

Semantic association effects in a mental comparison task

KEITH J. HOLYOAK

University of Michigan, Ann Arbor, Michigan 48104

SUSAN T. DUMAIS

Indiana University, Bloomington, Indiana 47401

and

ROBERT S. MOYER

Bates College, Lewiston, Maine 04240

Two reaction time experiments explored the effect of semantic associations on the time required to make mental size comparisons. Subjects in both experiments were able to judge relative size more quickly for associated than for nonassociated pairs. This association effect was found for a variety of different types of semantic relationships. Experiment 2 used a large pool of pairs that were scaled for degree of association and subjective size of the individual items. This experiment demonstrated that a high degree of association is more facilitative for pairs with small rather than large size differences. Two different explanations of the results are presented, and it is tentatively suggested that work on mental comparison may be integrated with a broader range of semantic memory research.

People are able to remember a wealth of information about the attributes of objects. For example, most adults probably have a fairly clear idea of the size of a rabbit, the shape of a kitchen chair, the color of coffee, and the speed of a horse. How is such information represented in semantic memory, and how is it located, retrieved, and used to answer questions?

Two independent lines of reaction time (RT) experiments have addressed these issues. In one type of experiment, subjects are timed as they judge which of two symbols is greater or lesser in magnitude along some continuous dimension. For example, the subject might see two animal names, such as "cow" and "dog," and be asked to choose the animal that is normatively larger (Moyer, 1973). The major finding in such experiments is that decision time is faster for pairs that are far apart in magnitude (e.g., cat-moose) than for pairs that are close in magnitude (e.g., cat-fox). This "symbolic distance effect" (Moyer & Bayer, 1976) has been replicated many times for the size dimension (Holyoak, 1977; Jamieson & Petrusic, 1975; Paivio, 1975), as well as for abstract dimensions such as intelligence (Banks & Flora, 1977), pleasantness, and monetary

value (Paivio, 1978). Similar results have also been found for comparisons of the magnitudes of digit pairs (Banks, Fujii, & Kayra-Stuart, 1976; Buckley & Gillman, 1974; Moyer & Landauer, 1967; Sekuler, Rubin, & Armstrong, 1971; Friedman, Note 1). In all of these experiments, RT was closely related to the scaled distance between the items on the judged dimension.

Other experiments concerned with representation in semantic memory involve verification of simple sentences of the form, "An S is a P" (e.g., A canary is a bird) or "S has P" (e.g., A canary has wings) (Collins & Quillian, 1969). Here, too, a type of "distance" variable influences RT. The more closely associated the critical terms (S and P), the faster an affirmative decision can be made. Degree of association (or semantic distance) has been assessed using a variety of indices, such as ratings of semantic relatedness (Rips, Shoben, & Smith, 1973), and measures of associative production frequency (Glass, Holyoak, & O'Dell, 1974; Loftus, 1973; Wilkins, 1971). In a closely related task in which subjects decide whether two nouns are members of the same conceptual category, a high degree of semantic association also facilitates RT (Schaeffer & Wallace, 1969, 1970; Shoben, 1976). In both the sentence and the word paradigms, a high degree of semantic association speeds affirmative decisions and, in many cases, slows negative decisions (see Hutchinson & Lockhead, 1977; Smith, Shoben, & Rips, 1974, for discussions of the relation between semantic distance and RT; see Holyoak & Glass, 1975, for a more detailed discussion of negative decisions).

This research was supported by NSF Grant BNS77-01211 to K. J. Holyoak. Reprint requests may be sent to K. J. Holyoak, University of Michigan, Human Performance Center, 330 Packard Road, Ann Arbor, Michigan 48104. We thank Frank Restle for providing the computer and experimental facilities at Indiana University that were employed in part of Experiment 2. Allan Paivio and an anonymous reviewer provided valuable comments on an earlier draft.

Both the mental comparison and the semantic verification paradigms are attempts to investigate how people make simple decisions based on the information stored in long-term memory. In both types of tasks, some index of psychological distance is the major predictor of RT. But, despite this similarity of purpose and empirical focus, virtually nothing is known about the relationship between the symbolic distance and semantic association effects consistently reported in these two paradigms. The only previously published attempt to manipulate both variables in a single experiment was reported by Paivio (1975, Experiment 1). In this experiment, subjects were presented with pairs of animal, object, and mixed animal and object names. For each pair, subjects were asked to choose the normatively larger item. RTs for object-object pairs were significantly longer than those for animal-object pairs, which did not differ from RTs for animal-animal pairs. Since there was no consistent RT advantage for same-category pairs over different-category pairs, Paivio concluded that the results were problematic for models of semantic memory that assume exemplars of the same category are stored closer together than are exemplars of different categories.

Although the failure to find a clear category effect in this experiment is somewhat discouraging, it is far from conclusive. It is not difficult to think of animal-object pairs (e.g., monkey-banana) in which the items are more closely associated than those in an object-object pair (e.g., shoe-banana). Since Paivio (1975) did not report a normative measure of degree of association for the pairs used in his experiment, the interpretation of his negative finding is left in doubt.

On closer inspection, then, the available evidence does not seem to preclude finding a semantic association effect in the size-comparison paradigm. Can both symbolic distance and association effects be observed in the same experiment? If so, how are the two variables related? Finally, is it possible to develop a somewhat more integrated model of a person's mental operations in this task? We try to answer these questions in the two experiments and discussion that follow.

EXPERIMENT 1

The purpose of the first experiment was to determine whether or not a semantic association effect can be observed in a size-comparison task. Subjects were asked to decide whether sentences such as "Dogs are bigger than cats" were true or false for noun pairs that varied in semantic relatedness. A sentence verification task was employed because it makes it possible to determine the effects of semantic association for both affirmative and negative decisions. Furthermore, the degree of semantic association between pairs of items was assessed directly by having subjects rate the degree to which each pair of items was associated.

Method

Pair selection. The test items were based on 24 pairs of sentences. Three types of associates were used, with eight exemplars of each. The first type were words for which size appears to be the dimension of a minimal semantic contrast (e.g., mountain-hill). For these items, size is the most salient and perhaps the only semantic distinction between the two categories. Associates of the second type, like those of the first, were members of the same immediate superordinate category (e.g., dog-cat); however, for these pairs size is not the only dimension on which the terms differ, nor perhaps the most salient. Associates of the third type were not members of the same immediate superordinate, but rather were syntagmatic associates—words that are often used together, but do not have the same semantic role (e.g., oven-turkey). This classification of type of association is not meant to be definitive; the main motivation for choosing items that fall into these different categories was simply to insure that items exemplifying a wide range of types of associates were included in the study.

