "increasing simultaneously and regularly in size and brightness," the kindergarteners, but not the fifth graders, would sometimes use it to describe their classifications.

Clearly, still more converging operations would benefit an understanding of what it means to perceive multidimensional stimuli in an integral versus a separable mode. They would strengthen the analysis of developmental problems as well. That both the operations and parallel developmental differences should be pursued is highlighted by the finding of converging evidence in the free classifications data of the present study. We suggest that the classification task tapped real and systematic differences in how younger and older children apprehend some multidimensional stimuli. Younger children are not just processing poorly; rather they are processing differently. And, it appears, the specific way that younger children are processing differently is identical to the way that adults process under other stimulus conditions. The way that young children are processing may also have parallels in the ways that natural concepts are organized in the mind: Rosch and Mervis (1975) have argued that many natural concepts are structured by overall perceptual similarity rather than by the possession of a few criterial features.

The position that younger children perceive some stimuli in a different manner from older children and adults is one that has been adopted by many developmental theorists. Often, it has been elaborated as the difference between holistic, undifferentiated perception in the young child and selective, differentiated perception in the older child (e.g., Gibson, 1969; Werner, 1961). The present conceptualization of developmental differences in perception offers real promise of refining and extending such notions. By characterizing the perceptual mode of young children as "integral" rather than as "undifferentiated," we adopt a substantive description that carries theoretical and operational weight in and of itself. Young children's perceptual mode has its own systematic properties. These properties are stated and potentially elaborated within a system, that of Garner, that is also being applied to stimulus effects in adult perception. In light of the present results, the term "undifferentiated perception," if it is to imply unstructured perception, does not capture the regularities in young children's mode of stimulus apprehension. Young children's perception is organized around a similarity structure as older children's perception is often organized around a dimensional structure.

REFERENCES

Crane, N. L., & Ross, L. E. A developmental study of attention to cue redundancy. *Journal of Experimental Child Psychology*, 1967, **5**, 1–15.

Doyle, A. Listening to distraction: A developmental study of selective attention. *Journal of Experimental Child Psychology*, 1973, **15**, 100–115.

Garner, W. R. The stimulus in information processing. *American Psychologist*, 1970, **25**, 350–358.

Garner, W. R. *The processing of information and structure*. Potomac, Md.: Lawrence Erlbaum Associates, 1974 (a).

Garner, W. R. Attention: The processing of multiple sources of information. In E. C. Carterette and M. P. Friedman (Eds.), *Handbook of perception* (Vol. 2). New York: Academic Press, 1974(b).

Garner, W. R., & Felfoldy, G. L. Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology*, 1970, **1**, 225–241.

Gibson, E. J. *Principles of perceptual learning and development*. New York: Appleton–Century–Crofts, 1969.

Hagen, J. W. The effect of distraction on selective attention. *Child Development*, 1967, **38**, 685–694.

Hagen, J. W., & Hale, G. H. The development of attention in children. In A. Pick (Ed.), *Minnesota symposia on child psychology* (Vol. 7). Minneapolis, Minn.: University of Minnesota Press, 1973.

Handel, S., & Imai, S. The free classification of analyzable and unanalyzable stimuli. *Perception and Psychophysics*, 1972, **12**, 108–116.

Kemler, D. G., Shepp, B. E. & Foote, K. E. The sources of developmental differences in children's incidental processing during discrimination trials. *Journal of Experimental Child Psychology*, 1976, **21**, 226–240.

Lockhead, G. R. Processing dimensional stimuli: A note. *Psychological Review*, 1972, **79**, 410–419.

Maccoby, E. E. The development of stimulus selection. In J. P. Hill (Ed.), *Minnesota symposia on child psychology* (Vol. 3). Minneapolis, Minn.: University of Minnesota Press, 1969. Pp. 68–96.

Rosch, E., & Mervis, C. Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 1975, **7**, 573–605.

Shepp, B. E., & Swartz, K. B. Selective attention and the processing of integral and non-integral dimensions: A developmental study. *Journal of Experimental Child Psychology*, 1976, **22**, 73–85.

Smith, L. B., Kemler, D. G., & Aronfreed, J. Developmental trends in voluntary selective attention: Differential effects of source distinctness. *Journal of Experimental Child Psychology*, 1975, **20**, 352–362.

Strutt, G. F., Anderson, D. R., & Well, A. D. A developmental study of the effects of irrelevant information on speeded classification. *Journal of Experimental Child Psychology*, 1975, **20**, 127–135.

Suchman, R. G., & Trabasso, T. Color and form preference in young children. *Journal of Experimental Child Psychology*, 1966, **3**, 177–187.

Tighe, T. J., & Tighe, L. S. Stimulus control in children's learning. In A. Pick (Ed.), *Minnesota symposia on child psychology* (Vol. 6). Minneapolis, Minn.: University of Minnesota Press, 1972. Pp. 128–157.

Werner, H. *Comparative psychology of mental development* (Rev. ed.). New York: Scientific Editions, 1961.

Wohlwill, J. From perception to inference: A dimension of cognitive development. *Monographs of the Society for Research in Child Development*, 1962, **72,** 87–107.

RECEIVED: July 9, 1976; REVISED: October 1, 1976.