JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR **13**, 237–254 (1974)

Production Frequency and the Verification of Quantified Statements[1]

ARNOLD L. GLASS, KEITH J. HOLYOAK, AND CARLA O'DELL[2]

Stanford University

In Experiment I subjects were required to generate nouns and adjectives which formed true completions of sentences of the form (*Quantifier*) *S are* ____. The quantifiers used were *All*, *Many*, *Some*, *Few*, and *No*. Three further experiments examined the effect of production frequency of the predicate on verification RT for quantified statements. True sentences with high-frequency category (Experiment II) or property (Experiments III and IV) predicates were verified more quickly than those with low-frequency predicates when the quantifier was *All*, *Many*, *Some*, or *Few*; this difference reversed for *No*-statements. False RT was fastest in all cases when the false sentence was semantically anomalous; but in Experiment IV, when degree of relatedness of subject and predicate words was varied within false but meaningful sentences, statements with high-related predicates were rejected more quickly. In general, the effect of semantic relatedness reversed for negative (*Few* and *No*) as opposed to positive quantifiers. An ordered attribute-search model was proposed to account for subjects' performance during both predicate production and sentence verification.

Recent studies have investigated the structure of semantic information and the process by which people use such knowledge in making semantic decisions (Collins & Quillian, 1969; Meyer, 1970; Rips, Shoben, & Smith, 1973). In such studies the dependent measure has been subjects' reaction time (RT) to verify simple propositions (for example, *All canaries are birds*). To test the generality of models for the verification task it is therefore crucial to identify those variables which influence verification RT, and to explore their effects over statements involving a range of semantic relations. This paper examines the effect on RT of the production frequency of the predicate—the probability that a particular predicate word is generated by subjects to complete a sentence. In order to vary the specific semantic relation to be verified, production frequency was manipulated in sentences containing five different quantifiers—*All*, *Many*, *Some*, *Few*, and *No*.

While English contains numerous quantifiers, only *All*, *Some*, and *No* have previously been studied in semantic-memory verification tasks (Meyer, 1970; 1973). Meyer (1970) found that statements such as *Some chairs are furniture* are verified more rapidly than the corresponding *All*-statement (*All chairs are furniture*), except when both statements are false. On this basis he proposed a two-stage model of the verification process according to which a subject presented with an *All*-statement is assumed to verify first the corresponding (implicit) *Some*-statement; only if the *Some*-statement is confirmed does he proceed to check the *All*-statement directly. It seems possible, however, that the differences in RT between *All*- and *Some*-statements obtained in Meyer's study were an artifact of his particular experimental design. He pre-

[1] This paper has greatly benefited from the extensive suggestions of our colleagues Gordon H. Bower, Herbert H. Clark, Robert J. Sternberg, and especially Edward E. Smith. While this study was conducted A. Glass held an N.S.F. graduate fellowship and K. Holyoak held a Stanford University fellowship. This research was supported by Grant MN13850-06 from the National Institute of Mental Health to our advisor, Gordon H. Bower.

[2] Now at the University of Oregon, Eugene, Oregon.

sented *Some-* and *All*-statements in blocks to different subjects, and while the false *All*-statements often contained semantically related subject and predicate words (*All stones are gems*), false *Some*-statements were generally anomalous (*Some typhoons are wheats*). This suggests that the longer latencies for *All*-statements recorded by Meyer may have been due to the more difficult semantic discrimination required to separate true and false *All*-statements. Recent studies by Glass and Holyoak (in press) and Rips (in preparation) have in fact found that Meyer's result does not replicate using procedures which avoid this confounding. Accordingly, in the present experiments statements of all quantifier-types were presented in mixed lists and semantically-related false statements were introduced for each quantifier-type. Under these conditions it seemed unlikely that subjects could resort to special strategies for particular quantifiers.

The production frequency measure used in the present experiments is similar to semantic relatedness variables (measures of the degree of overlap between the meanings of subject and predicate words) which have been found to influence verification RT in several previous studies. Wilkins (1971) reported that subjects' speed in confirming that an instance is a member of a category is positively correlated with the frequency with which the instance is generated as a response to the category name in an instance-production task. Conrad (1972) made use of a similar production-frequency measure in predicting true RT to property statements; while Loftus (1973) demonstrated that the frequency with which the category is generated as a response to an instance name can also predict categorization time. Rating scales have also been used to index semantic relatedness. It has been found that subjects categorize most quickly those instances which are rated as highly related to, or typical of, a given category (Rips et al., 1973; Rosch, 1973). Rating-scale and production-frequency measures of relatedness appear to be closely related. Rips et al. found a correlation of .85 between their relatedness ratings and the instance-production norms from which the items used by Wilkins were derived.

In the present study we obtained a measure of production frequency which took account of the way in which the meaning of a noun is modified by the quantifier associated with it. In our task subjects were required to generate true one-word completions for statements of the form (*Quantifier*) *S are* _____, where *S* is a simple concept. The frequency with which a particular response was given was taken as an index of its semantic relatedness to the quantified subject noun phrase. Experiment I deals with the collection of production frequency norms by this procedure and presents the results.

The remaining three experiments use these production-frequency norms to predict verification times for statements involving five different quantifiers. For *All-*, *Some-*, and *Many*-statements, it is predicted that true RT will decrease as production frequency increases. However, the quantifiers *No* and *Few* are negatives (Just & Carpenter, 1971), and, as will be argued below, the Clark and Chase (1972) model of negation suggests that the effect of semantic relatedness should reverse for negative as opposed to positive statements. Experiment II dealt with sentences in which the predicate was a category name, and Experiment III with the verification of property statements. Finally, Experiment IV examined in greater detail the process by which judgements about *Few-* and *Many*-statements are made.

EXPERIMENT I

Method

Subjects in Experiment I were presented with quantified statements in which the predicate was omitted, and for each sentence were asked to supply nouns and adjectives which would form true completions. Twenty nouns, with Thorndike-Lorge frequency greater than 20 per million and rated imagery value greater than 6.0 (on a 7-point scale) were selected from the norms of Paivio, Yuille, and Madigan (1968). Each of these words was then used as the subject word in five incomplete sentences of the form *(Quantifier) S are* , in which the quantifier was *All*, *Many*, *Some*, *Few*, and *No*. The resulting 100 incomplete sentences were collected into five booklets. Each page of the booklets contained a single sentence, with the headings *Adjectives* and *Nouns* typed below it. Booklets were 20 pages long, and contained equal numbers of sentences of each quantifier type. Each of the 20 subject words appeared in one sentence only in a given booklet. Five different versions of the booklets were constructed, such that across all versions each of the 100 incomplete sentences occurred once. The pages of the booklets were arranged in several different pseudorandom orders.