Each of the 24 pairs of associates was matched with a less associated pair. The nonassociated pairs were formed by replacing one of the words in the associated pair with another word that lessened the degree of association, and at the same time appeared to slightly increase the subjective normative size difference. In half the cases, the larger member of the associated pair was replaced, and in half the smaller member was replaced. The associates and nonassociates were closely matched on word frequency and length. The 48 word pairs used in Experiment 1 are listed in Table 1.

The experimenters' intuitions concerning degree of association and size difference were then compared with those of normative subjects. Twenty Stanford University undergraduates rated each of the 48 word pairs on a 1-7 scale according to "how often you think of the two objects together," with a rating of 7 indicating maximal frequency. The pairs were listed in random order on a piece of paper. For each of the 24 comparisons (with the exception of one tie), the associated pair was given a higher rating of associative frequency than was the nonassociated pair (means of 5.37 and 2.03, respectively). Table 1 presents the mean association rating for each word pair, as well as mean correct RTs obtained in the subsequent size-comparison experiment.

In addition, the same subjects were given a second questionnaire. For each of the 24 comparisons, the word used to form the nonassociated pair was listed alongside of the associated word it replaced (e.g., for dog-cat vs. goat-cat, "dog" and "goat" were listed). For each comparison, the subjects were asked "to circle the object that you think is generally bigger in the real world." They were told to base their decisions only on "average or typical instances of the two categories," and that they could choose to circle neither word if the two objects seemed equal in size. Their ratings made it possible to determine for each comparison whether the associated or the nonassociated pair had a larger normative size difference. For 22 of the 24 comparisons, an equal or greater number of sentences circled the word indicating the nonassociated pair had a larger normative size difference.¹ Collapsing over all comparisons, 79% of the subjects' responses indicated that the size difference was larger for nonassociates, 16% indicated that the difference was larger for associates, and 5% indicated that the size difference was equal. These questionnaire results thus confirmed that size difference and degree of association were pitted against each other in the present experiment.

Each of these word pairs was presented in four forms: with two orders of the words, and with either the adjective "bigger" or "smaller" (e.g., "Dogs are bigger than cats," "Cats are bigger than dogs," "Dogs are smaller than cats," and "Cats are smaller than dogs"). The total test set therefore comprised 192 sentences, with equal numbers of true and false sentences.

Procedure. Subjects were timed as they decided whether

Table 1
Mean Correct RTs and Association Ratings for Pairs Used in Experiment 1

Type of Association	Associates			Nonassociates		
	Pair	Association Rating	Mean RT*	Pair	Association Rating	Mean RT*
Same Category (Minimal Size Contrast)	mountain-hill	5.60	1638	continent-hill	1.70	1741
	river-creek	5.30	1676	freeway-creek	2.35	1799
	mansion-cottage	3.85	1600	mansion-boathouse	3.10	1787
	ocean-lake	4.25	1581	moon-lake	3.25	1808
	street-path	3.60	1675	river-path	3.60	1773
	bed-cradle	4.50	1674	bed-hen	1.10	1883
	tree-bush	5.75	1569	tree-stove	1.45	1991
	city-village	3.85	1681	city-meadow	2.45	1777
	dog-cat	6.30	1512	goat-cat	1.65	1588
	sun-moon	6.35	1600	sun-mountain	4.55	1737
Same Category	cake-cookie	5.10	1568	sink-cookie	1.30	1747
	cat-mouse	6.50	1675	fox-mouse	2.65	1611
	turkey-chicken	4.70	1574	seal-chicken	1.45	1734
	pipe-cigarette	5.40	1620	pipe-mail	1.60	1804
	table-chair	6.75	1731	car-chair	2.20	1646
	rat-mouse	5.10	1886	rat-thumb	1.15	1892
	oven-turkey	4.95	1780	oven-boot	1.15	1892
	book-pen	5.25	1737	toaster-pen	1.25	1789
	pistol-bullet	6.75	1550	pistol-thimble	1.05	1751
	hammer-nail	6.75	1644	fiddle-nail	1.25	1699
Syntagmatic	chicken-egg	6.35	1611	rabbit-egg	2.35	1653
	robin-worm	5.55	1690	squirrel-worm	1.50	1665
	postcard-stamp	6.15	1776	postcard-dime	3.20	1769
	pot-spoon	4.15	1668	pot-pen	1.35	1869

*In milliseconds.

tachistoscopically presented sentences of the form "As are bigger/smaller than Bs" were true or false. They were told to base their decisions only on typical exemplars of the categories. The subject pressed a center button to initiate a trial; the sentence appeared in a viewer after a 1-sec delay and was displayed until the subject pressed one of two decision buttons to end the trial. The subject was then informed if he had made an error. Assignment of dominant hand to the "true" response key was counterbalanced across subjects. All the subjects saw all 192 test sentences, which were presented in a different random order for each subject. Twenty practice trials preceded presentation of the test sentences.

Fourteen Stanford University undergraduates served as subjects in order to satisfy a course requirement.

Results and Discussion

Since neither the main effect nor any interactions involving type of association (viz., minimal size contrast, same category, syntagmatic) approached significance, the mean RTs reported in Table 2 collapse across this variable. (There was a trend toward a larger association effect for minimal size-contrast pairs, mainly due to the relatively slow RTs of the nonassociated pairs matched with this condition.) As the mean correct RTs given in Table 2 indicate, the size comparisons involving associated categories (e.g., Dogs are bigger than cats) were indeed made more quickly than comparisons between nonassociated categories (e.g., Goats are bigger than cats).² This 112-msec difference was highly significant using the quasi-F ratio statistic (Clark, 1973), which treats both subjects and items as random effects [$F(1,33) = 16.2, p < .001$]. This facilitating effect of

association occurred for both comparative adjectives, and for both true and false sentences. It was obtained in spite of the fact that the subjective size difference was slightly larger for nonassociates. No interaction involving degree of association approached significance.

