Perceptual comparisons through the mind's eye

ALLAN PAIVIO

University of Western Ontario, London, Ontario, Canada N6A 5C2

Four experiments tested a theory of memory and cognition which assumes that verbal and nonverbal information are processed in functionally distinct LTM systems. Subjects presented with pairs of pictures or printed names of animals and objects differing in rated real-life size were instructed to choose the conceptually larger member of each pair, the one that appeared to be farther away, or the one whose name was easier to pronounce. The following results were consistent with theoretical predictions: (1) RT to choose the larger member increased as the memory size difference decreased, for comparisons between as well as within conceptual categories; (2) memory size comparisons were faster with pictures than with words, whereas the reverse occurred for pronounceability comparisons; (3) with pictures, but not with words, size comparisons were significantly slower when real-life (memory) size relations conflicted with physical size relations than when the two were congruent; and (4) the size congruency effect was reversed for relative distance judgments of pictured pairs. These results cannot be easily explained by current verbal coding or abstract (propositional) theories of LTM representations.

Moyer (1973) reported a study in which subjects compared the sizes of named animals from memory. The animal names had been previously ranked by the subjects according to their judgments of the relative sizes of the animals themselves. In the experimental task, subjects were visually presented the names of two animals, such as *frog-wolf*, and were required to throw a switch under the name of the larger animal. The result was that the reaction time (RT) for the choice increased systematically as the difference in animal size became smaller. Specifically, RT was largely an inverse linear function of the logarithm of the estimated difference in animal size. Since this function is similar to the one obtained when subjects make direct perceptual comparisons of patterns differing in size, Moyer argued that subjects compare animal names by making an "internal psychophysical judgment" after first converting the names to analog representations that preserve animal size. Smaller size differences between animals presumably are represented as smaller differences between the internal analogs, and the resulting decreased discriminability is reflected in increased reaction times.

Moyer's study is unusually relevant to the question of how information about the perceptual world is represented in long-term memory. Although the experiment was not designed explicitly to test such a theory, the results are generally consistent with a dual coding approach according to which verbal and nonverbal information are represented and processed in functionally independent, though interconnected,

cognitive systems (e.g., Paivio, 1971, 1974). One system, the imagery system is presumably specialized for encoding, storing, organizing, transforming, and retrieving information concerning concrete objects and events. In brief, it represents our knowledge of the world in a form that is highly isomorphic with or analogous to perceptual knowledge (cf. Attneave, 1972; Cooper & Shepard, 1973). The other (verbal) system is specialized for dealing with information involving discrete linguistic units and structures. Independence implies, among other things, that the two systems can be independently accessed by relevant stimuli: the imagery system is activated more directly by perceptual objects or pictures than by linguistic stimuli, and conversely in the case of the verbal system. Interconnectedness simply means that nonverbal information can be transformed into verbal information, or vice versa, under appropriate controlling sets aroused by instructions or contextual cues. Thus Moyer's results would be interpreted in terms of indirect activation of representations in the visual image system through their interconnections with name representations in the verbal system. His results can be taken as evidence of one aspect of the structural and functional properties of the nonverbal representation system, namely, that they somehow represent the relative sizes of animals in long-term memory in a manner that parallels perceptual size differences. Whether these representations are like visual images or at least involve a visual memory system is one of the questions that the present series of experiments attempted to answer.

At least two other classes of interpretations are relevant to Moyer's data. One is the traditional verbal associative approach, according to which the results would be explained in terms of verbal response bias created by the experimental task itself or stemming

from pre-experimental associative habits. The first of these goes as follows. The experimental task involved all pairwise comparisons between seven animal names ranked in size. Thus the largest ranking animal was largest in all comparisons, whereas the smallest ranking animal was never largest in any comparison. Animals in the middle were sometimes the larger member and sometimes the smaller member of a pair. The accumulative response frequencies could bias the results in such a way as to produce the function that Moyer obtained. He was aware of this possibility and ruled it out because he obtained a similar function when he analyzed only the first comparison of each size difference made by the subject. The other associative possibility is that the function resulted from differences in pre-experimental verbal habits. One often says that bears and elephants are big and that ants and fleas are small. Thus, in a comparison such as *bear-flea*, "big" occurs quickly as an associative reaction to "bear" but not to "flea" and choice time should be accordingly fast. The possible effect of this kind of response bias was not ruled out in Moyer's experiment, but it was reduced in an extension to be reported here.

The third class of theories includes all those that postulate abstract entities of some kind as the representational substrate of semantic information. Most network models of semantic memory (e.g., Collins & Quillian, 1972) use only discrete semantic features to describe specific concepts, so they would have difficulty representing values on a continuous dimension, such as size. They might do so by adopting relational features like those proposed by Bierwisch (1970). The question, "Which is larger, X or Y?" presumably would be evaluated relative to the norm for the class of objects, in this case animals. If X is large and Y small relative to the norm for animals, the subject could make the choice on the basis of the relational semantic information. However, this would not account for the reaction time function without some additional assumptions. Something like Moyer's function could be deduced if we assume that each animal name has the feature *larger than the norm* or *smaller than the norm* in a feature list or stack, and that the more salient or relevant the size feature is, the higher it is in the stack. The subject could then work down the stack for each name and compare features, responding according to the kind of size marker he encountered first. One can think of various alternatives of this kind, but they all seem rather strained and they begin to look increasingly like the old-fashioned verbal associative model before they even begin to account for the reaction time function.

A somewhat different, though related, explanation might be in terms of descriptive propositions like those used by Anderson and Bower's (1973) human associative memory (HAM). Thus HAM might represent each relative size difference in a propositional tree consisting of a subject and a predicate, together with the relation "larger than." The propositional network would be extended to capture the idea that *bear* is larger than *wolf* is larger than *duck*, and so on (a possible format is suggested by HAM's representation of relative position information, Anderson & Bower, 1973, p. 408). This seems rather unparsimonious, but nonetheless plausible, when the comparisons are restricted to well-defined categories such as ANIMAL. Even so, the model would do no better than the semantic feature theory in explaining the reaction time function unless an analog process is assumed at some point in the system. That is, size differences must get represented in the propositional structure so that the smaller the size difference for a given comparison, the farther apart the concepts would be in a functional sense. No provision has been made in HAM or in other contemporary models of semantic or long-term memory for coping directly with such phenomena.

Reported here are four experiments that were designed to evaluate the alternative interpretations. In the first experiment, size ratings were obtained for a large number of items so that the range of objects to be compared could be extended to include not only animal-animal pairs but also animals paired with inanimate objects, as well as object-object pairs. The use of a large number of objects from different conceptual categories would make it increasingly unparsimonious to explain positive results in terms of relational features or propositional networks. The range of size differences was also sampled in such a way as to minimize the effect of associative biases favoring "big" or "small" as habitual verbal responses to particular items.