For each sentence, subjects were given 30 sec to list under the appropriate headings as many nouns and as many adjectives as they could think of that would complete the sentence to form a true statement. At a signal from the experimenter they turned the page and began work on the next item. Subjects were shown examples of the two response classes and were encouraged to try to provide both noun and adjective responses for all statements. Each subject completed two of the five booklet-types; a 2 min break intervened between booklets. Across subjects, each booklet-type was administered first and second equally often. Testing was done in two group sessions.

Fifty Stanford introductory psychology students participated in the experiment in order to satisfy a course requirement. Each quantified statement was completed by 20 subjects.

Results and Discussion

Both items and subjects were treated as random effects in all analyses of variance reported in this paper, and minimum quasi *F*-ratios were derived using the methods outlined by Clark (1973). This procedure results in significance levels substantially more conservative than those obtained in conventional subject analyses. The symbol F' will be used to denote minimum quasi-Fs, and t' will denote the related quasi-t statistic. Ordinary F and t values will also occasionally be reported, in which case the random variable on which they are based will be indicated.

The number of noun and adjective predicates generated for incomplete sentences of each quantifier-type were tabulated separately for each subject, pooling across the two booklets. Twenty-five percent of the total responses clearly did not form unequivocally true statements. These were eliminated from all analyses reported here; their exclusion did not substantially change the pattern of results. Analysis of variance revealed significant effects of quantifier-type, $F'(4, 185) = 48.4$, $p < .001$; predicate-type, $F'(1, 29) = 21.3$, $p < .001$; and also the interaction between these two variables, $F'(4, 167) = 23.9, p < .001$. The mean number of completions per sentence per subject are reported in Table 1.

TABLE 1

MEAN NUMBER OF PROPERTY AND CATEGORY PREDICATES PER SENTENCE PER SUBJECT GENERATED FOR INCOMPLETE SENTENCES OF EACH QUANTIFIER-TYPE

	Predicates	
Quantifier-Type	Property (Adjective)	Category (Noun)
All	.82	.63
Many	2.57	1.07
Some	2.76	1.07
Few	.77	.36
No	.67	1.00

To examine more closely the effect of the five different quantifiers in determining the number of responses, comparisons among the individual means were made using the Newman–Keuls method. *Some* produced the most responses, followed in order by *Many*, *No*, *All* and *Few*. Only the difference between *Some* and *Many* as opposed to the other three quantifiers was significant ($p < .01$). The low number of *All*-statement completions is presumably due to the fact that the criteria which delimit the range of true *All*-statements are more rigorous than those for any other quantifier.

Perhaps of more psychological interest is the relatively low number of responses given to the negative quantifiers *Few* and *No*. The low number of predicates produced as *No*-statement completions might especially appear surprising. *No S are P* implies that *Some S are P* is false, and one might suppose that for most subject words, there exist more possible predicates that would make *Some S are P* false than would make it true.

The fact that predicates positively related to the subject noun are nevertheless produced more readily has implications for a model of the generation process involved in the sentence-completion task. We suggest that for all quantifier-types the subject generates responses by searching through positive attributes of the subject noun. If the quantifier is positive (*All*, *Some*, or *Many*), he overtly produces predicates which express the attributes of the subject category. However, the generation of negative properties (true *Few*- and *No*-statements) is indirect. The positive attributes of the subject category are still searched, but now the person must also compute contradictions of these attributes, and respond with predicates which express these contradictions. For instance, in completing *No fires are* ______, the person might retrieve a positive attribute of fires, such as "hot," and from this derive the correct response "cold." The increased task difficulty resulting from the need for this extra step would result in fewer predicates being produced for negatively quantified statements. Similar hypotheses concerning the priority of positive information have been proposed by Trabasso, Rollins, and Shaughnessy (1971) and Wason (1959).

An extension of this model is suggested by the relative frequencies of noun and property responses for the different quantifiers. For *All*, *Many*, *Some*, and *Few*, people supplied more adjectives than nouns as completions ($p < .01$). Presumably there are many adjectives which describe the subject concept's attributes, but relatively few names of categories to which any number of subject exemplars belong. But as the significant interaction of predicate-type with quantifier-type indicates, for *No* there is a trend (though insignificant) in the opposite direction. Noun responses are at least as frequent as are property responses. It would appear that in generating true *No*-statements, people give a few adjectives which contradict highly salient attributes of the subject concept; but then they switch to searching for categories from which all subject exemplars are excluded.

The close relationship between responses to positive and negative quantifiers implied by this production model is borne out by an inspection of those predicates which occur with high frequency as responses to nouns quantified by *Few* or *No*. The frequency with which each word appeared as a response to a particular incomplete sentence was tabulated across subjects. High-frequency responses were defined as those given by five or more of the 20 subjects who completed each sentence (for *No*, property predicates produced by four out of the 20 were included). Table 2 lists all the high-frequency completions of *Few*- and *No*-statements, along with the number of subjects who gave each response. For each high-frequency *Few*-statement completion, there was a corresponding high-frequency *Many*-statement response which expressed positive attributes contradicted by the *Few*-statement response. Similarly, for each common *No*-statement response there usually was a high-frequency *Some*- or *All*-statement predicate from which the negative response appeared to derive. The derivation appears to occur either through a contradiction (*All women are females* yields *No women are males*), or else by way of a mutual and frequently-given superordinate (*All chairs are furniture* yields *No chairs are tables*. Derivations of the latter type appear to be instances of responses

