Role of Antonymy Relations in Semantic Judgments

Arnold L. Glass
Rutgers—The State University

Keith J. Holyoak
University of Michigan

John I. Kiger
Rutgers—The State University

Decision times for sentences or word pairs involving direct (e.g., boy–girl) or indirect (e.g., boy–sister) antonyms were measured in a sentence verification task and a lexical decision task. In Experiment 1 false sentences involving the direct antonyms were disconfirmed faster than sentences involving the indirect antonyms, even though the former were rated as more closely related in meaning. In Experiment 2 a smaller advantage for the direct antonyms was found in a lexical decision task, although “word” decisions were made more quickly for both types of antonym pairs than for pairs of unrelated words. Experiment 2 also investigated the degree of semantic facilitation obtained for category-instance word pairs. No significant latency differences were found for lexical decisions involving word pairs consisting of a category name (e.g., bird) and either a high frequency (e.g., robin) or low frequency (e.g., chicken) instance. The locus of the positive effect of semantic relatedness on falsification time for sentences involving antonyms is discussed in light of these results, as is the locus of the “typicality effect" obtained in previous semantic memory studies.

An important theoretical contrast that has emerged in models of semantic memory involves the processes by which people evaluate sentences asserting semantic relations such as category membership (e.g., A rose is a flower). Feature models assume that such decisions are based on a comparison of elementary features or attributes that define the subject and predicate concepts (McCloskey & Glucksberg, 1979; Schaeffer & Wallace, 1970; Smith, Shoben, & Rips, 1974). Such models assume that only elementary attributes are stored in memory and that other semantic relations, such as superordination, are inferred on the basis of a process of feature matching. In contrast, “direct storage” models propose that a variety of qualitatively distinct semantic relations are stored in memory (Collins & Loftus, 1975; Collins & Quillian, 1969; Glass & Holyoak, 1975; Meyer, 1970). The latter models emphasize the role that retrieval of specific semantic relations may play in the evaluation of sentences. (See Smith, 1978, for further discussion of this distinction.)

Much of the evidence for alternative types of decision processes is derived from studies of how subjects decide that sentences are false. One major finding is that anomalous sentences (e.g., A rose is a chair), in which the subject and predicate concepts are very dissimilar in meaning, can...
be rejected more quickly than other false sentences (e.g., *A rose is a fruit*) in which the two concepts are more similar in meaning (Glass, Holyoak, & O’Dell, 1974; Meyer, 1970; Rips, Shoben, & Smith, 1973). In terms of a feature comparison process, the concepts in anomalous sentences have very few shared properties and many contradictory ones. If the properties of the two concepts are compared, numerous mismatches will quickly be found regardless of the order in which comparisons are made. This accumulated negative evidence will indicate that the two categories are mutually exclusive (Collins & Loftus, 1975; Glass, Holyoak, & Kossan, 1977; McCloskey & Glucksberg, 1979).

There is also evidence, however, that some “false” decisions are based on other types of stored semantic relations. A study by Holyoak and Glass (1975; see also Glass et al., 1974) suggested two specific types of information that can be used to reject false sentences of the form *All/Some S are P.* When S and P were disjoint members of the same category, decision time was reduced if the P term was a relatively high frequency (as indexed by normative data collected from an independent group of subjects) *false* completion of the frame *All/Some S are ______.* That is, high frequency sentences such as *All fruits are vegetables* were rejected more quickly than low frequency sentences such as *All fruits are flowers.* Holyoak and Glass interpreted this result as evidence that subjects can reject such sentences by retrieving the information that S and P are disjoint instances of the same superordinate category.

Holyoak and Glass (1975) also investigated the processing of false sentences quantified by *all* in which S and P were *not* disjoint (e.g., *All flowers are roses*). Such sentences were rejected more quickly if normative subjects produced a relatively frequent *true* completion (other than the presented P) for the frame *Some S are ______.* Holyoak and Glass interpreted this result as evidence that subjects disconfirm such sentences by retrieving an instance of the S category that can serve as a counterexample to the presented sentence (e.g., *tulip* would be a counterexample to the sentence *All flowers are roses*).

