

# Semantic memory and the granularity of semantic relations: Evidence from speed-accuracy decomposition

JOHN KOUNIOS, ELIZABETH C. MONTGOMERY, and RODERICK W. SMITH  
*Tufts University, Medford, Massachusetts*

In the present study we examined whether semantic relations are atomistic unitary associations, or are complex concepts consisting of a number of relational elements. The complexity of the ownership relation was assessed by combining a relation verification task ("Many people own [cars/comets]") with the speed-accuracy decomposition procedure (Meyer, Irwin, Osman, & Kounios, 1988). The latter permits one to determine whether subjects achieve their final state of response accuracy in a single, discrete all-or-none transition, or whether the relevant processes yield partial information representing intermediate states of knowledge. The rationale was that the retrieval of a unitary relational link from a classical associative network should be an all-or-none affair. In contrast, a set of relational elements need not be processed as a unitary bundle, thereby allowing partial response-information states. In two experiments, we found evidence of such partial information (i.e., sensitivity in units of  $d'$ ), lending support to the notion that relations are complex. Furthermore, the results suggest that the accumulation of guessing sensitivity was linear over time, weighing against alternate theoretical interpretations.

One of the most persistent debates over more than two decades of semantic memory research has focused on the *structure versus process* question (Chang, 1986; Hollan, 1975; Kounios, 1993, 1994; Kounios & Holcomb, 1992, 1994; Kounios, Osman, & Meyer, 1987; Rips, E. E. Smith, & Shoben, 1975; E. E. Smith, Shoben, & Rips, 1974): Is semantic memory primarily a *storage* system or a *processing* system? For example, in verifying the truth of sentences such as "All dogs are animals" or "Many people are green," do people *search* a knowledge structure for a prestored relation, or do they access representations of the concepts corresponding to the subject and predicate terms of a sentence and then *compute* the truth value of the relation between them?

This question is important, because it speaks not only to the nature of semantic representations, but also to the processes that operate on them (Kounios et al., 1987; Rips et al., 1975). Accordingly, theorists have espoused a variety of views on this topic (see Chang, 1986). For example, some models posit featural representation of concepts, implying sentence verification by means of a computational comparison mechanism (e.g., McCloskey & Glucksberg, 1979; Ratcliff & McKoon, 1982; E. E.

Smith et al., 1974), while other models posit concept representation in classical associative networks, implying verification by search and retrieval of preestablished links connecting the subject and predicate concepts (e.g., Collins & Loftus, 1975; Collins & Quillian, 1969; Glass & Holyoak, 1975).

## Relations in Semantic Memory

As a part of this debate, Glass and Holyoak (1975) pointed out one limitation of feature-comparison approaches to semantic verification. Such models determine the truth value of a sentence by computing a measure of similarity between the subject and predicate concepts (e.g., E. E. Smith et al., 1974). This general approach does a reasonably good job of accounting for the database of empirical results from sentence-verification experiments (Chang, 1986). However, almost all of these studies used sentences involving class-inclusion relations (e.g., "All dogs are animals"). A similarity-based feature-comparison strategy would seem to be a plausible method for verifying such sentences. However, as Glass and Holyoak (1975) pointed out, a similarity-based feature-comparison mechanism would be hard-pressed to account for the verification of many types of sentences embodying relations other than similarity or class inclusion. One particularly salient example they give is the relation of possession or ownership. For instance, though "Many people own houses" is a true statement, this truth has nothing to do with the similarity of people to houses. Glass and Holyoak argued that the ability to correctly verify such sentences suggests a prestored association between the subject and predicate

Preparation of this article was supported by NIMH Grant CEP 1 R29 MH45447-01A3 to the first author. The article is based, in part, on a master's thesis by the second author. Thanks go to Robert Lorch and Roger Chaffin for comments on a previous version of the manuscript, and to Phil Holcomb and Mark Reinitz for helpful discussions about this research. Correspondence may be addressed to J. Kounios, Department of Psychology, Tufts University, Medford, MA 02155 (e-mail: jkounios@pearl.tufts.edu).

concepts. This can be easily accomplished by a classical associative network.

But what form might such associative relations take? Chaffin and Herrmann (1988) describe two competing accounts. The "traditional" approach conceives of relations as unitary links between concept nodes in a classical network. Of course, this implies a very large number of different types of atomistic links (viz., one type of link per type of relation). In contrast, they describe their *relation element theory*, which views relations themselves as complex concepts decomposable into featural elements. In this view, each type of relation is represented by a different subset of the total pool of relational features or elements. This notion is important to the structure-process debate, because unitary, atomistic relations are, by definition, prestored, whereas the semantic decomposition of relations into elements permits a range of relational computations to be performed (although such complex relations might still be prestored and retrieved as unitized bundles of elements).