TABLE 2

NUMBER OF SUBJECTS PRODUCING POSITIVE STATEMENTS RELATED TO HIGH-FREQUENCY NEGATIVE STATEMENTS

Negative Statements	Freq.	Related Positives	Freq.
Property predicates			
Few arrows are dull.	9	Many arrows are sharp.	7
Few arrows are crooked.	7	Many arrows are straight.	6
Few blossoms are ugly.	10	Many blossoms are beautiful.	10
Few diamonds are cheap.	8	Many diamonds are expensive/valuable.	10
Few factories are small.	6	Many factories are big/large.	5
Few professors are stupid.	5	Many professors are intelligent/educated/ smart.	9
No fires are wet.	4	No good lexical counterpart.	—
No diamonds are soft.	10	All diamonds are hard.	9
No fires are cold.	12	All fires are hot.	13
No forests are treeless.	5	No good lexical counterpart.	—
No prisoners are free.	9	All prisoners are jailed/incarcerated/ imprisoned/captive/restricted.	12
No horses are green.	4	Some horses are brown/black/gray/white.	23
Noun predicates			
Few professors are women.	8	Many professors are men.	8
Few doctors are women.	7	Many doctors are men.	7
No mothers are males.	7	All mothers are females/women.	17
No professors are children.	8	Some professors are men/women/fathers.	22
No valleys are mountains.	8	Some valleys are deep/shallow	14
No horses are cows.	10	All horses are animals/mammals	20
No chairs are tables.	5	All chairs are furniture	8
No birds are fish.	6	All birds are animals.	7

generated by a search for categories which exclude the subject category. But the obvious semantic relationships between the subject noun and frequently-produced predicates suggest that here too the search most often begins at the positive attributes of the subject concept. The two cases in which positive counterparts to high-frequency *No*-statement responses were absent may be ascribed to the fact that no common, one-word expression of the relevant attribute is available in English. Thus *wet*, as used in *No fires are wet*, contradicts a positive attribute of fires roughly equivalent to "incompatible with water."

The frequencies with which high-frequency *All*-statement predicates were given when the quantifier was changed to *Some* were also tabulated. Thirteen adjectives and 14 nouns occurred as *All*-statement completions five or more times out of the possible 20. Their mean frequencies were 7.7 and 7.1, respectively, when the quantifier was *All*. For *Some*, the corresponding frequencies were 4.7 and 1.6. This reduction in production frequency suggests that relatedness of subject and predicate words may be influenced by the quantifier.

If the production frequency norms compiled on the basis of Experiment I are valid measures of the semantic relatedness of the subject and predicate concepts in quantified statements, the norms should allow us to predict RTs in a verification task. Our next step was to test the effect of production frequency on verification latencies using statements in which the predicate was a category name.

EXPERIMENT II

Method

Subjects were timed as they responded true or false to visually-presented sentences of the form (*Quantifier*) *S are P*, where *P* was in each case a plural noun.

Materials. For each of the quantifiers *All*, *Many*, *Some*, and *No*, 12 true sentences were selected. Within each quantifier-type each subject noun was used twice, once with a high-frequency noun predicate (given by five or more subjects in Experiment I), and once with a low-frequency predicate (given by exactly one subject). The mean production frequency of high-frequency sentences was 7.3, and did not differ significantly across quantifiers. The quantifier *Few* was not used in this experiment, since not enough high-frequency noun completions to *Few*-statements were available. The design of the experiment, with two examples of each statement-type used, is outlined in Table 3.

According to some intuitions, the quantifier *Many* (and also *Few*) can sometimes produce ambiguous sentences. All *Many*-statements used in Experiment II (and the other quantified statements as well) were therefore selected on the basis of being consistently interpreted as true or false by an informal panel of graduate students.

The same subject nouns used to form true sentences (13 different nouns in all) were also used to form corresponding false sentences. Two levels of semantic relatedness of subject and predicate words were used for false as well as true sentences. High- and low-related false statements were also generated on the basis of the production frequency norms collected in Experiment I. High-related false *All*-statements were formed by using predicates which could form true *Some*- or *Many*-statements. For *Many*, high-related false statements were based on true *Few*- completions, while for *Some*-statements, responses to *No S are* were considered high-related falses. The basis of the latter assignments was the observation, supported by the examples of *No*-statements in Tables 2 and 3, that *No*-statement completions tend to be semantically-related to the subject noun. For all three of these quantifiers, low-related falses were generated by rearranging the same

TABLE 3

DESIGN OF EXPERIMENT II, WITH TWO EXAMPLES OF EACH STATEMENT TYPE

Quantifier	High-related	Low-related
True statements		
All	All mothers are women.	All mothers are parents.
	All blossoms are flowers.	All blossoms are plants.
Many	Many books are novels.	Many books are Bibles.
	Many factories are polluters.	Many factories are structures.
Some	Some birds are canaries.	Some birds are parrots.
	Some horses are mares.	Some horses are jumpers.
No	No horses are cows.	No horses are bees.
	No chairs are tables.	No chairs are feet.
False statements		
All	All buildings are houses.	All buildings are roses.
	All forests are parks.	All forests are houses.
Many	Many horses are albinos.	Many horses are tents.
	Many buildings are tents.	Many buildings are albinos.
Some	Some professors are children.	Some professors are birds.
	Some animals are plants.	Some animals are bows.
No	No arrows are weapons.	No arrows are objects.
	No horses are mammals.	No horses are runners.

predicates to form anomalous sentences.

By definition, the predicates of false *No*-statements must be true of the corresponding *Some*-statements. To form high-related false *No*-statements, therefore, predicate nouns were drawn from high-frequency *All*-, *Many*-, or *Some*-statements, while low-frequency positive predicates were used to form low-related false *No*-statements.

The complete item set, then, contained 48 true and 48 false sentences. Mean word frequency of the predicates (from Kučera & Francis, 1967) was closely controlled across all cells of the design, as was mean length of each sentence. Each sentence was typed in capital letters on a 4 × 6 in. white card.

Procedure. Sentences were presented by means of a tachistoscope. At a signal from the experimenter, the subject initiated a trial by pressing a start button. A dot then appeared in the viewer indicating where the sentence would begin. After 2 sec the sentence was presented. The subject then pressed one of two decision buttons indicating whether the sentence was true or false; this response stopped a timer and removed the sentence. Before the next trial began the experimenter informed the subject whether his response had been correct. Assignment of hand to true response button was counterbalanced across subjects. Subjects were instructed to respond as quickly as possible, but to make as few errors as possible.

Test sentences were presented in three blocks of 32 items; all experimental conditions were represented equally often in each block. Presentation order within these item blocks was pseudorandom and was changed for each subject. Thirty-two different practice sentences, illustrating all quantifier-types, were given prior to the test items.

Subjects were 16 Stanford undergraduates and junior college students (nine males, seven females), who participated either for pay or course credit.

Results

RTs which exceeded the subject's mean RT for that item type by 2 sec (<2% of responses) were counted as errors. The overall error rate was 5.5%; among conditions error rates were generally positively correlated with RT. Errors were not replaced. In order to equalize variances the data were transformed logarithmically. Data from true and false responses were analyzed separately.