The Holyoak and Glass results suggest that some “false” decisions are based on retrieval of superordinate and subordinate relations rather than simply on comparisons of features of the S and P concepts. McCloskey and Glucksberg (1979) have criticized the Holyoak and Glass experiment on a number of points related to the selection of items. However, Lorch (1978) has replicated the effects of availability of contradictions and counterexamples in a series of experiments that are not subject to these objections. In addition, Lorch’s study confirmed that these availability effects cannot be explained in terms of the semantic similarity of the S and P terms. On the other hand, Lorch’s results also suggest that property comparisons play a role in the rejection of false sentences.

The present study was designed to investigate the evaluation procedures used to reject a particular class of false sentences: those in which the subject and predicate words are antonyms. Test sentences were generated by taking advantage of the fact that English contains a number of pairs of words that differ in gender (e.g., *man-woman, male-female, brother-sister*). These terms can be used to form contradictory sentences with subject and predicate concepts that are direct antonyms (e.g., *All brothers are sisters*) or indirect antonyms (e.g., *All brothers are females*). Furthermore, it is possible to use each word equally often in both types of sentences across a number of related sentences, thus controlling for differences among individual words.

Three distinct predictions can be made about the outcome of a verification experiment using such sentences. First, note that the direct antonyms (e.g., *brother-sister*) are if anything more similar in meaning than the indirect antonyms (e.g., *brother-female*), since both pairs of concepts differ on the gender dimension, but the former pair contains the same relational component (e.g., “sibling”). Accordingly, if people primarily evaluate
A. GLASS, K. HOLYOAK, AND J. KIGER

Table 1

Contradictory Sentences Used in Experiment 1

<table>
<thead>
<tr>
<th>Quantified subject</th>
<th>Direct antonym predicate</th>
<th>Indirect antonym predicate</th>
</tr>
</thead>
<tbody>
<tr>
<td>All/Some girls</td>
<td>boys</td>
<td>brothers</td>
</tr>
<tr>
<td>All/Some women</td>
<td>men</td>
<td>kings</td>
</tr>
<tr>
<td>All/Some females</td>
<td>males</td>
<td>uncles</td>
</tr>
<tr>
<td>All/Some sisters</td>
<td>brothers</td>
<td>boys</td>
</tr>
<tr>
<td>All/Some queens</td>
<td>kings</td>
<td>men</td>
</tr>
<tr>
<td>All/Some aunts</td>
<td>uncles</td>
<td>males</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Male/Female order</th>
<th>Direct antonym predicate</th>
<th>Indirect antonym predicate</th>
</tr>
</thead>
<tbody>
<tr>
<td>All/Some boys</td>
<td>girls</td>
<td>sisters</td>
</tr>
<tr>
<td>All/Some men</td>
<td>women</td>
<td>queens</td>
</tr>
<tr>
<td>All/Some males</td>
<td>females</td>
<td>aunts</td>
</tr>
<tr>
<td>All/Some brothers</td>
<td>sisters</td>
<td>girls</td>
</tr>
<tr>
<td>All/Some kings</td>
<td>queens</td>
<td>women</td>
</tr>
<tr>
<td>All/Some aunts</td>
<td>aunts</td>
<td>females</td>
</tr>
</tbody>
</table>

these sentences by in some way amalga-
mating information about matches and mis-
matches between features (McCloskey &
Glucksberg, 1979; Smith et al., 1974),
then the sentences with direct antonyms
should be rejected relatively slowly.

Second, in all of these sentences the two
concepts are contradictory with respect to
the gender dimension. If only this specific
information about the two concepts is
retrieved and used to evaluate the sen-
tence, then both types of sentence will be
rejected equally quickly.

Finally, direct antonyms may be marked
as disjoint members of a specific super-
ordinate category (or equivalently, as
having contrasting values on a relational
feature dimension, such as sibling). If this
information is retrieved relatively quickly,
it is possible that sentences containing
direct antonyms will actually be rejected
more quickly than those containing in-
direct antonyms. The latter two outcomes
equal or faster reaction times for sen-
tences with direct antonyms) emphasize
the retrieval of specific information about
contradictions, whereas the first possibility
emphasizes the use of a more general
comparison process involving the amalga-
mation of information from multiple feature
comparisons. However, all of the above
hypotheses assume that antonymy will
have an effect on a stage of processing at
which a decision is made about the truth
value of the sentence. Experiment 2 will
explore the possibility that antonymy
affects earlier processing stages.