The other three experiments were intended to provide more direct evidence that the memory size comparisons involve visual processes. Experiment II contrasted pictures and words in a size-comparison task, with two expectations based on the dual-coding hypothesis. First, if size comparisons involve a visual memory system, the reaction times should be faster with the pictures than with the words because a picture has more direct access to the visual image system according to the dual coding theory. The second expectation was that a conflict could be created when real-life size relations are incongruent with physical size relations portrayed in the pictures (e.g., the conceptually larger member of a pair being depicted as smaller than its partner). The conflict should be absent or reduced in the case of words, because they must be read and interpreted before the relevant visual system is accessed.

Experiment III required subjects to make judgments of relative distance of the pictured objects, with the expectation that the reaction times for pairs with congruent and incongruent size relations should be the reverse of that obtained in the size comparison

experiments. That is, a conceptually large object would be quickly perceived as farther away when its picture size was small relative to its partner. Conversely, a congruent size relation should make the distance judgment difficult.

The verbal associative hypothesis would predict that the size comparisons would be faster with words than with pictures, contrary to the dual coding theory, and it predicts no conflict effect as a result of size incongruency because, according to such a view, the picture would have to be named before any comparison was made. The verbal hypothesis makes no clear prediction about the relative distance judgments. The abstract representation theories as presently formulated are generally irrelevant to the tasks involving pictures as stimuli.

Experiment IV involved judgments of the relative pronounceability of the names of items presented as pictures or words differing in size, as in Experiment II. Since the comparisons depend on verbal information, it was expected that RTs would now be faster for words than for pictures and that size congruity-incongruity would have no effect, contrary to the prediction for Experiment II.

SIZE NORMS

The experiments involved items drawn from a normative list of object names scaled for size on the basis of subjects' ratings. Since the first experiment was intended as an extension of Moyer's experiment, the norms encompassed approximately the range of sizes included in his list of 21 animal names. The procedure involved generation of a pool of items falling within that size range and ratings of the sizes of those objects.

The Item Pool

A total of 60 animal names were used as cues for the generation of additional items from other conceptual categories. Included among the stimuli were the 21 animal names used by Moyer (1973). The remaining animal names were selected from the Battig and Montague (1969) category norms. The 60 animal names were randomly divided into five sets of 12. The 12 items in each set were used as cues for nine subjects. The animal names in each set appeared on separate sheets of paper, one name per page repeated a number of times down the side of the page. Each of the nine subjects within a group received a different order of presentation of the 12 animal names.

The subjects were instructed to write down as many objects as they could think of that were approximately the same size as each of the given animals. The only restriction in selecting objects was that they were not to use the names of other animals. The subjects were also supplied with a list of seven sample category names, such as "things you would find in the kitchen" and "musical instruments," taken from the Battig and Montague norms. The subjects were encouraged first to think of items matched in size to the animals without reference to the category cues. When they had exhausted their own memory supply of suitable objects, they were to refer to the list of category names. They were also cautioned against dwelling too long upon objects from a single category. The subjects were allowed 2 min to write their responses for each of the animal names.

Following the association task, the subjects were asked to explain in writing the general strategies involved when they selected objects similar in size to the given animals. They were asked both to describe the mental process involved and how the actual decision was made when selecting objects. This task was untimed. Finally, they were given a questionnaire listing specific devices which may have been involved in the task. These devices were "Using verbal keys, such as 'greater than' or 'smaller than' to match the animals to similar sized objects," "General verbal strategies (e.g., letter similarity)," and "Imagery (i.e., mental pictures)." The subjects were instructed to indicate the extent to which they used each device on a 5-point scale, ranging from "used very infrequently or never" to "used with all or almost all of the animals' names."

The subjects in this phase of the study were 45 introductory psychology students, who volunteered as part of a course requirement.

The Size Ratings

The 60 animal names used in the first phase plus 116 of the subjects' associative responses that were reported with the greatest frequency to represent objects matched in size to the named animals were selected to be rated for size. The animal and object names, typed in capitals in two columns, appeared in random order on four pages, 44 words per page. Printed at the top of each page was a 9-point numerical scale, with the extreme ends labeled SMALL and LARGE, respectively. The items and instructions were presented as a mimeographed booklet, with the order of the four pages of items varied for different subjects. The first page contained the instructions and spaces for name, sex, etc.

The instructions asked the subjects to rate the sizes of objects on a 9-point scale, where 1 represented the smallest size and 9 the largest size. They were also told to look over all of the objects first in order to get an impression of the range of sizes before beginning the rating task. The instructions were read aloud by the experimenter and any questions were answered before the ratings began. The subjects were first given a sample of items to rate. This was followed by untimed ratings of the 174 experimental items.

Following the rating task, subjects were instructed to fill out a questionnaire listing specific devices which may have been involved in the task. This was the same questionnaire that was used in the first phase of the study.

The subjects for this phase of the study were 49 introductory psychology students who were fulfilling a course requirement.

Results

Means and standard deviations were calculated for each of the 174 items. The items are presented in Table 1, ranked from smallest to largest in mean rated size.

The small variances for most of the items give an indication of the reliability of the ratings across subjects. An intergroup reliability index was also computed by dividing the 49 subjects randomly into separate groups of 24 and 25 subjects. Mean scores for each item were calculated for each of the subgroups, and the two sets of means were then correlated. The Pearson product moment correlation was .99, confirming the high reliability of the ratings.

The subjects' responses to the questionnaires following both phases of the study provided suggestive evidence concerning the long-term memory processes involved in the tasks. Most of the responses indicated an overwhelming reliance on visual imagery at some point during size comparisons or estimates. Of the subjects who provided the associative data, 67% gave responses suggestive of the use of visual imagery in