The geometric mean RT for each condition is reported in Table 4. The true response data will be considered first. The effect of different quantifiers was significant, $F'(3, 36) = 4.85$, $p < .01$. Post hoc comparisons among individual means using the Newman–Keuls method revealed that true RT for *All*-, *Many*-, and *Some*-statements did not differ significantly, while *No*-statements were verified more slowly than all except *Many*-statements ($p < .05$). The main effect of semantic relatedness, as measured by production frequency, was not significant, but the interaction between relatedness and quantifier-type was, $F'(3, 37) = 4.61$, $p < .01$. The components of this inter-

TABLE 4

GEOMETRIC MEAN RT FOR DIFFERENT TYPES OF QUANTIFIED STATEMENTS WITH CATEGORY PREDICATES IN EXPERIMENT II

	True		False	
Quantifier	High-related	Low-related	High-related	Low-related
All	1571	1742	1882	1550
Many	1658	1854	1996	1657
Some	1557	1626	1607	1463
No	2032	1803	1911	1993

action were examined using orthogonal contrasts. As predicted, only the difference between the mean of the three positive quantifier-types and that of the *No*-statements proved to be significant, $t'(37) = 5.20$, $p < .001$. For the positive quantifiers high-frequency statements were verified most rapidly, $t'(37) = 2.90$, $p < .01$, whereas for the negative quantifier *No* the high-frequency sentences were verified slowest, $t'(37) = 2.51$, $p < .02$.

Basically the same pattern, only reversed, emerges in the data for false statements. Here low-related sentences produced significantly lower overall RT, $F'(1, 35) = 15.9$, $p < .001$, but the pattern of differences varied across quantifier-types, $F'(3, 37) = 4.74$, $p < .01$. Orthogonal comparisons revealed that the direction of the difference in RT between high- and low-related statements again reversed between the three types of positively-quantified statements, on the one hand, and *No*-statements on the other, $t'(37) = 4.76, p < .001$. High-related false sentences were more difficult to reject than were low-related statements when the quantifier was *All* or *Many*, $t'(37) = 4.58$, $p < .001$; or *Some*, $t'(37) = 2.05$, $p < .05$. For *No*-statements the trend was in the opposite direction, though not significant.

Overall RT for the various quantifier-types also differed for false sentences, $F'(3, 53) = 12.5$, $p < .001$. Newman–Keuls comparisons indicated that false *Some*-statements were rejected the quickest ($p < .05$), followed by *All*- and *Many*-statements, which did not differ significantly. *No*-statements required more time to discomfirm than all other quantifier-types ($p < .05$).

Discussion

The pattern of RTs in Experiment II appears to depend on two factors. First, the effect of negation (quantification with *No*) produces a consistent increase in difficulty across quantifiers. The overall difficulty of negative statements agrees with earlier results (Clark, 1970; Clark & Chase, 1972; Just & Carpenter, 1971). Secondly, semantic relatedness, as determined by the production frequency norms, and the procedure for generating related false statements, proved to be a reliable predictor of RT to verify quantified statements with category predicates. The results for the three positive quantifiers are completely consistent with previous research: High semantic relatedness speeds up verification of true statements (Conrad, 1972; Rosch, 1973; Rips et al., 1973); but slows down rejection of false statements (Kintsch, Crothers, & Berman, 1970; Meyer, 1970; Wilkins, 1971).

The reversal of the effect of relatedness in the case of *No*-statements has a clear interpretation in the context of the Clark–Chase model of negation (Clark, 1970; Clark & Chase, 1972; Just & Carpenter, 1971). *No S are P* would be represented as *False* (*Some S are P*). The truth of the embedded proposition is assumed to be evaluated first. The outer negation is then processed, reversing the computed truth value of the statement. A false *No*-statement, then, involves the implicit verification of a true *Some*-statement; consequently, any variable which facilitates verification of a *Some*-statement will also facilitate processing of the corresponding *No*-statement. And since high-related false *No*-statements were generated using frequent completions to *Some*- and *All*-statements, while the low-related falses contained infrequent predicates from positive sentences, it follows that the implicit *Some*-statement should be verified faster in the former case. This analysis explains the trend in favor of high-related false *No*-statements being rejected sooner. Meyer (1973) has also found that variables influencing the verification of true *Some*-statements produce similar effects on false *No*-statements.

A similar argument can account for the significantly longer latency to verify true high-related *No*-statements. According to the above analysis, a true *No*-statement requires the implicit rejection of a false embedded *Some*-statement. As our results for false data demonstrate, for positive quantifiers it is more difficult to reject false statements in which the subject and predicate are related than it is to reject anomalous sentences. And the comparison of high- versus low-related true *No*-statements amounts in fact to a comparison between sentences with related versus anomalous predicates. For as Tables 2 and 3 indicate, the predicates of high-frequency *No*-statements are semantically related to the subject concept (*No chairs are tables*); on the other hand, infrequent predicates tend to be anomalous (*No chairs are feet*). The higher relatedness of the predicates in high-frequency *No*-statements slows down rejection of the implicit *Some*-statement; consequently, verification RT should be increased.

The results of Experiment II indicate that production frequency is a reliable predictor of RT to verify statements with category predicates. Our hypothesis was that production-frequency norms, together with an analysis of negation, would be sufficient to successfully predict verification RTs with quantified property-statements as well. Accordingly, Experiment III extended this approach to sentences with adjective predicates, and also to an additional quantifier, the negative *Few*.

Experiment III

Method

Subjects were timed as they verified statements quantified by *All*, *Many*, *Some*, *Few*, and *No*. The selection of true sentences proceeded on the basis of production frequency norms in the same way as in Experiment II except that adjectival property predicates were used. Twelve true and 12 false *Few*-statements were included; otherwise, the number of items in each condition was the same as before. The overall mean frequency of high-frequency statements was 9.1; the means for *Few*- and

TABLE 5

Design of Experiment III, with Two Examples of Each Statement Type

Quantifier	High-related	Low-related
True statements		
All	All arrows are pointed.	All arrows are narrow.
	All diamonds are hard.	All diamonds are solid.
Many	Many blossoms are beautiful.	Many blossoms are soft.
	Many forests are green.	Many forests are wet.
Some	Some chairs are soft.	Some chairs are orange.
	Some armies are large.	Some armies are trained.
Few	Few arrows are crooked.	Few arrows are pink.
	Few blossoms are ugly.	Few blossoms are brown.
No	No forests are treeless.	No forests are purple.
	No horses are green.	No horses are airborne.
False statements		
All	All fires are yellow.	All fires are rusty.
	All birds are swift.	All birds are hard.
Many	Many books are round.	Many books are bleached.
	Many chairs are gigantic.	Many chairs are witty.
Some	Some horses are blue.	Some horses are rainy.
	Some arrows are soft.	Some arrows are mad.
Few	Few valleys are green.	Few valleys are settled.
	Few buildings are tall.	Few buildings are solid.
No	No streets are straight.	No streets are lighted.
	No horses are fast.	No horses are loved.