In addition to the association effect, people responded more quickly when the comparative adjective used was "bigger" rather than "smaller" [$F(1,31) = 38.8, p < .001$] and when the sentence was true rather than false [$F(1,31) = 9.58, p < .01$]. The advantage of true sentences tended to be reduced when the adjective was "smaller," although this interaction fell short of significance [$F(1,24) = 3.67, .05 < p < .10$]. The overall error rate in the experiment was 7%, and error rates were identical for associates and nonassociates.

In summary, Experiment 1 demonstrated that a large and highly reliable semantic association effect can be obtained in a size-comparison paradigm. Increasing degree of semantic association facilitated RTs for both affirmative and negative decisions.

Table 2
Mean RTs for Associated and Nonassociated Pairs in Experiment 1

Adjective		Associates	Nonassociates
Bigger	True	1486	1580
	False	1627	1731
Smaller	True	1729	1869
	False	1777	1887

EXPERIMENT 2

Although Experiment 1 demonstrated an associative facilitation effect in a size-comparison task, the nature of the relationship between semantic association and symbolic distance (size difference) could not be examined, since size difference was not systematically manipulated. In Experiment 2, association and size difference were varied factorially so that the nature of the relationship between the two variables could be examined. Since the associative facilitation effect characterized both affirmative and negative decisions in Experiment 1, a simple pair comparison task (which has been used in most previous mental comparison studies) was employed in the second experiment. The first step in selecting word pairs for this experiment was to collect systematic size norms for individual words, and association norms for word pairs. These norms were then used to generate a large set of associated and nonassociated pairs, each scaled for the size difference between the two words.

Method

Size norms. A pool of possible pairs of associates with clear size differences was first generated by the experimenters. The resulting 158 words (names of common objects and animals) were then scaled for subjective size. Forty-seven Indiana University undergraduates served as subjects in this phase of the experiment in partial fulfillment of a course requirement. They were tested in groups of 10 or fewer. Each subject was given a deck of computer cards, with one word printed on each card. The order of the words in the deck was randomized individually for each subject. In the first step of the scaling procedure, subjects were asked to sort the cards into nine different groups according to the size of the objects named on the cards, with the smallest items in the first group and the largest items in the ninth group. No constraints were placed on the number of cards that could be placed in each group. Subjects were told to base their groupings on the average or typical size of each item. When subjects had completed the sorting task, they were then asked to rank order the items within each group from smallest to largest, guessing if necessary. The entire task required from 35 to 50 min for completion.

Two methods were used to convert the resulting data into a set of size norms for the 158 items. The first method involved using the nine groups as a category rating scale and calculating the size rating that a subject gave each item, using a formula that also took account of the rank given to an item within its group:

$$\text{Size Rating} = \text{Group} + \frac{\text{Rank Within Group} - 1}{\text{Total Number of Items in Group}}$$

The resulting size ratings were then averaged over all subjects to yield a set of norms. These mean ratings and corresponding standard deviations are presented for all items in Table 3.

The second method of deriving a set of norms, successive interval scaling (Bock & Jones, 1968), avoids the need to assume that subjects treated the nine groups as a scale with equal intervals. This method uses the frequencies with which each item was placed in each of the nine groups (ignoring within-group rankings) and provides a simultaneous normalization of the distribution of responses to each item. Parameters for the boundaries between response categories and for item means on the resulting scale were estimated using a least-squares technique.

Table 3
Size Scales for 158 Items

Item	Mean Rating	SD	Successive Interval Value
flea	1.03	.12	.00
ant	1.09	.17	.28
fly	1.26	.30	1.08
pea	1.27	.25	.80
pearl	1.35	.34	.95
tack	1.37	.31	1.08
needle	1.45	.50	.60
pebble	1.46	.43	1.18
bee	1.54	.40	1.56
earring	1.55	.43	1.29
spider	1.64	.48	1.65
button	1.66	.45	1.59
match	1.67	.44	1.65
stamp	1.71	.39	1.62
penny	1.74	.45	1.65
ring	1.77	.43	1.73
nut	1.85	.52	2.09
acorn	1.87	.47	2.01
nail	1.94	.51	2.01
minnow	1.94	.50	2.05
petal	2.00	.73	2.25
eye	2.02	.51	2.22
quarter	2.06	.56	2.32
worm	2.08	.57	2.29
cork	2.09	.57	2.37
key	2.12	.57	2.32
bolt	2.12	.64	2.44
pod	2.16	.55	2.64
toe	2.32	.60	2.77
cigarette	2.32	.63	2.72
icecube	2.32	.52	2.64
thumb	2.44	.65	2.85
ear	2.47	.60	2.88
moustache	2.49	.73	2.97
finger	2.50	.68	3.03
oyster	2.55	.74	3.05
egg	2.55	.63	2.97
watch	2.69	.75	3.20
necklace	2.69	1.00	3.13
leaf	2.72	.75	3.26
pickle	2.73	.71	3.29
mouse	2.77	.71	3.24
cigar	2.78	.66	3.45
spoon	2.85	.70	3.36
buckle	2.86	.84	3.35
dollar	2.89	.79	3.46
flower	2.93	.79	3.55
knife	3.12	.85	3.80
frog	3.14	.84	3.91
carrot	3.14	.86	3.91
letter	3.18	1.00	3.91
apple	3.18	.70	3.90
fist	3.21	.69	3.92
wallet	3.23	.72	3.87
ashtray	3.28	.77	4.01
orange	3.32	.83	4.01
hand	3.40	.82	4.09
mitten	3.41	.78	4.12
hairbrush	3.41	.78	4.14
glove	3.42	.78	4.12
sock	3.46	.77	4.12
ladle	3.48	.90	4.17
grapefruit	3.48	.78	4.11
rock	3.49	1.60	3.89
beard	3.49	.82	4.14