Experiment 1

Method

Materials and design. The critical contradictory
sentences were constructed using six pairs of gender
antonyms: girl-boy, woman-man, female-male, sister-
brother, queen-king, and aunt-uncle. Semantically,
the first three pairs of words are relatively simple
concrete nouns, whereas the last three pairs are
more complex relational terms. In selecting these
pairs, terms with suffixes indicating gender (e.g.,
audress) were avoided. Using both orders of the
words as subject and predicate terms, 12 sentences
with direct antonyms were constructed. In 6 of
these, the subject and predicate were concrete
nouns, and in 6 they were relational nouns. For each
of these sentences, a matching sentence with in-
direct antonyms was created by replacing the
predicate term with a member of a different antonym
pair. Finally, all items were used with both of the
quantifiers all and some, for a total of 48 critical
false sentences. These items are listed in Table 1.

To verify that the pairs of direct antonyms were
more semantically similar than the pairs of indirect
antonyms, normative ratings were obtained from
two independent groups of University of Michigan
undergraduates. The two groups completed different
similarity questionnaires, both of which listed the
24 word pairs in random order. McCloskey and
Glucksberg (1979) have criticized the similarity
ratings used by Holyoak and Glass (1975), which
were ratings of "semantic association." Accordingly,
the first group of 23 subjects were given the in-
structions suggested by McCloskey and Glucksberg,
rating each pair (on a 1–7 scale) as to "how related
in meaning the two words are." The instructions
pointed out that one pair of words can be more
highly "associated" than another, yet less closely
related in meaning (e.g., cup–saucer vs. cup–mug).
The second group of 21 subjects rated each pair (on
a 0–10 scale) as to "how easy it would be for the
first concept to become the second." In both
questionnaires, lower numbers indicated higher
similarity.

Both sets of questionnaire results were unam-
biguous. The mean ratings for direct versus in-
direct antonyms were 3.37 versus 5.35 on the "re-
latedness" questionnaire, $F(1, 22) = 9.00, \ p < .01,$

1 This wording was suggested in a working paper
by E. Shoben, L. Rips, and E. Smith.
and 4.22 versus 7.34 on the "ease of change" questionnaire, $F(1, 20) = 12.13, p < .001$. In each set of ratings, all 12 of the direct antonym pairs in Table 1 were rated as more similar than the corresponding indirect antonym pairs. Accordingly, the feature comparison models of Smith et al. (1974) and McCloskey and Glucksberg (1979) predict that sentences containing direct antonyms will be rejected relatively slowly in the present study.

**Apparatus and procedure.** The experiment was controlled by a Digital Equipment Corporation PDP 11/40 computer. The sentences were displayed one at a time on a Panasonic monitor by a Princeton Electronics Products 400 scan converter. The 48 critical sentences were intermixed with other sentences quantified by *all* and *some*, some of which constituted a separate experiment not directly of concern here. The total item set consisted of 432 sentences, with equal numbers of true and false items for each quantifier. Each critical false sentence was matched with a filler true sentence with the same subject word, so that gender words did not always signal a "false" response. The additional sentences were of various types, including sentences with adjective predicates (e.g., *Some flowers are yellow*) and noncontradictory false sentences (e.g., *All birds are robins*). The test sentences were intermixed with other sentence types to prevent subjects from using special strategies to evaluate the critical items, which are all semantically very similar.

Twenty-one Rutgers undergraduates served as subjects. On each trial, the sentence was presented on one line on the monitor and remained on until the subject responded. The intertrial interval was 1,200 msec. The sentences were randomly ordered with the restriction that two critical sentences were never presented consecutively. The subjects were divided into three groups that were given different presentation orders. There were three brief rest breaks during the session, each lasting about 1 min. The first item after a break was never one of the critical sentences. Thirty different practice sentences were presented prior to the test items. A session was completed in 35-40 min.

**Results**

The mean correct reaction times (RT) and error rates for each type of contradictory sentence are presented in Table 2. Since the sentences used nearly exhausted the relevant item population, with the exception of a few very similar sentences, the items factor was treated as a fixed effect. The sentences based on direct antonyms were rejected 80 msec more quickly than those based on indirect antonyms, $F(1, 20) = 24.63, p < .001$ ($MS_e = 65,561$). This difference did not vary significantly as a function of any other variable. Collapsing over all other variables (including order of the male and female terms), the difference favored the direct antonyms for five out of six item comparisons, with the single reversal only 1 msec in magnitude. The magnitude of the difference ranged from −1 msec to 215 msec across items. This result therefore supports the hypothesis that sentences containing direct antonyms can be rejected relatively quickly on the basis of a specific contradiction between the meanings of the subject and the predicate words.