No-statements were slightly lower. Word frequency as indexed by the Kučera and Francis (1967) norms tended to be somewhat higher for the high-related predicates; word length in syllables was closely matched across conditions. Two levels of relatedness for false statements were created using the same procedure as was followed in Experiment II. Items for false *Few*-statements were constructed in the same way as were false *No*-statements; therefore, both of these quantifiers were used equally often with the same false sentences. Assignment of negative quantifier to sentence was counterbalanced across subjects, so that only one of the two quantifiers was presented to each subject with a given sentence. The design of the experiment, with two examples of each statement-type, is outlined in Table 5.

The apparatus used was identical to that of Experiment II, and the procedure highly similar. The 120 test items were presented in four balanced, randomly-ordered blocks of 30. These were preceded by a practice set of 40 different sentences representing all quantifier-types.

Subjects were 14 Stanford introductory psychology students (nine male) who participated in order to satisfy a course requirement.

Results

The RTs were transformed and analyzed in a manner similar to Experiment II. Geometric mean RTs for all conditions are given in Table 6. The overall error rate was 8.0%, and errors within conditions were again positively correlated with mean RT. For true responses the effect of quantifier-type was significant, $F'(4, 84) = 10.9$, $p < .001$. Newman–Keuls tests revealed that this effect was due to the longer latency of verifying negative (*Few*- and *No*-) statements as opposed to the three positively quantified types ($p < .01$). No other comparisons among the mean RTs for the different quantifier-types were significant.

Overall, high-related sentences were verified more quickly than low-related sentences, $F'(1, 63) = 17.4$, $p < .001$, but this effect differed significantly for *No*-statements as compared to all other quantified statements, $t'(71) = 3.60$, $p < .001$. For *All*, *Many*, *Some*, and *Few*, high-related sentences were verified more quickly, $t'(71) = 4.72$, $p < .001$; for *No*-statements there was a trend, although it was not significant, in the opposite direction.

For the false RT data, the interaction of relatedness and positive versus negative quantification was significant, $F'(1, 79) = 5.90$, $p < .025$. For *All*, *Many*, and *Some*, high-related false statements took longer to reject than low-related sentences. This effect was significant in the subject analysis, $t(52) = 3.33$, $p < .002$, and in an item analysis, $F(1,15) = 4.54$, $p = .05$, although the minimum F' value fell short of formal significance, $F'(1, 28) = 3.22$, $p < .10$. For *Few*- and *No*-state-

TABLE 6

GEOMETRIC MEAN RT FOR DIFFERENT TYPES OF QUANTIFIED STATEMENTS WITH PROPERTY PREDICATES IN EXPERIMENT III

	True		False	
Quantifier	High-related	Low-related	High-related	Low-related
All	1287	1687	1530	1446
Many	1251	1517	1629	1492
Some	1387	1650	1660	1460
Few	1763	2029	1802	1884
No	1866	1803	1609	1775

ments, on the other hand, high-related false sentences tended to be correctly rejected sooner. This effect was significant in the subject analysis, $t(52) = 2.58$, $p < .02$, but not in the corresponding item analysis, $F(1, 11) = 2.65$, $p < .20$. Overall, positive statements were rejected more quickly than were negative statements, $F'(1, 97) = 18.4$, $p < .001$.

Discussion

The same factors invoked to explain the verification of quantified statements with category predicates in Experiment II appear to account for the pattern of results obtained in Experiment III for statements with property predicates. Only the two negative quantifiers, *Few* and *No*, produced reliably longer overall response latencies. As in Experiment II, high-relatedness decreased true RT and increased false RT for *All*-, *Some*-, and *Many*-statements. This pattern was again reversed for *No*-statements. The results for *No*-statements provide further support for the hypothesis that sentences of the form *No S are P* are processed as *False* (*Some S are P*).

As Table 6 indicates, for false responses the effect of relatedness is the same for sentences involving the negative *Few* as it is for *No*—that is, high-related false sentences tend to be rejected most quickly. The explanation for this effect is analogous to that provided for processing of *No*-statements if we assume, following Just and Carpenter (1971), that *Few S are P* is understood as *False* (*Many S are P*).

But now comparison of the true RT data for the two negative quantifiers reveals an apparent paradox. For true statements, *Few* does not reverse the effect of relatedness as *No* does; rather *Few* behaves here like a positive quantifier. However, analysis of the items used in the true *Few* conditions suggests an explanation for this result that is in fact compatible with the hypothesis that *Few* is processed as a negative. As Table 2 demonstrates, high-frequency completions of *Few*-statements are invariably contradictions of high-frequency *Many*-statement predicates. The predicate of *Few arrows are dull*, for instance, contradicts a frequently-produced attribute of arrows, "sharp." Low-related *Few*-statements, such as *Few arrows are wide*, are also contradictions of positive concept attributes—in this case, "narrow." But the production frequency of "narrow" as a predicate for "arrow" is lower than that of "sharp."

Suppose that the attribute search process formulated to deal with the sentence-completion task (Experiment I) is applicable to the verification process as well. This assumes that an incomplete sentence of the form (*Quantifier*) *S are* causes the person to generate true predicates by an ordered search of the attributes of the subject concept (*Quantifier*) *S*. The order of the search is reflected in the frequency with which predicates expressing a particular attribute are produced. Suppose that verification also involves a search of the subject concept's attributes, and that the order of search is the same as in the predicate-production task. But now the search is not used to generate predicates, but rather to check whether any subject attribute is contradicted by the attributes of the given predicate. Assuming that the search is self-terminating, then sentences with predicates which contradict high-frequency positive attributes of the subject concept will most quickly be rejected.

This verification model is supported by the results for true *Few*-statements. Assume that *Few*-statements are processed as *False* (*Many S are P*). For true *Few*-statements the embedded *Many*-statement is false. A high-related predicate, as in *Few arrows are dull*, will quickly be found to contradict the attribute "sharp." But if the predicate is low-related, as "wide" is, it will take longer for the person to discover the contradiction. It follows that the implicit *Many*-statement will be rejected and the explicit *Few*-statement consequently verified more quickly in the high-related condition.