Table 3 Continued

Item	Mean Rating	SD	Successive Interval Value
page	3.57	.92	4.36
bone	3.62	1.15	4.35
rat	3.63	.88	4.44
glass	3.63	1.16	4.32
wrench	3.74	.87	4.35
foot	3.75	.80	4.43
bottle	3.82	.77	4.49
dish	3.82	.87	4.50
shoe	3.86	.82	4.52
bowl	3.87	.86	4.56
belt	3.93	1.03	4.61
hammer	3.97	.89	4.63
pie	3.98	.82	4.64
book	4.01	.80	4.71
hat	4.05	.86	4.70
teapot	4.18	.87	4.82
clock	4.18	1.07	4.86
squirrel	4.24	.82	4.99
purse	4.30	.91	4.99
head	4.41	.89	5.16
snake	4.43	1.26	5.15
toaster	4.52	.86	5.22
rabbit	4.55	.85	5.39
cat	4.67	.88	5.52
vest	4.69	.93	5.41
arm	4.72	1.01	5.57
chicken	4.75	.97	5.54
duck	4.78	.90	5.59
rooster	4.86	1.02	5.60
saw	4.90	.91	5.59
trumpet	4.91	.97	5.64
baby	4.95	.90	5.72
beaver	5.00	.96	5.74
shirt	5.03	1.02	5.79
goose	5.09	.94	5.80
sword	5.09	1.09	5.80
mailbox	5.12	1.05	5.79
pillow	5.13	.87	5.80
puddle	5.20	1.16	5.90
jacket	5.27	.94	6.00
fox	5.29	1.04	6.06
turkey	5.32	1.04	6.01
monkey	5.38	.90	6.11
dog	5.46	.96	6.14
penguin	5.48	.99	6.15
shelf	5.58	1.11	6.19
tire	5.62	.95	6.25
sink	5.62	1.01	6.28
coat	5.63	1.04	6.38
lamb	5.64	1.04	6.26
saddle	5.81	.91	6.49
window	5.86	1.04	6.46
cradle	5.86	.92	6.58
pig	5.88	.97	6.51
dwarf	5.98	1.07	6.68
tuba	6.04	1.04	6.66
chair	6.12	.92	6.81
barrel	5.22	1.09	6.89
oven	6.37	.97	7.11
jockey	6.54	.92	7.32
stove	6.55	.97	7.25
bath tub	6.74	.96	7.34
mother	6.75	.92	7.60
table	6.84	.96	7.45
pilot	6.87	.92	7.63

Table 3 Continued

Item	Mean Rating	SD	Successive Interval Value
man	6.89	.93	7.60
door	6.94	.90	7.56
lion	7.02	1.07	7.73
donkey	7.16	.90	7.82
bed	7.22	.93	7.87
cow	7.33	.90	8.08
gorilla	7.43	1.00	8.16
horse	7.48	.87	8.23
car	7.77	.82	8.81
haystack	7.85	.85	8.70
boat	7.96	.75	8.97
tree	8.15	.77	9.25
bus	8.16	.75	9.31
garage	8.18	.62	9.24
truck	8.19	.81	9.33
oak	8.20	1.14	9.33
giant	8.20	1.03	9.56
house	8.48	.61	9.57
pond	8.54	.52	9.62
hill	8.66	.65	9.90
whale	8.69	.76	9.94
airplane	8.82	.53	10.23
ship	8.98	.51	10.27
iceberg	9.23	.41	11.04
lake	9.27	.40	10.94
forest	9.41	.44	11.35
mountain	9.50	.32	11.69
ocean	9.81	.16	12.64

The item means obtained in this fashion are also presented in Table 3. The correlation between the norms derived by the two techniques was $r = .99$. The slight nonlinearity in the relation between the two scales is entirely due to the fact that the relative size of the spacings between the smallest and the largest items tended to be slightly larger in the successive interval norms.

Pair selection. Using the words that had been rated for size, 80 pairs of words judged by the experimenters to be highly associated were selected. These pairs, all of which are listed in Table 4, exemplified a wide range of types of association. Eighty matching pairs of nonassociates were then created by replacing one of the words in each associated pair (equally often the larger or the smaller item) by a less associated word (drawn from the same item pool) that slightly increased the size difference. Sixteen items were used twice in an associated pair; the remainder were used only once. Only one word was used twice as a replacement to form nonassociated pairs. Each associate word was matched as closely as possible with its nonassociate replacement for syllable length, word frequency (based on Kučera & Francis, 1967), and standard deviation of mean size ratings.

Association norms. To confirm and quantify the experimenters' intuitions about the degree of association, a set of normative data was collected. Thirty-eight Indiana University students, all of whom had served the previous day in the RT phase of the experiment (see below), provided association ratings. Each of the 160 word pairs that had been constructed (80 associates and 80 nonassociates) was typed on a separate computer card. Each subject received a randomized deck containing all of the pairs. For half of the subjects, the larger item was printed on the left, and for half the subjects, the larger item was on the right. Subjects were told to sort the cards into nine groups on the basis of "the degree to which you feel the two objects are associated," using the first group for the least

Table 4
Association Ratings, Size Differences, and Mean Correct RTs (in Milliseconds) for 160 Pairs Used in Experiment 2