Table 1
Means and SDs of the Size Ratings for 176 Items

Item	Mean	SD	Item	Mean	SD	Item	Mean	SD
Salt	1.00	.00	Foot	2.41	.57	Goat	4.57	1.31
Dot	1.02	.02	Book	2.43	.69	Tuba	4.59	1.79
Flea	1.02	.02	Vase	2.47	.62	Hog	4.61	1.14
Ant	1.04	.04	Turtle	2.47	1.11	Sheep	4.83	1.26
Crumb	1.12	.19	Lobster	2.51	.54	Stereo	4.84	2.22
Roach	1.20	.28	Dove	2.51	.74	Wolf	4.86	.82
Snail	1.20	.16	Shoe	2.55	.66	Trunk	4.88	1.13
Marble	1.22	.17	Hat	2.55	.90	Bicycle	4.90	1.55
Peanut	1.22	.17	Belt	2.57	.98	Dishwasher	5.02	1.24
Bee	1.22	.17	Squirrel	2.61	.77	Desk	5.06	1.65
Pea	1.22	.17	Iron	2.63	.76	Door	5.14	2.08
Moth	1.24	.23	Flute	2.67	.91	Stove	5.16	1.36
Nail	1.24	.18	Yardstick	2.73	.91	Table	5.37	1.50
Raisin	1.27	.19	Teapot	2.82	.52	Dresser	5.39	1.30
Eye	1.31	.25	Kettle	2.84	.63	Ostrich	5.45	2.00
Grape	1.31	.33	Football	2.88	.80	Ladder	5.45	2.08
Cherry	1.33	.22	Rope	2.90	1.11	Motorcycle	5.51	1.88
Butterfly	1.35	.23	Toaster	2.96	.57	Alligator	5.51	1.92
Thimble	1.35	.23	Purse	2.98	.71	Bathtub	5.52	2.46
Dime	1.35	.23	Blender	2.98	.63	Leopard	5.59	1.51
Earrings	1.37	.35	Snake	3.02	1.24	Panther	5.63	1.21
Eraser	1.39	.28	Coffeepot	3.02	.59	Man	5.69	1.48
Ring	1.39	.32	Quail	3.02	.92	Octopus	5.82	3.13
Toe	1.41	.28	Head	3.04	.73	Boulder	5.88	2.76
Penny	1.43	.37	Rabbit	3.06	.96	Tiger	5.90	1.77
Button	1.43	.53	Rooster	3.06	.75	Donkey	5.92	.97
Cigar	1.45	.33	Duck	3.08	.77	Mule	5.98	.96
Finger	1.45	.25	Skunk	3.10	.72	Tent	6.08	2.97
Prune	1.49	.49	Arm	3.10	.79	Bed	6.10	1.72
Quarter	1.49	.49	Cat	3.12	.96	Lion	6.16	1.36
Thumb	1.49	.37	Lamp	3.18	.68	Chest	6.20	1.75
Louse	1.51	1.36	Bread	3.22	1.11	Zebra	6.22	1.15
Caterpillar	1.53	.98	Clarinet	3.24	.59	Kangaroo	6.27	1.66
Ribbon	1.76	.55	Pumpkin	3.24	.72	Elk	6.33	2.83
Pencil	1.78	.46	Watermelon	3.27	.77	Shed	6.35	3.57
Spoon	1.84	.38	Footstool	3.35	1.29	Cow	6.47	.94
Doorknob	1.84	.46	Drum	3.40	.91	Refrigerator	6.57	1.39
Watch	1.84	.54	Violin	3.41	.85	Horse	6.65	1.21
Orange	1.86	.45	Pillow	3.41	1.06	Stall	6.71	1.71
Tomato	1.94	.35	Trumpet	3.47	.53	Freezer	6.76	1.74
Crab	1.98	.71	Goose	3.53	.66	Piano	6.88	1.28
Mouse	1.98	1.08	Porcupine	3.53	.86	Moose	7.06	2.22
Frog	1.98	.47	Beaver	3.61	.69	Boat	7.10	2.51
Apple	2.00	.73	Raccoon	3.67	.87	Tractor	7.20	1.55
Wallet	2.02	.55	Monkey	3.73	.93	Buffalo	7.27	1.66
Cup	2.04	.41	Eagle	3.82	.93	Bear	7.37	.76
Ashtray	2.06	.38	Sink	3.82	1.99	Car	7.41	1.83
Cucumber	2.06	.47	French Horn	3.88	.92	Camel	7.43	.90
Banana	2.08	.44	Tire	3.90	.91	Giraffe	7.69	1.03
Cord	2.10	1.11	Penguin	3.90	1.52	Rhinoceros	7.76	1.29
Fist	2.12	.64	Fox	3.98	.80	Garage	7.96	1.84
Knife	2.12	.96	Child	3.98	1.86	Tree	8.06	1.24
Mitten	2.14	.45	Window	4.02	1.33	Truck	8.12	.84
Pocket	2.17	.72	Tricycle	4.06	1.23	Bus	8.35	.84
Saucer	2.18	.48	Dog	4.06	.91	Elephant	8.35	.84
Hairbrush	2.22	.58	TV	4.10	1.23	Whale	8.55	.78
Hand	2.27	.52	Chair	4.49	1.11	Aeroplane	8.73	.36
Rat	2.27	.73	Bush	4.57	2.98	Iceberg	8.80	.24

their answers to the general question. Their answers to the results of the specific questionnaire showed mean ratings on the 5-point scale of 4.67 for the use of imagery, 2.42 for verbal keys, and 1.53 for general verbal strategies. Not only is the mean reported frequency much higher for imagery than for the other devices, but in addition 87% of the subjects showed the difference favoring imagery.

The answers to the strategy questionnaire by the rating group revealed means of 4.10 for imagery, 3.25

Table 2
Examples of Pairs Used in Experiment 1

Approximate Size Ratio	Animal-Animal	Object-Object	Animal-Object
1.17	Camel-Cow	Cup-Watch	Rhinoceros-Refrigerator
1.40	Dove-Porcupine	Tractor-Desk	Tricycle-Mule
1.75	Dog-Bear	Coffeepot-Motorcycle	Rooster-Table
2.33	Crab-Elk	Lamp-Toe	Foot-Panther
3.50	Wolf-Butterfly	Piano-Apple	Nail-Goat
7.00	Snail-Elephant	Eye-Aeroplane	Iceberg-Moth

for verbal keys, and 2.04 for general verbal strategies. Thus, in this phase, too, size judgments were most frequently based on visual imagery, according to the subjective reports. In this case, 49% of the subjects reported greater use of imagery than of the other devices.

More compelling evidence for the use of a visual memory system was provided by the following experiments, each of which involved items drawn from the normative list and presented either as words or pictures.

EXPERIMENT I

The purpose of the first experiment was to replicate and extend Moyer's experiment. The extensions included, first, the use of object-object and animal-object pairs in addition to animal-animal pairs. Second, to avoid any development of verbal response biases as a result of the experimental experience, each subject was exposed to a given item only once during the experiment. Third, to control for the possible effects of pre-experimental verbal habits, a range of pair ratios was sampled from the smallest, middle, and largest thirds of the size range. This means that items from the middle range were never paired with items from either extreme of the size range, and within the range sampled, items were paired equally often with smaller and larger items for different groups of subjects. This procedure, coupled with the fact that comparisons in the object-object and animal-object conditions cut across conceptual categories, makes it unlikely that positive results would have been strongly influenced by any differential pre-experimental verbal habits that systematically favored "large" or "small" as associative responses to the middle-range items in particular.

Finally, ratio differences based on the means of the size ratings, rather than ordinal size differences, were varied in the experiment. This, together with the selection of items from three different size ranges, made it possible to compare the effects of nominally similar ratio differences within small, medium, and large objects.

Method

Lists. The stimuli for each subject consisted of a list of 54 different pairs, including an equal number (18) of animal-animal (AA), object-object (OO), and animal-object (AO) pair types. Each type included three pairs from each of six different larger/smaller size ratios, namely, 1.17, 1.40, 1.75, 2.33, 3.50, and 7.0. These were selected so as to sample the range of size differences used in Moyer's (1973) experiment, with the sampled differences converted into ratios according to the normative ratings in Table 1 of the present study. Thus, the smallest ratio in the present experiment corresponds approximately to the average ordinal size difference between animals ranked 6 and 7 in Moyer's list (7/6 = 1.17), whereas the largest ratio corresponds to the largest difference in Moyer's list (7/1 = 7.0). Examples of each type of pair and ratio are presented in Table 2.