### Speed–Accuracy Decomposition

One strategy for ascertaining the complexity or granularity of semantic relations is to investigate properties of the processes that operate on them. This was the approach taken by Kounios et al. (1987) in a study of the verification of universally quantified class-inclusion sentences (e.g., "All dogs are animals"). They reasoned that if such sentences are verified by comparing sets of features corresponding to the subject and predicate concepts, then each comparison of a pair of features should yield some small increment in partial information about the truth or falsity of the sentence. On the other hand, if a sentence is verified by (successfully or unsuccessfully) searching for an appropriate prestored (unitary) link between the subject and predicate concepts, then at some point during each trial there should be a discrete transition in response information from a state of total ignorance (i.e., zero sensitivity) to the final state of relatively complete knowledge of the sentence's truth value.

Kounios et al. (1987) attacked this problem by means of *speed–accuracy decomposition* (SAD), a technique specially designed to detect and measure available partial response information (Meyer, Irwin, Osman, & Kounios, 1988). This technique utilizes two types of randomly interleaved trials: "regular" and "response signal." Regular trials are the type of trials used in conventional reaction time (RT) experiments: The subject responds to a test stimulus after taking the minimum amount of time necessary to ensure a high probability of an accurate response. The subject is trained to adopt this prioritization of goals at the beginning of each trial, with response speed being secondary to response accuracy. On response-signal trials, however, if the subject has not already responded by the time a response signal (e.g., a tone or a visual pattern) is encountered, then the presentation of the signal indicates that these priorities

should be reversed and an immediate "best guess" response should be made.

Therefore, the experimenter has two types of data. The regular-trial data give the accuracy and distribution of finishing times for the "normal" processing of the test stimulus, as would occur in a conventional RT experiment. The period of time elapsing between the test stimulus and the response signal (i.e., the signal lag) can be manipulated across trials yielding the other type of data, a conventional speed–accuracy tradeoff (SAT) function. This SAT function is presumed to consist of a mixture of two types of trials: those on which the subject was forced by the response signal to respond before ready (based on any available partial information), and those trials on which the subject responded not because of the response signal, but because he or she had finished normal processing of the test stimulus and would have responded at that point in time, even if no response signal had been presented on that trial. By means of a race model (the *parallel sophisticated-guessing model*; see Meyer et al., 1988), the accuracy and distribution of response times for the regular trials can be extracted from the accuracy and distribution of response times for the response-signal trials, yielding estimates of the residual accuracy and the RT distribution for *guesses*—that is, responses based on incomplete processing.<sup>1</sup> The critical equation for decomposing the signal-trial accuracies is

$$P(\text{cor}|\text{lag } s) = P(t_n \leq t_g)P(\text{cor}|t_n \leq t_g) \\ + P(t_g < t_n)P(\text{cor}|t_g < t_n),$$

where  $P(\text{cor}|\text{lag } s)$  is the probability of a correct response on response-signal trials with a particular signal lag, where  $P(\text{cor}|t_n \leq t_g)$  is the probability of a correct response on trials on which the normal processes won the race with guessing processes incited by a signal (of the same lag) and therefore determined the response, and where  $P(\text{cor}|t_g < t_n)$  is the probability of a correct response on trials on which the guessing process won the race against the normal processes and determined the response.  $P(t_n \leq t_g)$  is a weighting parameter giving the probability that the normal processes won the race and determined the response;  $P(t_g < t_n)$  is a weighting parameter giving the probability that the guessing process won the race. The above equation can be rearranged to solve for  $P(\text{cor}|t_g < t_n)$ —that is, the accuracy of the guessing process. This parameter reflects available partial response information on signal trials using a specific lag. For example, if this estimate of guessing accuracy is significantly greater than 50%, one can infer the existence of partial information. (For a detailed explication of SAD, see Meyer et al., 1988).<sup>2</sup>

### Overview

The present study takes up the suggestion of Glass and Holyoak (1975) that semantic memory researchers should investigate the verification of sentences em-

bodying relations other than class inclusion as a step toward clarifying the structure-process question. Following the logic of Kounios et al. (1987), we applied SAD to a sentence-verification task employing statements of possession (e.g., "Many people own galaxies"). The absence of partial response information during this task would support the classical associative-network view that such relations are prestored in the form of unitary ownership links between the subject and predicate concepts. The presence of partial information would support the notion that the ownership relation is a relatively complex collection of elements used in the computational verification of such sentences.