The only other significant effect was a 67-msec advantage for sentences quantified by *all* as opposed to *some*, $F(1, 20) = 17.76, p < .01$ ($MS_e = 63,253$). This difference indicates that the criteria used to evaluate sentences differ for the two quantifiers. Logically, a sentence quantified by *all* can be rejected on the basis of a single counterexample, whereas a sentence quantified by *some* can be rejected only on the basis of evidence that no positive instance can exist. One explanation for the faster rejection times for sentences with *all* is that counterexamples can sometimes be retrieved more quickly than direct contradictions. A slightly different possibility is that direct contradictions of the sort provided by the antonyms used in the present study take time to evaluate. For example, some very
similar sentences based on pairs of opposites are not actually contradictory (e.g., Some sons are fathers). If direct contradictions require time to be evaluated, the availability of additional falsifying information, such as counterexamples, may account for the faster false decisions obtained for sentences quantified by all rather than some.

Experiment 2

The results of Experiment 1 demonstrate that direct antonyms, despite their high degree of semantic relatedness, produce relatively rapid false decisions. At least on the surface, this finding is counter to the prediction of the feature comparison models of Smith et al. (1974) and McCloskey and Glucksberg (1979).

However, as we noted earlier, all of the hypotheses considered so far deal with effects at the decision stage of the sentence verification process. It is possible, however, that antonymy facilitates some earlier processing stage. Meyer and Schvaneveldt (1971, 1976a, 1976b) have demonstrated that when subjects are asked to judge whether a string of letters is a word, “word” decisions are made more quickly for pairs of related words (e.g., doctor–nurse) than for pairs of unrelated words (e.g., doctor–tree). This semantic facilitation effect in the lexical decision task has been explained in terms of activation spreading along associative pathways from the first concept to the second. That is, activation of the memory representation of one word will tend to quickly activate the representations of semantically related words. Although this explanation emphasizes lexical retrieval as the locus of the facilitation effect, there is some evidence that part of the effect may lie at an earlier encoding stage (Meyer, Schvaneveldt, & Ruddy, 1975).

The semantic facilitation effect found in the lexical decision task raises the issue of whether the positive effect of antonymy on falsification latency, obtained in Experiment 1, may be wholly or partly attributable to processes that occur prior to comparison of the S and P terms. In order to address this issue, one would ideally like to examine RT differences among critical items in a task that involves the same mental processes as sentence verification, except for the actual comparison process. A major problem in realizing this ideal is that candidate tasks are virtually certain to add processes not required in the sentence verification task. In addition, it is likely that the memory processes involved in the retrieval and comparison of word meanings operate at least partially in parallel. Nonetheless, it seemed reasonable to hope that an investigation of the role of antonymy relations in a different judgment task might constrain the interpretation of the effect found in Experiment 1. Accordingly, Experiment 2 used the critical gender antonyms in a lexical decision task.

Although our central concern involved the mechanism that might underlie a positive effect of semantic relatedness on falsification latency, similar questions arise in connection with the verification of true sentences. Numerous studies have demonstrated that for sentences asserting an instance–category relation, decision time is faster when the instance is a frequently produced or “typical” category member (e.g., All robins are birds) than when the instance is a less frequently produced or atypical member (e.g., All chickens are birds; Rips et al., 1973; Rosch, 1973; Wilkins, 1971). Although a variety of alternative models have been proposed to account for the positive effect of semantic relatedness on “true” RT, all share the assumption that the locus of the effect is the retrieval and/or evaluation of the relation between the S and P terms. However, the Meyer and Schvaneveldt studies raise the possibility that the locus is actually some earlier processing stage. Although this possibility has been recognized (Meyer & Schvaneveldt, 1976b), it has not been explored experimentally. Accordingly, Experiment 2 also investigated the effect of instance–category relatedness in a lexical decision task.

Method

Subjects were presented with pairs of letter strings and asked to decide as quickly as possible...
if both strings were words, or if one or both was a nonword. This “dual string” version of the lexical decision task, introduced by Meyer and Schvaneveldt (1971), is similar to the sentence verification task in that the two critical words are presented simultaneously.