The use of ratios in the present experiment permitted a comparison of RTs for equal ratio differences at three different size ranges from the normative list. Thus, the size ratings for individual items ranged from 1.0 to 3.99 for the smallest size range, 2.0 to 6.99 for the intermediate range, and 3.0 to 8.99 for the largest size range. Pairs were formed within each area so as to sample the four lowest ratios, viz. 1.17, 1.40, 1.75, and 2.33. The number of pairs from each area varied for different subjects, ranging from 13 to 18 for the smallest, from 22 to 27 for the intermediate, and from 16 to 23 for the largest objects. The remaining pairs involved ratios of 3.50 and 7.0, which extended beyond the values that were possible within any one area. Accordingly, the two largest ratio differences were not involved in comparisons of the three size ranges, although they were included in the overall analysis.

Five different lists with the above specifications were constructed using 148 items from Table 1. A given item occurred only once in any one list, but items were repeated across lists, always paired with different items and resulting in different pair ratios. The number of occurrences of a given item over the five lists varied from one to five. Because the item pool included more inanimate objects than animals, most of the sampled animal names were included in all lists whereas a given object was used less frequently across lists. Five additional lists were formed by reversing the left-right position of stimulus pairs.

Procedure. The stimuli, typed in Executive face capitals on a 7.6 x 12.7 cm filing card with about 2.5 cm separating each word, were presented in a Gerbrands two-channel mirror tachistoscope. Reaction times were recorded on a Hunter Model 120A Klockounter. The timer started when the stimulus was illuminated and stopped when the subject pressed either a left- or right-hand key. Instructions specified that the subject's task was to press the key on the same side as the larger of each pair of named objects. It was emphasized that they were to choose the item that was conceptually larger, i.e., larger in real life, and not the longer word. Each trial consisted of a "ready" signal, followed by a pre-exposure field with a black cross centered in the field. The subject understood that he was to fixate the cross. This was followed by the exposure of the stimulus pair. Nine practice pairs were presented prior to the 54 experimental pairs. The experimental pairs were
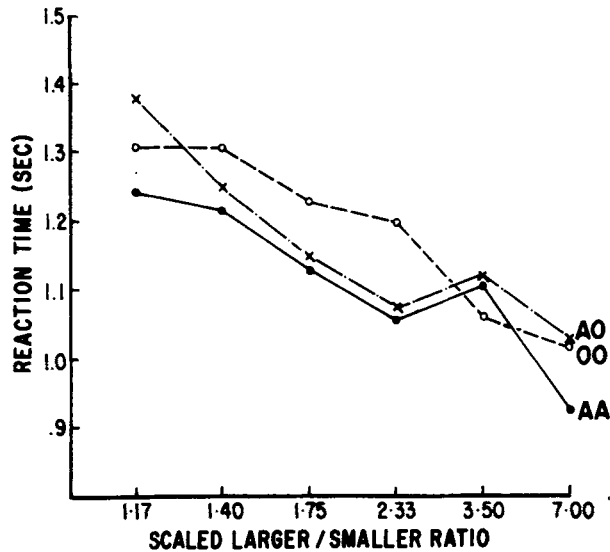
Figure 1. Mean RT for size comparisons of animal-animal (AA), object-object (OO), and animal-object (AO) pairs as a function of scaled ratio size difference.

presented in random order, and the left-right position of the larger member of the pair was also randomized over pairs.

Subjects. Twenty subjects participated in the experiment. Eleven were introductory psychology students fulfilling a course requirement, six were psychology graduate student volunteers, and three were staff members. All were naive to the experimental task. Four subjects were assigned randomly to each of the five basic lists (two to each list and its left-right reversal).

Results and Discussion

Overall effects. Mean correct RTs were determined for each size difference and pair type for each subject. Each score was based on a maximum of three pairs, with correct responses defined according to the normative size differences. The percentages of incorrect choices that were excluded from the analysis were 24, 7.8, 6.4, 5.0, 4.8, and 0 for ratio size differences 1.17, 1.40, 1.75, 2.33, 3.50, and 7.0, respectively, collapsing over pair types and lists. Examination of the mean RTs for errors indicated that their inclusion would not alter the overall pattern of results. The correct RT data were analyzed by a 6 (size ratios) by 3 (AA, OO, and AO pair types) by 2 (left-right order) by 10 (lists) analysis of variance, with repeated measures on the first three factors. The results showed significant main effects for size ratio [$F_{(5,25)}$ = 19.7, $p < .001$] and pair type [$F_{(2,10)}$ = 5.12, $p < .05$]. The only other significant effect was the triple interaction of Size Ratio by List by Pair Type [$F_{(40,50)}$ = 1.99, $p < .01$]. This interaction does not represent a serious qualification of the strong size ratio effect, particularly since the latter is not qualified by any significant interactions with list (F = 1.11) or pair type (F = 0.95). The triple interaction may be more important as a modifier of the pair type effect, but this possibility will not be pursued here because of other evidence that the pair type effect is

not reliable, e.g., it did not appear in the subsequent experiments.

The data for the size (ratio) difference and pair type effects are plotted in Figure 1. Each data point in the figure is based on a total of 44 to 63 observations. It can be seen that RTs generally decreased as the size difference increased, thus replicating the general function obtained by Moyer (1973). In addition, the negative function is uniform across ratios for all three types of pairs except for the rise at a ratio difference of 3.50 for AA and AO pairs. These exceptions presumably reflect the effect of idiosyncratic items, but their influence was too small to be manifested as a significant interaction of size difference and pair type. The data also show that the main effect of pair type was due primarily to the longer RTs for OO comparisons relative to AO and AA. The fact that RTs were on the average fastest for AA suggests that within-category comparisons are faster than between-category comparisons. This conclusion must be qualified, however, since AO comparisons cut across different categories as broadly as OO comparisons, yet the RTs for the former only slightly exceed AA and the overall OA-AA difference is in fact not significant.

Comparisons of similar ratio differences at different size levels. Mean RTs were calculated for larger/smaller ratios of 1.17, 1.40, 1.75, and 2.33, separately for the small, medium, and large objects. (The two largest ratios are not included because they necessarily spanned the three size areas.) The results are plotted in Figure 2. It can be seen that a comparable negative function between RT and size difference holds for the three size ranges. The only notable perturbation in what are otherwise relatively smooth functions is the upturn at the 2.33:1 ratio difference for the smallest objects—presumably due to idiosyncratic items. Also notable is the observation
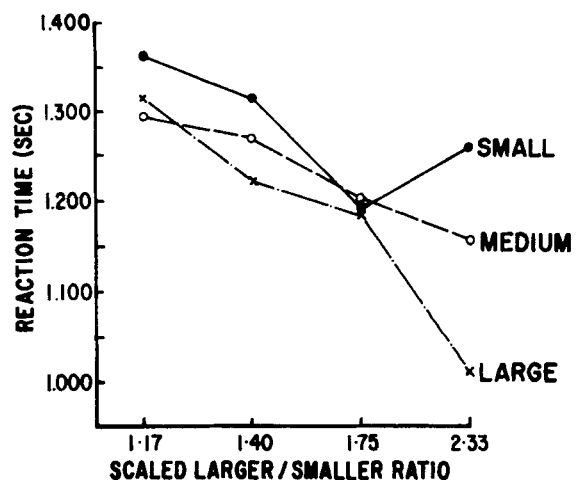

Figure 2. Mean RT for size comparison of small-, medium-, and large-sized animals and objects as a function of scaled ratio size difference.

that the RTs were generally fastest for the largest objects and slowest for the smallest objects. An analysis of variance of these data showed that the main effects of size ratio and size range are both significant [F(3,57) = 10.8 and F(2,38) = 18.7, p < .001], but their interaction is not (F = 1.31). Thus the psychophysical function is statistically equivalent regardless of the absolute size range of the comparison objects. The effect of size range is an interesting problem for further research, but it is not directly pertinent to the main issues in the present study so it will not be discussed further.