Why is it the case, then, that when the quantifier is *No*, which is also negative, it is the low-related items which are verified more rapidly? As the examples provided in Table 5 suggest, the true low-related *Few*-statements are relatively meaningful (*Few blossoms are brown*), but the corresponding *No*-statements are generally semantically anomalous (*No chairs are angry*). This suggests that anomalous sentences do not require an extensive attribute search, as low-related but meaningful sentences do; rather such anomalous sentences are rejected on the basis of initial information which precludes further processing. Experiment IV was designed to test these hypotheses concerning the process by which people make decision about false propositions.

EXPERIMENT IV

In Experiment IV we substituted the quantifier *Many* in *Few*-statements, making explicit the false *Many*-statements which we argue are computed implicitly in order to verify true *Few*-statements. To this we added anomalous *Many*-statements. We hypothesized that anomalous sentences would again be rejected most rapidly. Sentences with high-related but false predicates should require longer processing, terminating with the discovery of a contradiction. Low-related but still sensible false predicates would require a longer attribute search, and hence be slowest to reject. For the corresponding true *Few*-statements we predicted parallel effects, since we assume that the false *Many*-statements are nevertheless computed implicitly.

Method

Twelve subject nouns were each paired with factorial combinations of five property predicates (high- and low-frequency *Few*- and *Many*-responses, and an anomalous predicate) and the quantifiers *Many* and *Few*. Examples of the 10 resulting statement-types are given in Table 7. For each level of relatedness, the selected *Few*- and *Many*-predicates both expressed attributes related to the same semantic dimension. For instance, for sentences based on *Many/Few blossoms are* the high-related *Many*- and *Few*-predicates were "beautiful" and "ugly," respectively, while the corresponding low-related predicates were "soft" and "hard." The high frequency predicates, both for *Many* and *Few*, were always produced by five or more subjects in Experiment I. At least one of the low-frequency predicates was given by one subject; if a corresponding contradictory property did not appear for the other quantifier, we generated one intuitively. The truth values of all sentences were again assessed by several graduate students, and ambiguous items were replaced.

The fifth predicate-type, which produced a semantically-anomalous sentence, was formed by rearranging the *Few*- and *Many*-predicates across subject nouns. The anomalous sentences were considered to be false when quantified by *Many* and true when quantified by *Few*. The resulting set of 120 sentences contained equal numbers of *Few*- and *Many*-statements. Sixty per cent of the *Few*-statements and 40% of the *Many*-statements were true items. For *Many*-predicates, mean word-frequency as determined by the Kučera and Francis (1967) norms did not vary as a function of relatedness; word frequency was somewhat lower for *Few*-predicates, especially those in the high-related condition. Predicate word-length was controlled across all conditions.

The method of stimulus presentation was identical to that of Experiment II. Early results showed that subjects often had difficulty in interpreting anomalous *Few*- statements (*Few arrows are intelligent*) as being true. It was therefore necessary to explicitly instruct subjects that "Few" was to mean "Not many, and possibly even none." Twenty-four different practice sentences were presented prior to the test items.

Twenty Stanford undergraduates and junior college students (10 male) served as subjects. Each subject received either pay or course credit for participation. Data from four subjects with error rates over 20% overall, or

50% in a single condition, were excluded, and four extra subjects were tested to replace them.

Results and Discussion

RTs were transformed logarithmically prior to statistical analysis. Geometric mean RTs for all conditions are presented in Table 7. The overall error rate was 8.8%, and across conditions error rates generally increased with RT.

The anomalous conditions were excluded from the first analysis, which compared two levels of each of three variables—quantifier-type (*Many* vs *Few*), predicate-type (*Many*- vs. *Few*-responses), and production frequency (high vs low). Overall, *Many*-statements were verified more quickly than *Few*-statements, $F'(1,26) = 59.3, p < .001$. Of the two predicate-types, *Few*-predicates took longer to verify, $F'(1, 16) = 7.62$, $p < .05$. The result of major interest concerns the effect of production frequency. Statements with high-frequency predicates were verified more quickly than those with low-frequency predicates, $F'(1, 14) = 10.1$, $p < .01$, and production frequency did not interact significantly with any other variable.

A second analysis included the anomalous conditions. Predicate-type was treated as a single variable with five levels. Two additional comparisons were derived from this analysis. The interaction between true and false responses and quantifier-types, based on the four meaningful predicate-types, was tested by a polynomial contrast. This interaction was significant, $t'(103) = 5.34$, $p < .001$, indicating that the longer latency of false responses was greater for the positive quantifier *Many* than for the negative quantifier *Few*. This result is consistent with the Clark and Chase (1972) model of negation, which expects false responses to be relatively fast for negative statements.

An orthogonal comparison also revealed that for the three false *Many*-statement types, anomalous sentences were rejected significantly faster than were the two meaningful sentence-types (with high- or low-frequency *Few*-predicates) considered together, $t'(103) = 2.17$, $p < .05$. The corresponding comparison for true *Few*-statements was not significant. As indicated in Table 7, *Few*-statements with high-frequency predicates were verified more quickly than those with low-frequency predicates, but mean RT for verifying anomalous sentences fell midway between these two conditions.

This unexpectedly long latency to verify

TABLE 7

GEOMETRIC MEAN RT FOR DIFFERENT TYPES OF FEW AND MANY- STATEMENTS IN EXPERIMENT IV, WITH AN EXAMPLE OF EACH STATEMENT TYPE

Quantifier-type	Many-predicates High-freq.	Many-predicates Low-freq.	Few-predicates High-freq.	Few-predicates Low-freq.	Anomalous predicates
	True	True	False	False	False
Many arrows are . . .	Sharp	Narrow	Dull	Wide	Intelligent
	1346	1523	1634	1712	1561
	False	False	True	True	True
Few arrows are . . .	Sharp	Narrow	Dull	Wide	Intelligent
	1846	1970	1814	1953	1869

anomalous *Few*-statements is presumably related to the fact that a majority of subjects initially interpreted such sentences as *Few arrows are intelligent* as being false. Two of the four subjects whose data were excluded on the basis of high error rates were completely unable to accept such sentences as true. This difficulty is consistent with our retrospective linguistic intuition that *Few S are P* generally implies that while not many *S* are *P*, it is true that some *S* are *P*. Furthermore, the closely related expression *A few S are P* is unequivocal in asserting the existence of some *S* which are *P*.