Materials and design. The total item set comprised 284 pairs of letter strings. Of these, 132 consisted of two words, 132 of a word and a nonword, and 20 of two nonwords. The word-word pairs formed two sets of items, both listed in Table 3. The first set of 36 was based on the antonym pairs used in Experiment 1. These included 6 direct antonym pairs, 6 indirect antonym pairs, and 6 unrelated pairs formed by replacing one of the gender terms with a semantically dissimilar word. These replacement words were selected so that the antonym pairs and unrelated pairs were matched closely for word length and frequency (logarithm of the sum of the singular and plural forms in Kucera & Francis, 1967). (These factors were automatically equated for the direct and indirect antonym pairs, since these were based on exactly the same words.) All pairs were presented in both possible orders.

The second set of word-word pairs consisted of 96 category-exemplar pairs. These were based on 12 categories included in the Battig and Montague (1969) norms of instance generation frequency. Each category name (e.g., *bird*) was paired with two instances, one produced frequently (e.g., *robin*) and one produced infrequently (e.g., *chicken*), and with two noninstances, one produced frequently (e.g., *hammer*) and one infrequently (e.g., *axe*), as instances of another category (in this case, *tool*). All categories and exemplars were used equally often in both category-instance and category-noninstance pairs. The high frequency instances were produced by at least 209 respondents in the Battig and Montague sample \( (m = 334) \), whereas the low frequency instances were produced by no more than 97 respondents \( (m = 37.4) \). Mean word frequency, based on the logarithm of the sum of the singular and plural forms in the Kucera and Francis (1967) count, was virtually identical for the high frequency and low frequency instances. All pairs were presented in both possible orders.

Categories were matched so that animate categories were paired with exemplars of inanimate categories (and vice versa) to form the category-noninstance pairs. This procedure minimized the semantic relatedness of the category-noninstance pairs. The division of the category-instance pairs into two levels of instance production frequency

<table>
<thead>
<tr>
<th>Antonymy set</th>
<th>Direct antonyms</th>
<th>Indirect antonyms</th>
<th>Unrelated</th>
</tr>
</thead>
<tbody>
<tr>
<td>girl-boy</td>
<td>girl-brother</td>
<td></td>
<td>girl-door</td>
</tr>
<tr>
<td>woman-man</td>
<td>woman-king</td>
<td>female-uncle</td>
<td>female-snow</td>
</tr>
<tr>
<td>female-male</td>
<td>sister-boy</td>
<td>queen-king</td>
<td>queen-money</td>
</tr>
<tr>
<td>sister-brother</td>
<td>aunth-uncle</td>
<td></td>
<td>auntpoem</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Category-Exemplar set</th>
<th>Instances High frequency</th>
<th>Low frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>bird-robin</td>
<td>bird-chicken</td>
<td></td>
</tr>
<tr>
<td>tool-hammer</td>
<td>tool-axe</td>
<td></td>
</tr>
<tr>
<td>insect-mosquito</td>
<td>insect-butterfly</td>
<td></td>
</tr>
<tr>
<td>weapon-knife</td>
<td>weapon-arrow</td>
<td></td>
</tr>
<tr>
<td>fruit-apple</td>
<td>fruit-lime</td>
<td></td>
</tr>
<tr>
<td>sport-football</td>
<td>sport-boating</td>
<td></td>
</tr>
<tr>
<td>vegetable-carrot</td>
<td>vegetable-onion</td>
<td></td>
</tr>
<tr>
<td>metal-iron</td>
<td>metal-brass</td>
<td></td>
</tr>
<tr>
<td>fish-trout</td>
<td>fish-eel</td>
<td></td>
</tr>
<tr>
<td>clothing-shirt</td>
<td>clothing-belt</td>
<td></td>
</tr>
<tr>
<td>flower-tulip</td>
<td>flower-poppy</td>
<td></td>
</tr>
<tr>
<td>vehicle-bus</td>
<td>vehicle-canoe</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Noninstances High frequency</th>
<th>Low frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>tool-robin</td>
<td>tool-chicken</td>
</tr>
<tr>
<td>bird-hammer</td>
<td>bird-axe</td>
</tr>
<tr>
<td>weapon-mosquito</td>
<td>weapon-butterfly</td>
</tr>
<tr>
<td>insect-knife</td>
<td>insect-arrow</td>
</tr>
<tr>
<td>sport-apple</td>
<td>sport-lime</td>
</tr>
<tr>
<td>fruit-football</td>
<td>fruit-boating</td>
</tr>
<tr>
<td>metal-carrot</td>
<td>metal-onion</td>
</tr>
<tr>
<td>vegetable-iron</td>
<td>vegetable-brass</td>
</tr>
<tr>
<td>clothing-trout</td>
<td>clothing-eel</td>
</tr>
<tr>
<td>fish-shirt</td>
<td>fish-belt</td>
</tr>
<tr>
<td>vehicle-tulip</td>
<td>vehicle-poppy</td>
</tr>
<tr>
<td>flower-bus</td>
<td>flower-canoe</td>
</tr>
</tbody>
</table>
corresponds to the variations in relatedness examined in many semantic verification experiments (e.g., Wilkins, 1971).