Associates				Nonassociates			
Pair	Association Rating	Size Difference	Mean RT	Pair	Association Rating	Size Difference	Mean RT
flea-dog	7.68	6.14	1243	flea-lamb	3.32	6.26	1142
duck-geese	6.97	.21	1394	cat-geese	3.50	.28	1443
rabbit-fox	5.50	.67	1378	rabbit-monkey	3.61	.72	1471
rooster-pig	4.00	.91	1380	toaster-pig	1.55	1.29	1356
minnow-whale	4.95	7.09	1175	minnow-airplane	1.42	8.18	1237
mouse-cat	7.68	2.28	1409	pickle-cat	1.26	2.23	1314
carrot-donkey	3.32	3.91	1351	carrot-cow	1.74	4.17	1164
nut-squirrel	7.84	2.90	1243	penny-squirrel	1.42	3.34	1311
saddle-horse	7.82	1.74	1224	saddle-car	1.97	2.32	1237
pearl-oyster	8.26	2.10	1355	pea-oyster	2.18	2.25	1564
egg-chicken	8.05	2.57	1107	egg-baby	4.40	2.75	1276
nail-hammer	8.37	2.62	1208	acorn-hammer	1.82	2.62	1263
button-shirt	7.74	4.20	1134	button-geese	1.29	4.21	1182
match-cigar	7.24	1.80	1465	button-cigar	1.53	1.86	1471
book-shelf	7.84	1.48	1299	foot-shelf	1.63	1.76	1320
teapot-stove	6.97	2.43	1249	teapot-bathtub	2.11	2.52	1319
glove-hat	6.95	.58	1496	hand-hat	3.47	.61	1787
trumpet-tuba	6.95	1.02	1295	trumpet-barrel	1.84	1.23	1385
stamp-letter	8.61	2.29	1284	match-letter	2.03	2.26	1394
chair-table	8.53	.64	1291	chair-door	3.21	.75	1375
car-truck	7.68	.52	1125	horse-truck	2.71	1.10	1285
needle-haystack	7.03	8.10	1407	needle-boat	1.34	8.37	1152
sock-shoe	8.16	.40	1380	hairbrush-shoe	2.71	.38	1506
ring-watch				ring-leaf	1.29	1.53	1442
earring-necklace	7.58	1.84	1449	bee-necklace	1.58	1.57	1396
garage-house	7.29	.33	1418	garage-ship	1.90	1.03	1367
puddle-pond	6.42	3.72	1333	pillow-pond	1.79	3.82	1405
toe-foot	7.74	1.66	1254	toe-bottle	1.29	1.72	1387
thumb-hand	7.42	1.24	1336	toe-hand	5.53	1.32	1359
hairbrush-head	7.32	1.02	1394	hairbrush-arm	3.53	1.43	1359
foot-shoe	7.90	.09	1475	wrench-shoe	1.50	.17	1570
leaf-tree	8.05	5.99	1164	leaf-bus	1.26	6.05	1275
pilot-airplane	8.08	2.60	1235	table-airplane	1.50	2.78	1206
apple-pie	7.84	.74	1490	apple-book	2.40	.81	1351
pickle-barrel	6.68	3.60	1269	watch-barrel	1.34	3.69	1347
acorn-oak	7.08	7.32	1343	acorn-house	2.18	7.56	1245
baby-mother	7.55	1.88	1166	saw-mother	1.26	2.01	1357
buckle-belt	7.87	1.26	1439	buckle-hammer	1.97	1.28	1355
page-book	8.08	.35	1317	sock-book	1.45	.59	1424
vest-jacket	7.37	.59	1460	vest-dog	1.63	.73	1420
sink-bathtub	6.84	1.06	1279	tire-bathtub	1.32	1.09	1406
icecube-glass	6.29	1.68	1462	icecube-shoe	1.32	1.88	1415
spoon-bowl	7.21	1.20	1213	cigar-bowl	2.37	1.11	1390
cork-bottle	7.05	2.12	1289	cork-dish	1.84	2.13	1309
cigarette-ashtray	7.68	1.29	1472	pod-ashtray	1.40	1.37	1562
pea-pod	7.34	3.21	1452	pea-thumb	1.82	2.05	1275
boat-ship	7.97	1.30	1273	haystack-ship	1.34	1.57	1353
knife-sword	7.47	2.00	1294	knife-pillow	1.61	2.00	1392
monkey-gorilla	7.24	2.05	1207	turkey-gorilla	3.13	2.15	1268
pig-cow	5.26	1.57	1276	pig-gorilla	3.11	1.65	1155
chicken-turkey	6.29	.47	1312	vest-turkey	1.76	.60	1393
bolt-wrench	6.82	1.91	1382	bolt-bowl	1.82	2.12	1590
penny-quarter	6.79	.67	1234	stamp-quarter	3.61	.70	1509
pebble-rock	7.16	2.71	1311	needle-rock	1.58	3.29	1231
lake-ocean	7.32	1.70	1290	hill-ocean	3.24	2.74	1240
hill-mountain	6.97	1.79	1258	pond-mountain	3.71	2.07	1247
orange-grapefruit	6.13	.10	1252	orange-beard	1.76	.13	1625
bee-flower	7.29	1.99	1274	bee-knife	1.32	2.24	1340
dwarf-giant	6.37	2.88	1110	dwarf-whale	2.55	3.26	1361
worm-snake	6.18	2.86	1438	quarter-snake	1.50	2.83	1434
ant-spider	5.68	1.37	1373	ant-ring	1.53	1.45	1376
cradle-bed	7.29	1.29	1288	saddle-bed	2.18	1.38	1403
moustache-beard	7.79	1.17	1363	moustache-glass	1.79	1.35	1513

Table 4 Continued

Associates				Nonassociates			
Pair	Association Rating	Size Difference	Mean RT	Pair	Association Rating	Size Difference	Mean RT
spoon-dish	7.13	1.14	1182	spoon-belt	1.76	1.25	1429
eye-ear	7.53	.66	1463	minnow-ear	1.47	.83	1467
fist-arm	6.90	1.65	1281	fist-chicken	1.45	1.62	1425
penguin-iceberg	6.13	4.89	1280	fox-iceberg	1.24	4.98	1172
lamb-lion	5.58	1.47	1317	lamb-bed	1.90	1.61	1311
ant-fly	5.16	.80	1458	ant-pearl	1.66	.67	1576
carrot-rabbit	7.32	1.48	1330	flower-rabbit	2.55	1.84	1313
toaster-oven	6.97	1.89	1299	toaster-jockey	1.34	2.11	1361
key-door	8.08	5.24	1180	cork-door	1.71	5.19	1328
wallet-purse	7.66	1.12	1371	wallet-head	1.66	1.29	1292
tack-nail	7.34	.93	1466	fly-nail	1.68	.93	1505
pillow-bed	8.26	2.07	1228	mailbox-bed	1.58	2.08	1468
glove-coat	6.92	2.26	1354	glove-window	1.42	2.34	1300
penny-dollar	6.92	1.81	1151	penny-frog	1.61	2.26	1173
watch-clock	7.74	1.66	1300	egg-clock	2.11	1.89	1311
cat-dog	7.92	.62	1143	cat-penguin	3.00	.63	1319
jockey-horse	7.92	.91	1299	oven-horse	1.50	1.12	1394

associated pairs. Unlike the size-sorting task, subjects were not asked to rank order the items within groups. As would be expected given that the experimenters had selected pairs according to a dichotomy, subjects placed most items in the extreme groups and fewest in the central groups.

Association norms were constructed for all pairs by averaging subjects' ratings and also by successive interval scaling. Since the correlation between the two versions of the norms was virtually perfect ($r = .995$), only the mean ratings were subsequently used. For each of the 80 comparisons, the associated pair received a higher rating than the corresponding nonassociated pair (means of 7.08 and 2.03, respectively). Table 4 presents all the pairs used in the experiment, together with their association ratings, size difference (based on the successive interval norms), and mean correct RT in the size-comparison experiment reported below.