The results demonstrate that the psychophysical function obtained by Moyer was generally replicated for size comparisons across as well as within conceptual categories, for equivalent ratio differences within three different size ranges, and with only one presentation of a given item to a particular subject. The functions presumably would be smoother and their precise forms clearer with more subjects and pair samples, but the main purpose of the present experiment was to establish the generality of the findings and this was achieved. The results decrease the plausibility of verbal associative interpretations of the functions, although they probably do not rule them out entirely. They also increase the difficulties for semantic or propositional network theories of semantic memory inasmuch as comparison times and functions were not greatly affected by increasing the range of conceptual categories involved in a given comparison. Note, for example, that a Collins and Quillian type of model clearly implies that comparisons would be slower across different categories than within them. This did not occur consistently. Finally, the subjective reports obtained during the normative study were consistent with an imagery interpretation of size comparisons and estimations. Such an interpretation was tested more directly in the following experiment.

EXPERIMENT II

Experiment II tested the dual coding theory using pictures and words as stimuli in the size comparison task. According to the theory, long-term visual memory representations (images) corresponding to concrete objects are activated more directly by pictures than by names. Thus, if size comparisons primarily involve the visual memory system, the decisions should be generally faster with pictures than with words as stimuli. Secondly, subjects should respond to the pictures themselves as perceptual analogs so that they would tend initially to choose the member of the pair that is depicted as physically larger. Thus, correct RTs should be fast when physical size and memory size are congruent, but relatively slow when the size relations are incongruent because of conflicting response tendencies to relevant
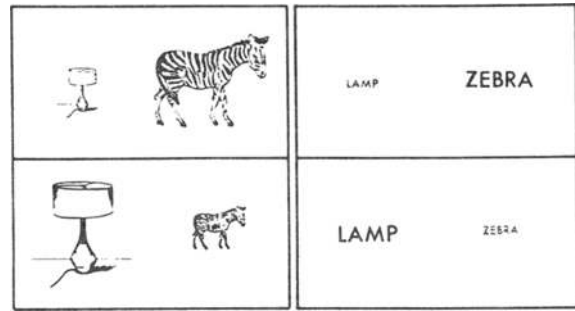

Figure 3. Examples of congruent (top of figure) and incongruent (bottom of figure) physical-size/memory-size relations between picture and word pairs used in Experiment II. Same-size pairs were constructed by pairing the two large and the two small pictures or words.

perceptual and memory information. The conflict should be absent or at least reduced in the case of words differing similarly in printed size because the words must be read and interpreted before they can activate the visual memory representations necessary for the size comparison; word size is presumably irrelevant to the conceptual comparison. A verbal coding theory implies, conversely, that size comparisons would be faster with words than with pictures, and that there should be no conflict in the case of pictures or words as a result of size incongruity because in either case the comparison would be based on verbal associative reactions (e.g., the strength of the response "large") to the object names. Abstract representation theories, such as Anderson and Bower's (1973), generally would not generate clear predictions regarding picture-word differences or the congruity variable without additional assumptions.

Method

Each subject was presented a list of either 48 picture pairs or 48 word pairs. A list included 12 pairs of each of four physical size relations, namely, small-small (SS), large-large (LL), small-large (SL)-congruent, and SL-incongruent, where congruency refers to the relation between picture and memory size relations. A list included six different pairs, two each of AA, OO, and AO types selected from the norms so that the absolute size difference between the members of each pair was approximately 3 scale units. The pairs were: ant-dog, lobster-leopard, mitten-stove, arm-bed, lamp-zebra, and toaster-mule. Each pair was presented twice in each of the four size-relation conditions (SS, LL, SL-congruent, SL-incongruent), so that the left-right position of a pair on the first presentation was reversed on the second presentation (e.g., SL lamp-zebra became LS zebra-lamp).

The physical size relations for the SL conditions were arbitrarily set at somewhat more than 2:1 in depicted area. The stimuli were on 23 x 35.5 cm cards of white board. The large pictured items were approximately 5.75 cm in height and width, and the small items were about 3.18 cm in both dimensions. The large and small words were printed in capitals, 1.27 and .64 cm high, respectively. In both picture and word conditions, the stimuli were separated by approximately 2.5 cm. Figure 3 shows examples of the SL-congruent and SL-incongruent conditions for the pair, zebra-lamp, both in picture and word versions. The SS and LL conditions were constructed from the small and large versions, respectively, of each stimulus.

Table 3
Mean RTs in Seconds for Memory Size Comparisons of Animal (AA), Object (OO), and Animal-Object (AO) Pairs Presented as Pictures or Words so that Drawing or Print Size Differences are Congruent or Incongruent with Real-Life Size Differences or so that the Stimuli are About Equal in Size

| | Pair Type | Different Size | | Same Size | |
		Con-gruent	Incon-gruent	Both Small	Both Large
Pictures	AA	.609	.710	.679	.691
	OO	.590	.666	.610	.590
	AO	.624	.714	.714	.659
Words	AA	.823	.793	.893	.866
	OO	.818	.811	.855	.844
	AO	.768	.797	.817	.815

Four different orderings of the 48 pairs were used in different lists, with the restriction that each sequence of 12 pairs always included three presentations of each of the four size conditions. Moreover, the second half of each list contained the left-right reversals of the pair from the first half.

The general presentation procedure and instructions were as in Experiment I, but with additional emphasis on the necessity of reacting to the conceptually larger member of each pair, regardless of the physical size of the stimuli. The experimental list was preceded by nine practice items.

The subjects were 13 graduate psychology students and 3 faculty members with no previous experience with the task. They were assigned randomly to either the picture or word condition.

Results and Discussion

Mean RTs for correct choices were computed for each subject for each size relationship and pair type (the percentages of incorrect choices excluded from the congruent, incongruent, small-small, and large-large conditions, respectively, were 3.1, 7.3, 4.2, and 1.0 for pictures, and 0, 9.4, 3.1, and 2.1 for words). The overall means for each condition are shown in Table 3, separately for pictures and words. The data were analyzed by a 2 by 4 by 3 analysis of variance involving mode of presentation (pictures vs. words), size relations, and pair type as factors, with repeated measures on the last two.