To form the 132 word–nonword pairs, one word was dropped from each of the word–word pairs in both the antonymy and category–exemplar sets. For each of the words to be replaced, a pronounceable nonword (e.g., skart) of equal length in letters and syllables was constructed. These nonwords were then randomly paired with the remaining words from the word–word pairs. Every word used in a word–word pair occurred at least once in a word–nonword pair. (Words occurred in more than one word–word pair.) The nonwords occurred equally often as the left and as the right pair member. Finally, 20 pairs containing two pronounceable nonwords were added to the item set.

**Apparatus and procedure.** Up to four subjects were tested simultaneously in semi-isolated booths containing TV screens controlled by an IBM 1800 computer. Each trial began with presentation of a fixation cross that remained on for 1 sec and then was replaced by the pair of letter strings. The strings appeared side by side on the screen. The leftmost letter of the left string was placed at the position previously occupied by the fixation cross, and the two strings were separated by one character space. The fixation cross was positioned in this way so as to encourage left to right processing, as would presumably occur in a sentence verification task such as that used in Experiment 1.

Subjects were instructed to press the right response button if both strings in a pair were words and the left button if either or both strings were nonwords. An error message was displayed for 1 sec if the subject pressed the wrong button; the next trial then began automatically. The 284 test trials were presented in a different random order for each subject. Subjects were given a short break and informed of their cumulative error percentage after each block of 71 trials. The test trials were preceded by 20 practice pairs. Eighteen University of Michigan undergraduates served as paid subjects.

**Results**

The mean correct RTs and error rates for all conditions are presented in Table 4. The three levels of semantic relatedness represented in the antonymy set produced significant variations in decision time. The pairs of direct antonyms were classified as words 41 msec more quickly than were the pairs of indirect antonyms, which in turn were classified 65 msec more quickly than were the unrelated pairs. Using planned comparisons, the contrast between the direct and indirect antonyms was significant, \( t(34) = 2.14, p < .05 \), as was the contrast between all antonym pairs and the unrelated pairs, \( t(34) = 5.15, p < .001 \) (\( MS_e = 6622 \)). As in Experiment 1, the concrete and relational antonym pairs did not differ significantly in mean RT. Collapsing over order of the male and female terms, the direct antonym pairs were classified more quickly than were the indirect antonym pairs in four of the six item comparisons, with the two reversals 21 and 31 msec in magnitude. The magnitude of the effect ranged from −31 msec to 166 msec across the six item comparisons.

The RT analyses for the category–exemplar set were performed using both subjects and categories as random effects, and minimum quasi-\( F \) ratios were calculated (Clark, 1973). The only significant effect was a 69-msec advantage for the category–instance over the category–noninstance pairs, \( F_{min}(1, 22) = 11.5, p < .01 \) (\( MS_e = 5,366 \) for subjects, 6,590 for items). The small trend toward faster decisions for pairs containing high frequency rather than low frequency instances (24 msec for category–instance pairs) did not approach significance in either the subject or the category analysis, either as an interaction or as a main effect, \( F_{min}(1, 24) = .92 \) (\( MS_e = 5,391 \) for subjects, 3,840 for items). There was also a nonsignificant 27-msec trend toward faster responses when the exemplar rather than the category occurred on the left in the pair, \( F_{min}(1, 23) = 2.23, p > .10 \) (\( MS_e = 4,149 \) for subjects, 4,659 for items).