Size-comparison experiment. In the RT phase of the experiment, subjects were presented with pairs of words and timed as they judged which word named the larger object. The total item set comprised 320 pairs (80 associates and 80 nonassociates, each presented once with the larger item appearing on the left and once with the larger item on the right). The experiment began at Indiana University, with the pairs presented on 12-in. TV screens controlled by an IBM 1800 computer. Subjects were tested in groups of up to eight, each in a semi-isolated booth. At the beginning of the experiment, the subjects read the instructions on the TV screens. If there were no questions, they went on to 20 practice trials. These consisted of 10 associated and 10 nonassociated pairs, constructed using words that did not occur in the test set.

The 320 test pairs were then presented. These were divided into eight blocks of 40. The division was random with the constraint that equal numbers of associates and nonassociates appeared in each group, with the left or right item being larger for equal numbers of pairs of each type. Also, all 160 unique pairs (collapsing over left-right order) were presented in the first four blocks, and then repeated in their reversed order in the second four blocks. Across different subjects the eight blocks were presented in four different orders. Subjects were given error feedback and a rest break after each block of 40 trials.

On each trial the two words appeared side by side and centered on the TV screen. The subject then pressed one of the two response keys, with the correct response on the same side as the word naming the larger object. The next pair of words was presented approximately 1.5 sec after the subject's response. The entire session lasted approximately 45 min.

Thirty-eight Indiana University undergraduates were first tested in the experiment. At this point, the computer was halted by lightning, erasing the data for half of the subjects. This setback was then compounded when the computer was sold. Undaunted by these technical difficulties, we then moved the experiment to the University of Michigan, where it so happened that another IBM 1800 system was available. After reprogramming, 21 additional subjects were tested in essentially the same manner as described above. The results reported below are therefore based on data for a total of 40 subjects from the two universities.

Results and Discussion

The results were first examined separately for the Indiana and Michigan subjects. The Michigan subjects tended to respond more quickly overall; however, the basic pattern of results was identical for the two groups. Accordingly, the results reported here are based on the data averaged over all 40 subjects. The results also collapse over the four counterbalancing conditions and left-right pair order. Separate analyses of variance were performed treating subjects and items as random effects, and minimum quasi-F ratios (min F') (Clark, 1973) were calculated. All analysis are based on RTs for correct decisions only. RTs longer than 3,500 msec (.8% of the responses) were truncated to that value. Due to a typing error, one pair presented to subjects contained an incorrect word; accordingly, the results are based on a maximum of 79 comparisons between associated and nonassociated pairs.

Mean RTs for all pairs are presented in Table 4. Overall, subjects chose the larger item an average of 50 msec more quickly for associated than for nonassociated pairs (means of 1,312 and 1,362 msec, respectively) [$\min F'(1,115) = 10.8$, $p < .001$]. This difference replicates the facilitating effect of association found in Experiment 1. However, the crucial concern in Experiment 2 was the question of whether and how the variables of degree of association and size difference would interact.

The relationship between the two variables was assessed in two ways. First, the correlations between mean RT and size difference were calculated separately for the 79 associates and 79 nonassociates.³ For both types of pairs, RTs tended to decrease with increasing size difference. However, the correlation was considerably larger for the nonassociates ($r = -.58, p < .01$) than for the associates ($r = -.26, p < .05$). The difference between the two correlations was significant ($p < .02$). Size difference thus has a greater effect for nonassociates than for associates.

The relationship between degree of association and size difference was also assessed by analysis of variance. To perform this analysis, the pairs were divided into three levels of size difference—small, medium, and large—on the basis of the successive interval size norms (see Table 5). Of the 79 possible comparisons, 3 were dropped to eliminate any overlap between the association ratings for associates vs. nonassociates. In addition, all comparisons in which the matched associated and nonassociated pairs fell into different size categories (according to the cutoffs given in Table 5) were also excluded from the analysis. The remaining 63 comparisons were distributed in approximately equal numbers over the three size groups. As Table 5 indicates, mean association ratings were fairly constant across levels of size difference.

Figure 1 presents the mean RTs for associates and nonassociates across the three levels of size difference. Overall, decisions were made more quickly for associated than for nonassociated pairs [$\min F'(1,89) = 11.0, p < .001$]. In addition, decision time decreased with increasing size difference. While this monotonic trend was significant for associates [$\min F'(1,99) = 5.76, p < .025$] as well as for nonassociates [$\min F'(1,97) = 23.8, p < .001$], the effect of size difference was significantly reduced for the associates [$\min F'(1,88) = 5.24, p < .025$]. For pairs with small size differences, associated pairs were evaluated 108 msec more quickly than nonassociated pairs. This difference is virtually identical in magnitude to that obtained in Experiment 1 (where all size differences were relatively small). In contrast, the advantage of the associates was reduced to a nonsignificant 19 msec for the pairs with the large size differences. In fact, for the seven comparisons between pairs with the largest size differences (range = 4.89–8.37), RT averaged 35 msec less for the nonassociates than for the associates. The analysis of variance results

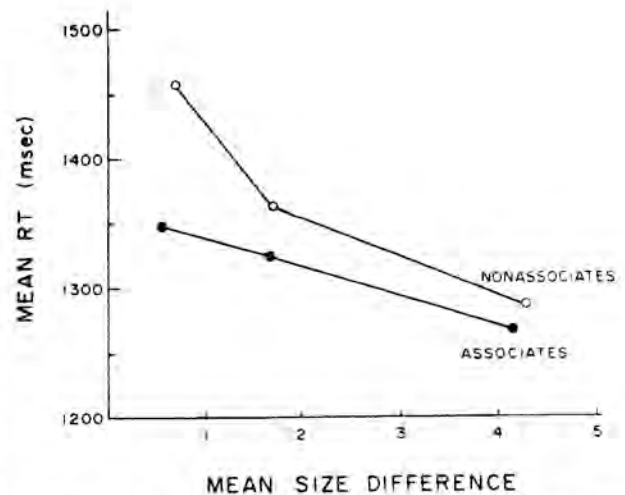


Figure 1. Mean correct RT as a function of size difference for associated and nonassociated pairs.

are thus entirely consistent with the correlation analysis reported above.

The error rate in the experiment was 4.3% for both associates and nonassociates. Over all 158 pairs, the correlation between mean RT and error rate was $r = .54$ ($p < .01$).