The two results of special interest were the main effect of mode of presentation and the interaction of mode with size relations. Both effects were significant. The RTs were faster with pictures than with words [F(1,14) = 11.6, p < .01], as predicted from the dual coding hypothesis. The difference favoring pictures held for every pictured pair-size relation and pair type. In fact, mean RTs were faster for pictures than for words for each of the individual pairs under every conditions. i.e., 24 paired comparisons. The Mode by Size interaction [F(3,42) = 4.08, p < .05] reflects contrasting patterns for pictures and words particularly in regard to the congruent and incongruent pairs: for pictures, RTs were on the average fastest for congruent (.608 sec) and slowest for incongruent (.697 sec) pairs, as predicted; for words, mean RTs were essentially equal for congruent and incongruent pairs (.803 and .800 sec, respectively).

A further analysis of variance involving only the congruent and incongruent size conditions showed that the crucial interaction remained significant [F(1,14) = 8.22, p < .02]. This effect occurred with all three types of pairs, as indicated by the fact that the triple interaction of Mode by Size Relations by Pair Type was not significant [F(2,28) = 0.49]. The incongruent-congruent difference was in fact greater for pictures than for words in the case of each of the six pairs (p = .015 by an exact test). Thus the theoretically critical effect was highly reliable over subjects and items.

Other results from the overall analysis showed that the main effect of pair type was not significant [F(2,28) = 2.42, p > .10], but that the Mode by Pair Type interaction was [F(2,28) = 4.59, p < .05]. These results are of interest because they are inconsistent with those obtained in Experiment I. The fastest mean RTs in the present analysis occurred with AO pairs in the case of words (.799 sec) and OO pairs in the case of pictures (.614 sec), wheras RTs were fastest for AA pairs in Experiment I. Thus no reliable effect can be attributed to pair type over the two experiments.

The above analyses were based on data averaged over eight trials for a given item in different size relations and left-right reversals of pairs. To determine whether the critical effect held for unrepeated items, mean RTs were computed for congruent and incongruent pairs involving items presented for the first time. The means for the respective pair types were .739 and .817 for pictures and .928 and .936 for words. The incongruent-congruent difference is greater for pictures than for words (.078 vs. .008). No analysis was performed on these data because they are based on a small number of observations, but the differences are consistent with the results of the overall analysis and with the prediction.

Replication experiment. To check the reliability of the above findings over subjects and items, the experiment was replicated with the six original pairs as well as two new sets of six pairs each. One of the new sets involved a larger-to-smaller scaled size ratio of 2.33. This is similar to the mean ratio of 2.4 for the original six pairs, thereby providing a means of evaluating the generality of the results over comparable sets of items. The other set of six pairs involved a size ratio of 1.4 in order to permit comparisons of results for different size ratios. Each set included two pairs each of AA, OO, and AO types. The 2.33 ratio pairs were kangaroo-squirrel, buffalo-duck, tree-violin, table-hand, car-cat, and lion-hat. The 1.4 ratio pairs were horse-goat, rhinoceros-alligator, shoe-watch, kettle-apple, piano-pig, and fox-coffeepot. The three sets of items were

presented to different groups of introductory psychology students. half the subjects in each group receiving picture pairs and the other half word pairs. Eight subjects were tested in each condition. The procedure was identical to that in the main experiment. with each pair appearing in congruent, incongruent. and same-size formats.

The major results of the main experiment were completely replicated. The RTs were again faster for pictures than for words ($\bar{X}$ = .810 and 1.001 sec, respectively) [F(1,42) = 20.7, p < .001]. Input mode again interacted with pair-size relation [F(3,126) = 2.76, p < .05]. A further analysis involving only the critical congruent and incongruent pairs revealed a significant main effect of congruency [F(1,42) = 20.9, p < .001]. This was qualified by a significant Mode by Size Congruency interaction [F(1,42) = 5.75, p < .05]. which indicates that the congruency effect was greater for pictures (the means for congruent and incongruent pairs were .766 and .846 sec. respectively) than for words (.974 and .999 sec). Neither the picture-word effect nor its interaction with size congruency were qualified by the list variable (Fs ≤ .19).

The generality of the critical interaction was further evaluated by Newman-Keuls tests of the incongruent-congruent difference. separately for pictures and words as well as for each of the three groups receiving the different sets of items. The difference was significant (p < .05) in each case for pictures, but not for words. Thus the theoretically crucial results confirmed those of the main experiment and are consistent with the prediction from the dual coding theory. The modality-specific conflict was somewhat more subtle in the replication than in the original experiment. however. in that the difference of differences (i.e.. the incongruent-congruent difference for pictures minus the incongruent-congruent difference for words) was about 55 msec in the replication as compared to 92 msec in the case of the original experiment. The difference could be due to the use of more sophisticated subjects in the main experiment. The fact that the RTs were generally faster in the original experiment is also consistent with the subject differences.

The only interesting new result from the overall analysis was a significant main effect of list [F(2,42) = 6.17, p < .01]. This reflects the slower mean RT (1.01 sec) for the groups receiving the pairs with the smaller conceptual size difference than for the two groups receiving pairs differing more in conceptual size. The latter two groups had comparable overall mean RTs of .843 and .865 sec. The effect of the conceptual size difference. which occurred for both pictures and words. confirms and extends the general size difference-RT function observed in Experiment I.

In summary. the results confirm both the predictions from the dual coding hypothesis:

comparison times are faster with pictures than with words. as predicted from the assumption that the comparisons involve the imagery system and that representations in that system are more directly activated by pictures than by words; and the conflict created when the perceptual size difference is incongruent with the memory size difference was significant in the case of pictures but not of words. again as predicted from the theory.

EXPERIMENT III

Experiment III was designed to produce a reversal of the RT pattern for the picture conditions of the previous experiments on the basis of the same (hypothetical) visual representations that presumably mediate size comparisons. This was done by asking subjects to indicate which of two pictured objects appeared to be farther away. Since the pictures contained no contextual distance cues other than relative size. the judgments had to be based on the latter together with information concerning the relative sizes of the real objects in long-term memory. Given the known inverse relation between retinal size and distance. the obvious predictions were that the incongruent pictures would result in fastest RTs and congruent pictures. longest RTs. because big things could be quickly judged as relatively farther when they appear smaller. but not when they appear larger than their picture partners. The RTs should be intermediate when the object pairs are pictured as approximately equal in size. The RTs should also be generally longer in the case of distance than size judgments because the former involves a more complex decision process which is dependent on the initial processing of pictures and memory size information.

Method

The materials were the picture lists of the original six pairs used in Experiment II. The pairs were presented to the subjects in exactly the same manner as in Experiment II. Only the instructions differed in the present experiment. The subjects were told that they would be presented pairs of pictures of various objects. that they should think of each as two things appearing together in a scene. and that they were to press the key on the same side as the object that appeared to be farther away. Eight subjects were used. These included one undergraduate. four graduate students. and three staff members.