<table>
<thead>
<tr>
<th>Stimulus pairs</th>
<th>RT</th>
<th>% error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct antonyms</td>
<td>684</td>
<td>.9</td>
</tr>
<tr>
<td>Indirect antonyms</td>
<td>725</td>
<td>.5</td>
</tr>
<tr>
<td>Unrelated</td>
<td>790</td>
<td>1.9</td>
</tr>
<tr>
<td>Category–instance (H)</td>
<td>730</td>
<td>1.4</td>
</tr>
<tr>
<td>Category–instance (L)</td>
<td>754</td>
<td>1.6</td>
</tr>
<tr>
<td>Category–noninstance (H)</td>
<td>805</td>
<td>2.8</td>
</tr>
<tr>
<td>Category–noninstance (L)</td>
<td>817</td>
<td>2.6</td>
</tr>
<tr>
<td>Nonword–word</td>
<td>794</td>
<td>3.0</td>
</tr>
<tr>
<td>Word–nonword</td>
<td>930</td>
<td>7.7</td>
</tr>
<tr>
<td>Nonword–nonword</td>
<td>787</td>
<td>2.0</td>
</tr>
</tbody>
</table>

**Note.** RT = reaction time in msec; H = high frequency; L = low frequency.

The mean correct RT and error rate for lexical decisions in Experiment 2.
Pairs containing a nonword were classified 136 msec more quickly when the nonword occurred on the left rather than on the right, \( F_{\text{min}}'(1, 81) = 46.5, p < .001 \) (\( MS_e = 1,450 \) for subjects, 6,964 for items). In addition, the word–nonword condition produced the only substantial error percentage in the experiment (7.7%). These results confirm that subjects processed the pairs from left to right. Also, mean RTs were virtually identical for the nonword–word and nonword–nonword conditions, suggesting that decisions were made sequentially for the left and then the right letter strings.

General Discussion

Decision times for items involving direct (e.g., boy–girl) and indirect (e.g., boy–sister) gender antonyms were compared in two speeded judgment tasks. In Experiment 1, false sentences involving direct antonyms were rejected 80 msec more quickly than false sentences involving indirect antonyms. This finding supports previous work indicating that a high degree of semantic relatedness does not inevitably lead to relatively slow false decisions in a sentence verification task (Glass et al., 1974; Holyoak & Glass, 1975; Lorch, 1978).

In Experiment 2, using a dual-string lexical decision task, a 41-msec advantage was obtained for the direct versus indirect antonym pairs (although both types of pairs were classified more quickly than pairs of unrelated words). In view of the latter result, how should the positive effect of semantic relatedness on falsification latency (Experiment 1) be interpreted? At least nominally, the lexical decision task involves decisions about individual words rather than about the relation between words. The results of Experiment 2 therefore suggest that at least part of the antonymy effect may involve processes that precede the evaluation of the relation between the S and P concepts (e.g., initial encoding of the S and P words or retrieval of their meanings). The fact that the effect tended to be larger in the verification task than in the lexical decision task (80 vs. 41 msec) might suggest that antonymy relations facilitate evaluation in addition to earlier processes. However, this difference in the size of the facilitation effect might simply reflect the fact that RTs were overall longer and more variable in the verification task than in the lexical decision task.

One problem in interpreting our results is that it is not entirely clear to what extent the processes involved in the two tasks overlap. For example, it is plausible to suppose that the facilitation of lexical decisions for a pair of direct antonyms is a result of activation spreading along associative pathways linking the two concepts. For such pairs, retrieval of the individual word meanings may in fact depend on retrieval of the relation between the concepts—the same information that would be used to evaluate the corresponding sentence. If this is so, then the retrieval and evaluation processes may be sufficiently interdependent as to be extremely difficult to separate empirically.

A further complication arises when the present results are related to those of Lorch (1978). We noted at the outset that direct antonyms are disjoint members of an immediate superordinate category and that direct antonyms share more properties than do indirect antonyms. Lorch demonstrated that although a highly available superordinate facilitates false decisions, a high degree of property overlap slows them down. It is therefore possible that two opposing factors influence the relative speed of the evaluation process for sentences with direct versus indirect antonyms. The facilitation produced by a readily available superordinate contradiction may be partially offset by the negative effect of feature overlap; as a result, the obtained 80-msec advantage for sentences with direct antonyms may be an underestimate of the facilitatory effect of a superordinate contradiction on falsification time. Since the size of this effect nonetheless tended to be larger than the comparable effect obtained in the lexical decision task (in which direct antonymy presumably has an
entirely positive effect on judgment time),
direct antonymy may indeed facilitate the
decision stage in sentence verification, as
well as earlier processes.