With this exception, the pattern of results obtained in Experiment IV is completely in accord with predictions. The longer latency to verify *Few*-statements, the faster verification of sentences with high-frequency predicates for true *Many*- and true and false *Few*-statements, and the faster rejection of anomalous false *Many*-statements, all represent replications of findings obtained in Experiment III. In addition, the results for the three types of false *Many*-statements provide evidence that the verification process allows quick rejection of semantically-anomalous sentences, and involves a comparison process based on an ordered search through the attributes of the subject concept.

General Discussion

The results of our three sentence-verification experiments demonstrate the powerful predictive validity of measures of semantic relatedness based on production tasks. When the effect of negative quantification was taken into account by means of the negation model, the RT to verify true quantified statements, whether explicit or implicit, was invariably shorter for sentences with high- rather than low-frequency predicates. This was true both for statements with category and with property predicates and for statements involving five different quantifiers. We also found that false RT was fastest in all cases when the false sentence was semantically anomalous; but in Experiment IV, when degree of relatedness was varied within false but meaningful sentences, statements with high-related predicates were rejected more quickly.

These findings pose serious challenges to previous models of how semantic memory is organized and searched in verification tasks. The model proposed by Meyer (1970) to account for the verification of *All*- and *Some*-statements is not supported by the present data. A critical assumption of Meyer's model is that the implicit verification of a *Some*-statement is the necessary first stage in *All*-statement verification. The prediction which follows from this assumption is that true *All*-statements will take longer to verify than true *Some*-statements. The results of Experiments II and III indicate that in a mixed-statement design with related false statements of all quantifier-types—that is, in a context in which the relationship specified by the quantifier must be fully processed—there are no reliable differences in overall RT among any of the three positive quantifier-types, when production frequency of the predicate is controlled. For category statements, mean RT was 65 msec longer for *All*- than for *Some*-statements; for property statements there was a 32 msec reversal. Neither of these differences achieved statistical significance. At any rate, Meyer's model has no clear extension that would account for the verification of *Few*- and *Many*-statements.

Smith, Shoben, and Rips (1974) have proposed a feature-comparison model of the verification process for categorization tasks. This model involves two stages. In Stage 1, the overall similarity of subject and predicate words is assessed in terms of defining features (those which strictly define category membership) and characteristic features (those which are shared by most category instances) of the two categories. Smith et al. hypothesize that Stage 1 is the locus of semantic distance effects. If the overall similarity of the subject and predicate falls below a lower criterion, a quick false response is made; if an upper criterion is exceeded, the subject responds true immediately. Only if the overall similarity falls between these two bounds is Stage 2 executed, resulting in longer RT. Stage 2, which involves actually verifying that the defining features of the predicate are present in the subject category, plays a minimal role in determining semantic distance effects.

While this model as explicitly formulated applies mainly to *All*-statements with category predicates, it would seem that this general account of the effect of relatedness on RT could be extended to property statements and other quantifier types. Such an extension would correctly predict all the present results for true sentences. However, this model would be unable to accommodate the results obtained for false *Many*-statements in Experiment IV. The Smith et al. model implies that RT must always be monotonically related to the overall relatedness of subject and predicate words as indexed by typicality ratings and similar techniques. For false statements, their model predicts that the probability of Stage-1 rejection, and hence fast RT, is necessarily decreased by relatedness. Our extension of this model predicts the finding that anomalous sentences (*Many arrows are intelligent*), where the subject and predicate are least related, are rejected most quickly. But it seems likely that high-frequency *Few*-predicates, as in *Many arrows are dull*, are more closely related semantically to the subject category than are low-frequency *Few*-predicates, as in *Many arrows are wide*. But here relatedness does not inhibit rejection; rather, the former statement-type elicits faster false RT than does the latter.

The failure of this extension of the two-stage feature-comparison model to predict this result appears to relate to a basic processing assumption. According to this model, semantic relatedness does not determine RT by influencing the actual verification of the relation specified by the quantifier, but rather by determining the probability that the subject makes a fast correct response on the basis of minimal processing of the statement. This assumption suggests that in a task where false statements often contain highly-related subject and predicate words, people would either have very high error rates, or else would set their criterion (regarding subject and predicate similarity) for quick true responses so high that almost all meaningful sentences would require second-stage processing. In the latter case it would seem that degree of relatedness should no longer have a strong effect on true RT. The mixed quantifier-types and related false statements used in the present experiments, however, produced unexceptional error rates and robust effects of semantic relatedness.

The model of verification to be proposed assumes that the meaning of a word is composed of a set of more elementary attributes (Katz & Fodor, 1963). As Katz (1972) suggests, the sets of attributes (or semantic markers) representing lexical items may exhibit a degree of organization exceeding that of simple feature lists. Attributes may vary in abstractness from specific properties to general categories. The attributes of "canary," for example, might include not only "yellow" but also "avian." Abstract attributes would themselves imply more elementary meaning components, which the person can derive by means of semantic redundancy rules (Katz, 1972). Many individual attributes will therefore imply whole sets of other attributes; the attribute "avian," for instance, might imply such further attributes as "feathered" and "animate." Figure 1 shows the hypothesized relations between words (capital letters) and attributes (lower case letters). Attributes may be directly connected to words (double arrows) or implied by other attributes (single arrows). The latter type of association is a graph-theoretic realization of redundancy rules.

The introduction of redundancy rules as a psychological construct serves to simplify the search process we assume takes place during verification, since such rules can eliminate the need to postulate exhaustive searches through

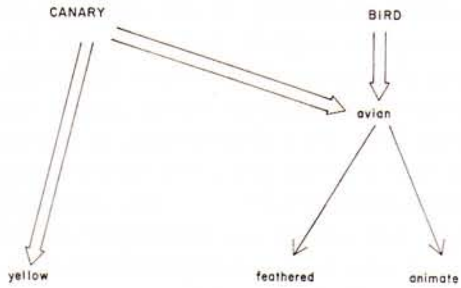

FIG. 1. Semantic representation of hypothetical word-to-attribute and attribute-to-attribute associations.

indefinitely long feature lists in order to account for the verification of true propositions.