Several post hoc analyses were also performed on the RT data. One analysis probed for possible differences in the degree of facilitation produced by various types of semantic associations. Most of the items in the associated pairs were related in one of four ways: as members of the same superordinate category (e.g., cat-dog), by the "part of" relation (e.g., toe-foot), the "on" relation (e.g., teapot-stove), or the "in" relation (e.g., pearl-oyster). All of these association types speeded response times by approximately the same amount (a range of just 20 msec in the size of the association effect over the four types). As in Experiment 1, a wide variety of different semantic relationships produced similar decreases in comparison time. It remains possible, of course, that more sensitive tests may yet reveal differences in performance attributable to type of association.

The effect of the item classification used by Paivio (1975) was also examined. All of the pairs used in the present experiment were classified as animal-animal, object-object, or animal-object. An initial examination did not reveal any clear RT differences among the three

Table 5
Division of Pairs Into Three Levels of Size Difference (Small, Medium, and Large)

	Small Range: .00-1.20		Medium Range: 1.21-2.20		Large Range: 2.21-8.50	
	Associates shoe-foot	Nonassociates shoe-wrench	Associates clock-watch	Nonassociates clock-egg	Associates door-key	Nonassociates door-cork
Examples:						
Mean Size Difference	.57	.69	1.67	1.79	4.15	4.29
Mean Association Value	7.21	2.29	7.14	1.88	7.20	1.70

Note— $N = 20, 23$, and 20 for the small, medium, and large size differences, respectively.

types of pairs, either for associates or nonassociates (a maximum difference of only 30 msec). A multiple-regression analysis was then performed, using three variables to predict mean RTs for all pairs: size difference, association rating, and a dichotomous variable differentiating same-category (animal-animal and object-object) and different-category (animal-object) pairs. In a simultaneous regression analysis, the same-category vs. different-category variable achieved a partial correlation with RT of only $-.06$ ($p = .46$). In contrast, both the size difference and association variables remained significant beyond the .001 level even after the variance attributable to same vs. different category was removed. The present results are thus entirely consistent with Paivio's (1975) failure to find any consistent RT differences between same-category and different-category pairs.

GENERAL DISCUSSION

The two RT experiments reported here demonstrate that people can judge relative size more quickly for associated than for nonassociated object concepts. The results of Experiment 2 indicate that a high degree of association is more beneficial for pairs with small rather than large size differences. Conversely, the size difference between the items being compared has a greater influence for nonassociated than for associated pairs.

A model of the size-comparison task must therefore account for three basic facts: (1) Large size differences facilitate the comparison process, (2) associations also facilitate the comparison process, and (3) associations are more facilitative for pairs close in size. One approach to developing a model is to assume that people use multiple decision strategies to compare the sizes of pairs of objects. These strategies might be broadly classified as a general strategy based on analog information, and a set of specific strategies based on propositional information. The former strategy, which in various forms has often been incorporated into models of mental comparisons, presumably relies on analog size values stored with the memory representations for object concepts. Such analog size information can be used to compare both associated and nonassociated items. Decisions based on analog information will be made more quickly the greater the difference in analog values (see Holyoak, 1978; Moyer & Dumais, 1978).

The second set of strategies would rely on specific facts stored in memory about the relationships between object pairs. Consider, for example, a pair of closely related categories, such as dogs and cats. Since people often think of dogs and cats together, it seems likely that the normative size relation between the two categories may be part of their preexperimental knowledge of the world. So, if they are asked which category is generally larger, they will not have to compute the relation from information about the size of exemplars

of each category. Rather, they can simply retrieve the stored proposition "Dogs are generally larger than cats." In other cases, some known relationship that does not directly supply size information might guide an inference about relative size. Examples of such relationships include the "part of" relation (e.g., toe-foot), since wholes are always larger than one of their parts; the "in" relation (e.g., pearl-oyster), since containers are usually larger than their contents (although there are exceptions, such as "sock" and "foot"); and the "on" relation (e.g., teapot-stove), since the supporting object is usually (although not always) larger than the supported object. All of these are types of facts stored in semantic memory that could facilitate inferences about relative size. It is obvious that such information would be more likely to link highly associated items than words that are rarely associated. Comparison strategies based on such information would therefore produce faster decisions for associated pairs.

To account for the obtained interaction between size difference and degree of association, the "multiple decision strategies" approach assumes that both analog and propositional information become available to the subject in parallel (see also the analog-discrete parallel processing model proposed by Kosslyn, Murphy, Bemdeserfer, & Feinstein, 1977). If this is so, then a correct decision will be made relatively quickly either if the size difference between the two objects is large (based on analog information) or if the objects are closely associated (based on propositional information). As a result, degree of association will have less effect for pairs far apart in size, while size difference will have less effect for pairs that are highly associated. Thus there should be main effects of both size difference and association value, and the two variables ought to interact—exactly the results we observed.

It is clear that a variety of questions need to be answered in order to develop a model of the comparison process based on multiple decision strategies. In particular, the manner in which various types of propositional information might actually be used to derive inferences about relative size needs to be specified more clearly. In general, one would expect that some types of associations would yield size inferences more directly (and hence more quickly) than others. While the present study found no clear evidence of differences in degree of facilitation attributable to type of association, it may prove possible to derive more specific predictions of this nature from a more detailed version of the multiple decision strategies hypothesis.

An alternative approach attempts to account for the fact that the association and size-difference variables interact without postulating multiple decision processes. According to the "scan plus comparison" model of mental comparisons proposed by Moyer and Dumais (1978), subjects first access the stored analog size values for each of the presented items (scan) and then compare the two values. If the two size values differ by some

critical amount, the appropriate response is executed; if not, the difference is added to an internal accumulator and the scan plus comparison process is repeated. This model can account for the symbolic distance effect, since the greater the size difference between the two items being compared, the fewer iterations of the scan plus comparison process will be required to reach a decision.

With respect to the present results, the critical assumptions of the model are that the memory search process (scanning) is repeated on each iteration, and that this scanning process is facilitated by semantic association. That is, once one of the items has been located in memory, the second will be found more quickly if it is an associate (see Meyer & Schvaneveldt, 1971; Meyer, Schvaneveldt, & Ruddy, 1974, Note 2). Since the model assumes that the search process is repeated on each scan plus comparison iteration, the advantage of associated over nonassociated pairs will be greatest for pairs with small size differences (since these will require a relatively large number of iterations). Some readers may think this model implausible, since it requires that memory be searched all over again whenever an additional iteration is required. But suppose that the comparison itself cannot be made in the storage space, but only in some working memory space. Then exemplars of the to-be-compared items, once found, would be transferred to the working memory space for comparison. Under this arrangement, it seems quite reasonable for the search to begin anew, perhaps from the working memory space, if the comparison fails.