Results and Discussion

There was no clear objective criterion for deciding which choice was correct in the case of the congruent size conditions. so all responses in that condition were scored as correct (but see Note 1. below). The conceptually larger member of each pair was considered as the correct "farther" choice in all other conditions (examination of error RTs indicated that their inclusion would have had only trivial effects on the pattern of results). The means and SDs of the

Table 4

Mean RTs in Seconds for Relative Distance Judgments of Pictured AA, OO, and AO Pairs Presented so that Depicted Size Differences are Congruent or Incongruent with Real-Life Differences or Pairs are the Same Size

Pair Type	Different Size		Same Size	
	Con-gruent	Incon-gruent	Both Small	Both Large
AA	2.142	1.347	1.648	1.562
OO	2.486	1.452	1.726	1.916
AO	2.261	1.427	1.602	1.977

Note—The corresponding SDs range from .341 to .825 sec.

correct RTs for each size relation and each pair type are shown in Table 4. It can be seen that, as predicted, RTs were longest in the congruent condition and shortest in the incongruent condition for all three types of pairs. The RTs were generally intermediate for same-size pairs. A two-factor analysis of variance confirmed the significance of the size-relations effect [$F(3,2l) = 19.4$, $p < .001$]. No other effect approached significance ($Fs \leq 1.82$). Newman-Keuls comparisons showed that only the difference between the small-small condition and the large-large condition failed to reach significance. All other comparisons were significant. Thus the pattern of significant RT differences was congruent > equal size > incongruent. This is precisely the reverse of the pattern for size comparisons in Experiment II, as can be seen in Figure 4. The graphical representation also shows that the effect over the three size relations is essentially linear, that the magnitude of the effect is much greater for distance than for size judgments, and that (as predicted) the RTs are generally longer for distance than for size comparisons.

The experiment was repeated with a sample of eight introductory course students, with almost identical results. A further replication with 16 subjects involved the 12 additional picture pairs from the replication study in Experiment II. Thus six pairs had a larger-to-smaller ratio of 1.40 and six, a ratio of 2.33.[1] Again, the pattern of results was highly similar to the original experiment, with the mean RTs being 2.475, 1.999, 2.170, and 1.555 sec for congruent, SS, LL, and incongruent pairs, respectively [$F(3,42) = 12.3$, $p < .001$]. In addition, the judgments were generally faster for the pairs with the larger conceptual size difference. This was qualified by an interaction [$F(3,42) = 3.77$, $p < .05$] in which the ratio-size effect was small for the difficult congruent pairs and substantial for the other size relations. The most important point here, however, is that the replications established the reliability of the predicted effects for distance judgments.

EXPERIMENT IV

Experiments II and III involved judgments

presumably based on long-term memory information in the nonverbal image system, to which pictures were expected to have more direct access. Experiment IV attempted to reverse the situation, using comparisons that presumably depend on information in the verbal system. Such information should be accessed more directly with printed words than with pictures as stimuli. Moreover, size discrepancies in the presented pair should have no effect on such decisions because it is irrelevant to the long-term memory information involved in the task. To test this, subjects were presented with pairs judged to differ somewhat in pronounceability and they were asked to press the key under the word which they thought was the easier to pronounce. Half the subjects were tested with picture pairs and the other half with words, with each condition involving equal numbers of the four pair-size relations used in Experiments II and III.

Method

Twelve different pairs, four each of the AA, OO, and AO types were selected so that the members of each pair differed somewhat in ease of pronounceability. The pairs were ostrich-rooster, whale-fox, leopard-rhinoceros, cow-snake, stove-mitten, truck-bed, shoe-watch, table-tricycle, zebra-lamp, squirrel-nail, dove-apple, and hat-lion. Each subject was presented with a list of the SS, LL, congruent and incongruent size relations, as in Experiment II. Each pair was presented twice in each of the four conditions, so that the left-right position of a pair on the first presentation was reversed on the second presentation. Four different orderings of the 96-pair sequence were used with different subjects. The construction and presentation of the stimulus pairs were identical to those of Experiment II.

The subjects were 16 introductory psychology students (8 males) who volunteered for the experiment in partial fulfillment of a course requirement.
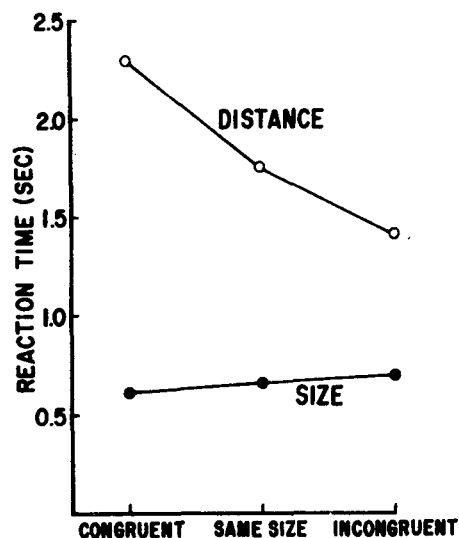

Figure 4. Mean RT for memory size comparisons and relative distance judgments of pictured pairs in which the items are depicted as equal in size or differing so that the pictured size difference is congruent or incongruent with the real-life difference.

Results and Discussion

The mean RTs for congruent, SS, LL, and incongruent size conditions were 1.868, 1.834, 1.818, 1.855 for pictures, and 1.026, 1.034, 1.075, 1.050 for words. As predicted, the comparison time for pronounceability was slower for pictures than for words [$F(1,14) = 4.82$, $p < .05$]. It is apparent that the difference held for each pair-size relation. Again as predicted, the pair-size effect was entirely absent ($F = .179$). Input mode and size also did not interact ($F = 1.25$). The only other effect was one showing that the pair types differed somewhat in comparison time, but this is of no interest here since no attempt was made to equate or control pronounceability differences across pairs. The main point is that the pattern of results contrasted with those from the size and distance judgment experiments in precisely the way predicted from the dual coding hypothesis.

GENERAL DISCUSSION

The results of Experiment 1 confirm and extend Moyer's (1973) findings concerning memory size comparisons. They showed, first, that the negative relation between memory size difference and reaction time for choosing the larger member of a pair applies not only to comparisons within a particular conceptual category (animals), but also to comparisons across diverse categories. Second, the RT appears to be related systematically to size ratios, so that similar psychophysical functions emerge for comparable ratio differences between members of pairs within relatively small, medium, and large objects. Third, the general function occurred without repeated exposures to a given item. Finally, the results were obtained using size norms based on ratings by an independent group rather than by the experimental subjects themselves. Thus the psychophysical function has considerable generality.

The most important new findings, however, emerged from the experiments involving pictures as items, for these provided compelling support for the hypothesis that the comparisons involve *visual* analog representations in long-term memory, rather than verbal or amodal abstract entities of some kind. The supporting facts are that the size comparisons were consistently faster with pictures than with words as stimuli; the pictured size relation facilitated or hindered the choice, depending on whether it was congruent or incongruent with memory size; and this Stroop-type effect reversed when subjects were required to indicate which member of the pair looked farther away than which was conceptually larger. The distance judgments, like the size comparisons, must have been based on remembered size, because the only distance cue in the pictures was the relative physical sizes of the depicted objects. The reversal of the reaction time pattern in the two tasks is important because it indicates that the same visual memory

information can be used for different purposes in the same situation, depending on the demands of the task.