Second, we were interested in the imageability of the stimuli, because a verification strategy based on image retrieval could complicate the isolation of partial information generated solely by the processing of abstract (e.g., propositional) semantic information. Accordingly, an effort was made to ascertain the role of imageability in ownership statement verification. Specifically, high- as well as low-imageability stimuli were included. If high- and low-imageability sentences yield different partial-information results, then the results from the low-imageability sentences might better reflect partial outputs from the processing of abstract semantic information. If high- and low-imageability stimuli yield the same partial-information results, then imageability is not affecting the processes in question, thereby indicating that the information used by subjects is abstract rather than concrete.

### Stimuli

**Stimulus generation.** The first step was to collect appropriate stimuli. Since there were no available lists of such stimuli, this was accomplished by means of questionnaires and normative ratings available from the literature.

One hundred and eighty Tufts University students from an introductory psychology class participated as volunteers in a group setting. Each subject was randomly assigned to one of two experimental conditions, resulting in 90 subjects per condition. Each subject was given a "blue book" and a set of instructions corresponding to their experimental condition. In the instructions they were asked to write down "all the single-word nouns that you can think of that refer to objects that *many* people own" (first condition), or "that *few, if any*, people own" (second condition) within 8 min. These latter words were collected to form the basis for false sentences in the verification task.

The word-generation task yielded over 1,200 unique items. Many of these items were discarded because they were multiple-word items, nonnouns, nonwords, or were frivolous responses.

**Questionnaires.** A questionnaire was administered to a separate group of 100 Tufts undergraduates in order to assess the imageability and truth value of sentences of the form "Many people own [noun]." The subjects were instructed to rate the imageability of the entire sentence

on a 7-point scale. They were also instructed to circle the appropriate letter (T or F) to indicate whether each sentence was generally true or generally false. There were 10 such questionnaires, each containing a subset of the words yielded by the initial generation task.

Another questionnaire (modeled after McCloskey, 1980) assessed the familiarity of the individual predicate nouns on a 7-point scale. Ten such questionnaires were administered to a different group of 100 Tufts University undergraduates, each containing a subset of the items.

A preliminary list of high- and low-imageability true and false sentences was drawn from this pool. High-imageability sentences were drawn from those given an average imageability rating greater than 4; low-imageability sentences were drawn from those given an average rating less than 4.

True sentences of high imageability were easily compiled. All 52 of the selected sentences in this category were rated as true by 100% of the raters; all of these sentences received an imageability rating of 7 from all the subjects. False sentences of low imageability were also easily compiled. The 52 sentences in this category were rated as false by 100% of the raters; the imageability values of these sentences were more variable, spanning the allowed range of 1–3.

On the other hand, high-imageability false sentences were more varied in their characteristics. At least 85% of the subjects rated each of these 52 sentences as false; the imageability ratings ranged from 5 to 7. True sentences of low imageability were the most difficult to compile. The initial word list contained only 9 words from which appropriate sentences could be constructed. The experimenters therefore generated an additional 43 words for this condition. Consequently, the familiarity and imageability questionnaires were repeated on two new groups of 50 Tufts undergraduates (per group) for the final list of 208 test words. There was one significant change in the imageability questionnaire. In the first questionnaire, the subjects were asked to assess the imageability of the entire sentence containing the critical predicate noun. It became obvious from discussions with the raters that many of them did not understand the distinction between rating the imageability of the whole sentence and rating the imageability of just the predicate noun. In the second questionnaire, therefore, they were asked to rate only the isolated nouns that would later be inserted into the test sentences as predicates.

Finally, there is an important point that should be made about the final stimulus list (given in the Appendix). Inspection of the words in each condition suggests semantic confounds that subjects might be able to use as a heuristic for verifying the sentences. For instance, the true/low-imageability condition contains a number of financial or business terms. The false/high-imageability condition contains a number of animals. Also, the false/low-imageability condition contains a number of astronomical and meteorological terms. Ideally, a stimulus set should contain no such confounds. However, there

were not enough stimuli available from which a totally confound-free list could be constructed. Fortunately, this proved not to be a problem, as the design of the experiment made it possible to determine whether subjects were using such confounds. This is because each confound occurs with respect to both the truth and imageability factors. Consequently, if subjects were to use such a heuristic, it would necessarily be reflected in differential guessing sensitivities for high- and low-imageability stimuli, since the semantic confounds were correlated with the imageability factor. As the results show, the imageability factor had no discernible effect on guessing sensitivity until quite late during the time course of verification. Therefore, the subjects did not base their quicker guesses on such heuristics.