Although the basis for the antonymy
effect on falsification remains uncertain,
the results of Experiment 2 appear more
conclusive in assessing the locus of the
“typicality effect” commonly obtained for
“true” verification decisions. In the lexical
decision task, only a small nonsignificant
trend was found toward faster word judg-
ments for pairs consisting of a category and
a frequently produced instance, as opposed
to a category and an infrequently produced
instance. It therefore seems that the
reliable differences in verification times
found for such pairs in previous studies
(Rips et al., 1973; Rosch, 1973; Wilkins,
1971) arise during the process of evaluating
the category–instance relations.

In assessing the effects of antonymy on
semantic judgments, our conclusions are
necessarily limited by the restricted set of
items used in the present study. It seems
clear, however, that a high degree of
semantic relatedness does not necessarily
lead to relatively slow rejection of false
sentences. Further research will be re-
quired to clarify the locus of the positive
effect of specific meaning relations, such as
antonymy, on the time required to make
semantic decisions.

References
Battig, W. F., & Montague, W. E. Category norms
for verbal items in 56 categories: A replication
and extension of the Connecticut category norms.
Journal of Experimental Psychology Monograph,
1969, 80 (3, Pt. 2).
Clark, H. H. The language-as-fixed-effect fallacy:
A critique of language statistics in psychological
research. Journal of Verbal Learning and Verbal
Collins, A. M., & Loftus, E. F. A spreading-activa-
tion theory of semantic processing. Psychological
Collins, A. M., & Quillian, M. R. Retrieval time
from semantic memory. Journal of Verbal Learn-
Glass, A. L., & Holyoak, K. J. Alternative con-
ceptions of semantic memory. Cognition, 1975,
J, 313–339.
Glass, A. L., Holyoak, K. J., & Kossan, N. E.
Children’s ability to detect semantic contradic-
Glass, A. L., Holyoak, K. J., & O’Dell, C. Produc-
tion frequency and the verification of quantified
statements. Journal of Verbal Learning and Verbal
Holyoak, K. J., & Glass, A. L. The role of contra-
dictions and counterexamples in the rejection of
false sentences. Journal of Verbal Learning and Verbal
Behavior, 1975, 14, 215–239.
Kučera, H., & Francis, W. N. Computational analysis
of present-day American English. Providence, R.I.: Brown
Lorch, R. F., Jr. The role of two types of semantic
information in the processing of false sentences.
Journal of Verbal Learning and Verbal Behavior,
1978, 17, 523–537.
McCloskey, M., & Glucksberg, S. Decision processes
in verifying category membership statements:
Implications for models of semantic memory.
Meyer, D. E. On the representation and retrieval
of stored semantic information. Cognitive Psychol-
Meyer, D. E., & Schvaneveldt, R. W. Facilitation
in recognizing pairs of words: Evidence of a
dependence between retrieval operations. Journal
Meyer, D. E., & Schvaneveldt, R. W. Meaning,
memory structure, and mental processes. Science,
1976, 192, 27–33. (a)
Meyer, D. E., & Schvaneveldt, R. W. Meaning,
memory structure, and mental processes. In C. N.
Cofer (Ed.), The structure of human memory.
San Francisco: Freeman, 1976. (b)
Meyer, D. E., Schvaneveldt, R. W., & Ruddy,
M. G. Loci of contextual effects on visual word
recognition. In P. M. A. Rabbit & S. Dornic
(Eds.), Attention and performance V. London:
distance and the verification of semantic rela-
tions. Journal of Verbal Learning and Verbal
Rosch, E. On the internal structure of perceptual
and semantic categories. In T. E. Moore (Ed.),
Cognitive development and the acquisition of
Schaeffer, E., & Wallace, R. The comparison of
word meanings. Journal of Experimental Psy-
chology, 1970, 86, 144–152.
Smith, E. E. Theories of semantic memory. In
W. K. Estes (Ed.), Handbook of learning and
cognitive processes (Vol. 6). Hillsdale, N.J.: Eri-
Smith, E. E., Shoben, E. J., & Rips, L. J. Structure
and process in semantic memory: A featural
model for semantic decisions. Psychological Review,
Wilkins, A. T. Conjoint frequency, category size,
and categorization time. Journal of Verbal
Learning and Verbal Behavior, 1971, 10, 382–385.

Received April 5, 1979
Revision received June 26, 1979