The results reported above cannot decide between the two approaches we have outlined, since both proposals can account for the three basic facts mentioned earlier. Since both hypotheses posit an analog strategy, manipulation of the presumed propositional links appears to be the most direct empirical test of these notions. For instance, if the multiple decision strategies notion is correct, supplying pair-specific propositional size information about nonassociates ought to speed comparison of those items, as should providing additional pair-specific propositional size information about the associates. Conversely, selectively interfering with the retrieval of propositional information about size would be expected to slow RT. But according to this proposal, pair-specific propositional information unrelated to relative size should not affect decision time. The scan plus comparison model, however, predicts that pair-specific propositional information will facilitate size comparisons by virtue of information, even when the new information is not relevant to size (or to inferences about size).

Although we cannot yet favor one of our proposals over the other, we have answered the questions posed at the outset. Both symbolic distance and semantic association effects can be observed in the same experiment, they are not additive, and both the scan plus comparison model and the hypothesis of multiple

decision strategies can account for the observed results. It should now be possible to begin to integrate the research on mental comparisons with the broader range of work on semantic memory.

REFERENCE NOTES

1. Friedman, A. Comparing words: An "internal psychophysics" for a nonphysical dimension. Paper presented at the XXIst International Congress of Psychology, Paris, July 1976.
2. Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. Activation of lexical memory. Paper presented at the annual meeting of the Psychonomic Society, Saint Louis, Missouri, November 1972.
3. Katz, A., & Paivio, A. Personal communication, 1979.

REFERENCES

- BANKS, W. P., & FLORA, J. Semantic and perceptual processes in symbolic comparisons. *Journal of Experimental Psychology: Human Perception and Performance*, 1977, 3, 278-290.
- BANKS, W. P., FUJII, M., & KAYRA-STUART, F. Semantic congruity effects in comparative judgments of magnitudes of digits. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, 2, 435-447.
- BOCK, R. D., & JONES, L. V. *The measurement and prediction of judgment and choice*. San Francisco: Holden-Day, 1968.
- BUCKLEY, P. B., & GILLMAN, C. B. Comparisons of digits and dot patterns. *Journal of Experimental Psychology*, 1974, 103, 1131-1136.
- CLARK, H. H. The language-as-a-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 335-359.
- COLLINS, A. M., & QUILLIAN, M. R. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1969, 8, 240-248.
- GLASS, A. L., HOLYOAK, K. J., & O'DELL, C. Production frequency and the verification of quantified statements. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 237-254.
- HOLYOAK, K. J. The form of analog size information in memory. *Cognitive Psychology*, 1977, 9, 31-51.
- HOLYOAK, K. J. Comparative judgments with numerical reference points. *Cognitive Psychology*, 1978, 10, 203-243.
- HOLYOAK, K. J., & GLASS, A. L. The role of contradictions and counterexamples in the rejection of false sentences. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, 215-239.
- HUTCHINSON, J. W., & LOCKHEAD, G. R. Similarity as distance: A structural principle for semantic memory. *Journal of Experimental Psychology: Human Learning and Memory*, 1977, 3, 660-678.
- JAMIESON, D. G., & PETRUSIC, W. M. Relational judgments with remembered stimuli. *Perception & Psychophysics*, 1975, 18, 373-378.
- KOSLYN, S. M., MURPHY, G. L., BEMESDERFER, M. E., & FEINSTEIN, K. J. Category and continuum in mental comparisons. *Journal of Experimental Psychology: General*, 1977, 106, 341-375.
- KUČERA, H., & FRANCIS, W. N. *Computational analysis of present-day American English*. Providence, R.I.: Brown University Press, 1967.
- LOFTUS, E. G. Category dominance, instance dominance, and categorization time. *Journal of Experimental Psychology*, 1973, 97, 70-94.
- MEYER, D. E., & SCHVANEVELDT, R. W. Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 1971, 90, 227-234.
- MEYER, D. E., SCHVANEVELDT, R. W., & RUDDY, M. G. Loci of contextual effects on visual word-recognition. In P. Rabbit

- (Ed.), *Attention and performance V*. New York: Academic Press, 1974.
- MOYER, R. S. Comparing objects in memory: Evidence suggesting an internal psychophysics. *Perception & Psychophysics*, 1973, **13**, 180-184.
- MOYER, R. S., & BAYER, R. H. Mental comparison and the symbolic distance effect. *Cognitive Psychology*, 1976, **8**, 228-246.
- MOYER, R. S., & DUMAIS, S. T. Mental comparison. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 12). New York: Academic Press, 1978.
- MOYER, R. S., & LANDAUER, T. K. Time required for judgments of numerical inequality. *Nature* (London), 1967, **215**, 1519-1520.
- PAIVIO, A. Perceptual comparisons through the mind's eye. *Memory & Cognition*, 1975, **3**, 635-647.
- PAIVIO, A. Mental comparisons involving abstract attributes. *Memory & Cognition*, 1978, **6**, 199-208.
- RIPS, L. J., SHOEN, E. J., & SMITH, E. E. Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 1-20.
- SCHAEFFER, B., & WALLACE, R. Semantic similarity and the comparison of word meanings. *Journal of Experimental Psychology*, 1969, **82**, 343-346.
- SCHAEFFER, B., & WALLACE, R. The comparison of word meanings. *Journal of Experimental Psychology*, 1970, **86**, 144-152.
- SEKULER, R., RUBIN, E., & ARMSTRONG, R. Processing numerical information: A choice time analysis. *Journal of Experimental Psychology*, 1971, **89**, 75-80.
- SHOEN, E. J. The verification of semantic relations in a same-different paradigm: An asymmetry in semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1976, **15**, 365-379.
- SMITH, E. E., SHOEN, E. J., & RIPS, L. J. Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, 1974, **81**, 214-241.
- WILKINS, A. T. Conjoint frequency, category size, and categorization time. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 382-385.

NOTES

1. Eliminating the data for the other two comparisons did not alter the results reported below.
2. This basic result has also been obtained by Katz and Paivio (Note 3).
3. These correlations were calculated using both the size norms based on mean ratings and those based on successive interval scaling. The results were virtually identical. All the size-difference analyses reported here are based on the successive interval norms.

(Accepted for publication April 1, 1979.)