Table 1 summarizes the lexical characteristics of the nouns used in the experiment. In brief, the true and false predicate nouns were highly similar. It is very unlikely that the subjects could detect and utilize the subtle difference in, say, word frequency (Francis & Kučera, 1982) during the course of one 30-min experimental session, especially given that other characteristics of the stimulus situation were much more salient. This conclusion is supported by the partial-information results described in the following text.

## EXPERIMENT 1

### Method

**Subjects.** Ten subjects from the Tufts University community served as subjects. They were paid \$7 per session, plus bonuses for excellent performance. Two of these subjects performed erratically on the final test session; their RTs on regular trials decreased dramatically, while their regular-trial accuracies plummeted. This apparently occurred as a result of the change from the practice stimuli to the test stimuli. Their data are not included in this report.

**Apparatus.** The experiment was controlled by IBM-compatible AT and XT microcomputers. The stimuli were presented on amber monochrome monitors. The subjects responded by pressing buttons on a buttonbox with their thumbs. The stimulus-response mapping was counterbalanced across subjects.

**Procedure.** Considerable training is necessary for subjects to learn to conform to the requirements of the SAD technique.<sup>3</sup> The main requirements are near asymptotic performance on the regular trials and very rapid responding after response signals. Since the stimulus set was not large enough to permit stimulus repetition without contaminating the results of interest, the subjects practiced with an alternate pool of stimuli. Eight stimulus sets of 208 sentences each were constructed for this purpose. Each sentence was of the form "Many people [verb] [noun]." The verb *own* was one of several verbs used in these practice sentences. The 208

critical predicate nouns were also used in the practice sentences, but never with the verb *own*. In this way, the subjects acquired experience verifying sentences and acquainted themselves with the verb *own* and the predicate nouns used in the final test session. The latter were included in the true as well as the false sentences during the practice sessions to discourage the subjects from associating specific predicates with a particular truth value.

The subjects were seated in front of the computer monitor and were told that they would see a series of sentences on the screen, each of the form "Many people [verb] [noun]." Each trial began with the sentence frame "Many people" presented as a warning signal 1,000 msec before the verb and the predicate noun. (During the final test session, the frame included the verb *own*.) The frame was presented on the 10th line of the monitor, horizontally adjusted so that the critical noun would begin on the 34th space from the left. The entire sentence remained on the screen until a response was made or the sentence was masked by the response signal. The intertrial interval was 2,000 msec. Each subject viewed the stimuli in a unique random order.

Each subject's initial practice session involved only regular trials (as in an RT experiment). The mixture of regular and signal trials characteristic of SAD was introduced in the second session. The subjects were given as many sessions of practice as necessary to achieve the criteria of at least 90% accuracy on regular trials with few, if any, "slow" trials (defined below) on at least two consecutive sessions. The subjects required about seven sessions of practice apiece. After each session, they were given feedback concerning their performance.

One third of the trials were of the regular variety. The subjects were told that accuracy was their highest priority on these trials and that speed was of secondary importance. They were told to respond in the minimum time necessary to ensure an accurate response. They were given error feedback on error trials.

Two response-signal lags were utilized (i.e., one third of the total number of trials per lag). "Early" signals were presented 100 msec after the onset of the predicate noun; "late" signals were presented 280 msec after predicate onset.<sup>4</sup> The subjects were told to respond immediately after a response signal with their best guess. Speed was stressed above accuracy, though accuracy was given a secondary importance.

The different types of trials were randomly intermixed. The subjects did not know whether a response signal would occur on a given trial until and unless it was actually presented.

The response signal was a series of pound signs (#####) masking the entire sentence. If a subject did not respond within 325 msec of the onset of this mask, a tone was presented to inform the subject that the response was too slow. The tone was terminated by the subject's response. The subject's goal was to respond quickly enough on signal trials so that the tone would rarely, if ever, occur. Error feedback was given on signal trials.

### Results

**Overview.** The three types of SAD data that are most important for the present purposes are (1) the regular-trial data (RTs and accuracies), (2) the derived guessing-time distributions, and (3) the derived guessing accuracies. After describing the regular-trial data, the derived guessing times were examined in order to determine whether there were any significant violations of the parallel sophisticated-guessing model upon which the SAD analyses were based. Then, the derived guessing accuracies were examined in order to determine whether there was any partial information. And last, the guessing sensitivities were compared with sensitivity on regular trials.