Using this structural representation, we can outline an ordered attribute–search process to account for our results for both the predicate-production and for the verification task. Predicate generation and verification are intimately related in this model, which is why production frequency predicts RT for the various statement-types used. Both during predicate production and during verification, the subject is presumed to search the attributes of the subject concept in a serial order. For all quantifiers, the same subject attribute-set is searched; however, the person takes account of the meaning of the quantifier in selecting out and considering just those subject attributes which satisfy the relation expressed by the quantifier. For instance, if the subject word *people* is quantified by *Some* or *Many*, then attributes associated to subsets (for example, "female") or instances of the subject category are relevant. But the same attributes will not be selected for consideration when the subject is *All people*; rather, only defining attributes—those coded as true generically for any member of the category—will be activated. We have no detailed model of how the quantifier restricts the subject attributes considered, but for present purposes no such model is required; we simply need the result that a set of subject attributes are selected according to the quantifier.

During the production task the subject outputs words derived from these selected subject attributes as sentence completions. We assume that in the verification task the subject serially compares the set of derived attributes of the subject concept to the attribute set representing the predicate. If all the attributes of the predicate are found in the set derived from the subject category, the person responds "true." This search need not proceed laboriously through vast numbers of attributes; since an attribute set may be dominated by a single astract marker, a true decision may often be based on a match between a single pair of attributes. For example, when presented with the statement *All canaries are birds*, a person may first derive "avian" from the subject attribute-set; since "bird" would also be represented in memory as an attribute set dominated by "avian," a match would follow immediately.

Such an attribute search can account for the effect of relatedness (or production frequency) upon the speed of verifying true sentences. When the subject and predicate concepts are closely related, as in *All canaries are birds*, the first attributes derived from the subject (for example, "avian") are likely to be those allowing confirmation of the statement. On the other hand, these same highly salient attributes will be insufficient to produce a decision when considering the less-related sentence *All canaries are animals*. In the latter case search must continue until the subject attribute "animate" is found later in the search order. Continued search would involve decomposing the first subject attribute ("avian") into its elements (see Figure 1); in this case "animate" would then be recovered via redundancy rules in the first generation of attribute expansion. If decomposition of the first subject–attribute yields some overlap with the predicate attribute-set, but not a complete match, further comparisons will be made using a new attribute derived from the subject. The first subject attribute selected need not be the most specific, and hence this process has no necessary connection to category size. For instance, the attributes necessary to confirm that *All dogs are animals* are presumably derived from "dog" prior to those which allow verification of *All dogs are mammals* (Rips et al., 1973).

At what point will the subject decide to reject a false statement? Our results for false but meaningful *Many*-statements in Experiment IV suggest that at least some false decisions result from the discovery of a contradiction. If the person finds an attribute of the predicate that contradicts an attribute of the subject, he terminates the search and responds "false." Thus, sentences containing predicates which contradict a subject attribute early in the search order will be rejected more quickly than sentences contradicting later attributes.

At least two explanations might account for the fast rejection of anomalous sentences. Perhaps anomalous predicates invariably produce an immediate contradiction. Alternatively, the person may reject anomalous sentences on the basis of an initial holistic comparison of the relatedness of subject and predicate words (Smith et al., 1974).

It is clear that not all meaningful false sentences can be rejected on the basis of the type of contradiction discussed above. For instance, the high-related false *All*-statements used in Experiment II (for example, *All buildings are houses*) were produced by substituting *All* in statements which were true when quantified by *Some*. For this statement-type the predicate will not contradict any attribute (such as "physical object") which can be derived from the subject. Such sentences are perhaps rejected when a counterexample is found. For instance, a church is an example of a building that is not a house, so it provides a falsifying counterexample to *All buildings are houses*. This hypothesis concerning false decisions not based on direct contradictions is conjectural; but it does suggest that detailed examination of how sentences are disconfirmed may play an important role in guiding models of semantic memory.

The ordered attribute-search model outlined here seems sufficient at least to provide a framework for the present findings. Further research will determine the fruitfulness of this approach.

References

Clark, H. H. How we understand negation. Paper presented at COBRE workshop on Cognitive Organization and Psychological Processes, Huntington Beach, California, 1970.

Clark, H. H. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 335–359.

Clark, H. H., & Chase, W. G. On the process of comparing sentences against pictures. *Cognitive Psychology*, 1972, **3**, 472–517.

Collins, A. M., & Quillian, M. R. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 241–248.

Conrad, C. Cognitive economy in semantic memory. *Journal of Experimental Psychology*, 1972, **92**, 149–154.

Glass, A. L., and Holyoak, K. J. The effect of *some* and *all* on reaction time for semantic decisions. *Memory & Cognition*, 1974, in press.

Just, M., & Carpenter, P. Comprehension of negation with quantification. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 244–253.

Katz, J. J. *Semantic theory*. New York: Harper & Row, 1972.

Katz, J. J., & Fodor, J. A. The structure of a semantic theory. *Language*, 1963, **39**, 170–210.

Kintsch, W., Crothers, E. J., & Berman, L. N. The effects of some semantic and syntactic properties of simple sentences upon the latency of judgements of semantic acceptability. In *Studies in mathematical learning theory and psycholinguistics*, University of Colorado, 1970.

Kucera, H., & Francis, W. N. *Computational analysis of present-day American English*. Providence, RI: Brown University Press, 1967.

Loftus, E. F. Category dominance, instance dominance, and categorization time. *Journal of Experimental Psychology*, 1973, **97**, 70–74.

Meyer, D. E. On the representation and retrieval of stored semantic information. *Cognitive Psychology*, 1970, **1**, 242–300.

Meyer, D. E. Verifying affirmative and negative propositions: Effects of negation on memory retrieval. In S. Kornblum (Ed.), *Attention and performance IV*. New York: Academic Press, 1973.

PAIVIO, A., YUILLE, J. C., & MADIGAN, S. Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, 1968, **76**, Monograph Supplement No. 1, Part 2.

RIPS, L. J. Quantification and semantic memory. In preparation.

RIPS, L. J., SHOBEN, E. J., & SMITH, E. E. Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 1–20.

ROSCH, E. R. On the internal structure of perceptual and semantic categories. In T. M. Moore (Ed.), *Cognitive development and acquisition of language*. New York: Academic Press, 1973.

SMITH, E. E., SHOBEN, E. J., & RIPS, L. J. Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, 1974, in press.

TRABASSO, T., ROLLINS, H., & SHAUGHNESSY, E. Storage and verification stages in processing concepts. *Cognitive Psychology*, 1971, **2**, 239–289.

WASON, P. C. The processing of positive and negative information. *Quarterly Journal of Experimental Psychology*, 1959, **11**, 92–107.

WILKINS, A. T. Conjoint frequency, category size, and categorization time. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 382–385.

(Received August 3, 1973)