It could be argued that the effects of pictured size congruency-incongruency in the size comparison studies resulted simply from greater difficulty of recognizing the conceptually smaller of a pair when it was pictured as the larger member. This view is supported by subjective impressions in some instances, e.g., in the pair *ant-dog*, the ant looks rather strange when pictured larger than the dog. However, any systematic effect of such recognition difficulty should have resulted in relatively slow RTs for incongruent pairs in the distance and pronounceability comparisons as well. Since this did not occur, recognition difficulty can be ruled out as a significant confounding variable.

The pronounceability comparison results can be viewed as relatively trivial in themselves because they are so predictable, if only from the fact that naming latencies for pictures are slower than reading latencies for words (Fraisse, 1968). They are theoretically important, however, because they contrast sharply with the patterns of results obtained for size comparisons, which presumably could have been similarly affected by the difference in nameability of pictures and words if the size comparisons depended primarily on the verbal system. Thus the reversal of the overall picture-word latency differences in the two cases and the difference in patterns as a function of size congruency provide support for both the image and verbal components of the dual coding theory.

The results of any one of the experiments challenge single-code explanations whether these are expressed in terms of verbal processes, semantic features, or abstract propositional representations of knowledge of the world. The total set of findings appears to be beyond the predictive and explanatory scope of such theories as they are presently formulated. A simple verbal associative theory would erroneously predict faster RTs with words than with pictures as stimuli, and it is difficult to see how the theory could make any predictions concerning the size conflict or distance judgment experiments. Verbal associations would be relevant only if "big" were a dominant reaction to a pictured object class or its name, so that this verbal habit could conflict with the tendency to respond similarly to the physically larger object. The experiments, however, required responses based on the *relative* sizes of pairs of objects in long-term memory, ignoring physical, depicted size in the case of real-life memory size comparisons and simultaneously taking both into account in the case of distance judgments. An absolute response (verbal or other) to either member of the pair seems insufficient to account for the results in this case, as it is in other tasks involving perception of stimulus relations (Reese, 1968).

Abstract entity theories presumably would have to

incorporate some kind of analog process into their assumptions in order to account for the RT data with object names as stimuli. The faster RT with pictures than with names as stimuli would require the further assumption that the common abstract representation was more accessible to pictures than to words as stimuli. Both of these assumptions would have the effect of rendering the abstract theory functionally similar to dual coding theory, which accepts these assumptions at the outset. Finally, abstract-entity theories which assume that the long-term memory representations are amodal (e.g., Anderson & Bower, 1973; Pylyshyn, 1973) would have difficulty explaining the conflict and distance judgment data, inasmuch as the processes involved in these tasks appear to be specifically nonverbal and visual in nature.

The data are, however, easily explained by a theory which assumes that the effects are mediated by internal analog representations that contain relative size information along with other attributes of the objects themselves. These representations apparently are manifested sometimes as consciously experienced visual images, as in the case of the subjects who provided the normative size judgments that eventuated in the scaled values shown in Table 1, but it is conceivable that the analog information can also be functional in a task without any conscious experience of imagery. This suggestion simply acknowledges a necessary logical distinction between a functional cognitive system and its behavioral manifestations. In this case, the hypothetical imagery system is assumed to have various functional properties that are not necessarily dependent on conscious imagery, although the latter is obviously a salient expression of the activity of the underlying system. These and other related issues are discussed elsewhere in more detail by various authors (e.g., Cooper & Shepard, 1973; Hebb, 1968; Paivio, 1971, 1974). The important point here is that consciousness is not viewed as a necessary defining attribute of the imagistic representations presumably involved in size comparisons and other tasks, although it often provides supplementary evidence that such a process is functionally activated.

It is important to note also that the present data provide no information concerning the precise form of the postulated analog representations. In particular, it is not claimed that size is represented in some absolute form, so that the image of an ant literally takes up less brain space than the image of an elephant. Presumably such information is always relational or contextual, just as the objects themselves are always experienced in a situational context involving other objects. However, the memory size comparisons need not depend always on directly experienced relationships between two objects. They may be based instead on comparisons in a context common to both. For example, kitchen utensils and farm animals may never have been experienced perceptually together, but they have been experienced in the context of common objects that are relatively constant in size, such as doors, windows, people, or simply the perceiver's own body images. Such possibilities and their implications are empirically testable.

REFERENCES

ANDERSON, J. R., & BOWER, G. H. Human associative memory. Washington, D.C: Winston, 1973.
ATTNEAVE, F. Representation of physical space. In A. W. Melton and E. Martin (Eds.), Coding processes in human memory. New York: Winston-Wiley, 1972.
BATTIG, W. F., & MONTAGUE, W. E. Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. Journal of Experimental Psychology Monograph, 1969, 80, 1-46.
BIERWISCH, M. Semantics. In J. Lyons (Ed.), New horizons in linguistics. Baltimore: Penguin, 1970.
COLLINS, A. M., & QUILLIAN, M. R. How to make a language user. In E. Tulving and W. Donaldson (Eds.), Organization of memory. New York: Academic Press, 1972.
COOPER, L. A., & SHEPARD, R. N. Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), Visual information processing. New York: Academic Press, 1973.
FRAISSE, P. Motor and verbal reaction times to words and drawings. Psychonomic Science, 1968, 12, 235-236.
HEBB, D. O. Concerning imagery. Psychological Review, 1968, 75, 466-477.
MOYER, R. S. Comparing objects in memory: Evidence suggesting an internal psychophysics. Perception & Psychophysics, 1973, 13, 180-184.
PAIVIO, A. Imagery and verbal processes. New York: Holt, Rinehart, and Winston, 1971.
PAIVIO, A. Language and knowledge of the world. Educational Researcher, 1974, 3, 5-12.
PYLYSHYN, Z. W. What the mind's eye tells the mind's brain: A critique of mental imagery. Psychological Bulletin, 1973, 80, 1-24.
REESE, H. W. The perception of stimulus relations: Discrimination learning and transposition. New York: Academic Press, 1968.

NOTE

1. This experiment provided a basis for a sensitive evaluation of the validity of the scaled ratios using the data for the condition in which pictured and memory size differences are congruent. The choices could be speculatively classified as correct or incorrect by assuming that the normative ratings accurately reflect ratio differences. Since the conceptually smaller member of the pair was pictured about half the size of the larger member, a rough predictive index can be arrived at by dividing the normative size ratios by two. Thus, the index is .70 for the pairs with the 1.40 ratio size difference and 1.17 for the pairs with the 2.33 difference. An index of 1.0 would mean neutrality, i.e., that the two objects were perceptually equidistant. Balues below and above 1.0 would mean that the physically smaller object should appear farther and closer, respectively, relative to its partner. The prediction was supported by the subjects' choices in this experiment. In the case of ratio 1.40 pairs, the object pictured as the smaller of the pair was selected as farther away in 87 out of 95 instances; the opposite was true in the case of ratio 2.33, i.e., the

pictorially *larger* object was selected as the farther one in 73 out of 96 instances. Moreover, the RTs were somewhat faster, on the average, for "correct" choices than for "incorrect" ones. The implications are interesting and worth pursuing in their own right (e.g., they confirm the validity of the size ratings), but they do not modify the relevant conclusions from the present experiment and will not be discussed further here.