**Table 1**  
Lexical Characteristics of the Predicate Nouns

	True	False
Familiarity	5.95	5.92
Frequency	76.36	57.76
Length	7.38	6.93

Note—Familiarity values were obtained from a 7-point (1 = low; 7 = high) scale. Frequency values refer to occurrences per million words. Length values are in number of letters per word.

**Table 2**  
**Mean Reaction Times (RT, in Milliseconds) for**  
**Correct Responses, Standard Errors (SE), and**  
**Accuracies for Regular Trials From Experiment 1**

	True			False		
	RT	SE	Accuracy	RT	SE	Accuracy
Low imageability	957	142	.89	989	171	.96
High imageability	837	158	.99	1,115	232	.82

**Regular-trial data.** The mean correct RTs and accuracies for the regular-trial data are presented in Table 2. A  $2 \times 2$  (truth  $\times$  imageability) repeated measures analysis of variance (ANOVA) was performed on the RT data. There was a significant main effect of truth [ $F(1,7) = 6.09, p = .04$ ] and a significant interaction between truth and imageability [ $F(1,7) = 17.95, p < .01$ ]. An ANOVA on the accuracy data also yielded a significant main effect of truth [ $F(1,7) = 6.09, p = .04$ ] and a significant interaction between truth and imageability [ $F(1,7) = 16.01, p = .01$ ].

The regular-trial data exhibited a speed-accuracy tradeoff between the two conditions. Specifically, the true/low-imageability responses were 32 msec faster and 7% less accurate than the responses to false/low-imageability sentences, complicating any comparison between these two conditions.

**Guessing times.** The guessing times were derived by separately applying Equation 2 from Meyer et al. (1988; also given in note 1 of this report) to the data from each subject's test session.<sup>5</sup>

The parallel sophisticated-guessing model that constitutes the foundation on which SAD is based assumes that neither signal lag nor type of stimulus has an effect on the shape of the derived guessing-time distributions (or, practically speaking, an effect large enough to bias the derived guessing-time and guessing-accuracy parameters). The significance of these effects was tested by a multivariate analysis of variance (MANOVA) on the times corresponding to the .00, .10, .30, .70, .90, and 1.00 percentiles of the guessing-time distributions. Lag, imageability, and truth value were used as factors. In order to detect effects of lag on the shape (rather than on the latency) of the distributions, 180 msec was subtracted from the late-signal times before executing the MANOVA (thereby equating the early- and late-signal distributions by removing horizontal displacement due to the different signal onset times). This analysis (of

course) yielded a significant main effect for percentile [Wilks's lambda = .01;  $F(5,3) = 91.75, p < .01$ ]. None of the other main effects or interactions approached significance [ $.29 < F(5,3) < 2.51; .24 < p < .89$ ].

It is particularly interesting to note just how stable the medians of the derived guessing-time distributions were. For the early lag, the average guessing-time medians ranged from 408 to 414 msec across conditions, with a mean of 410 msec across conditions. The late-signal medians ranged from 585 to 591 msec, with a mean of 589 msec across conditions. Subtracting the average early-signal median from the average late-signal median yields 179 msec, almost perfectly recovering the 180-msec difference between the onsets of the signals.

It is also important to note the point in time at which guessing times fell in relation to the distributions of regular-trial RTs. For early-signal trials, the average median guessing time of 410 msec fell well before the average regular-trial minimum RT of 493 msec. In contrast, the average late-signal guessing median of 589 msec fell well within the regular-trial RT distribution; 14% of the regular-trial responses occurred before the guessing-time median. So, a failure to find partial information in the late-signal condition would imply a discrete transition from a state of no response information to the final state of relatively complete response information for (at least) the quickest 14% of regular-trial responses. One could therefore not explain any failure to find partial information for the late lag as resulting from too early a temporal placement of the late response signal.

The fact that early-signal guesses fell well before regular-trial responses indicates that the SAD analysis was unnecessary for these trials, because signal-trial responses on early-lag trials were not influenced by any contribution from normally completed processing. In effect, virtually all of these responses were guesses. In contrast, the late-lag guesses fell well within the regular-trial RT distribution. Therefore, many of the late-lag signal-trial responses were actually produced by the normal processes that produced the responses on the regular trials. Consequently, the failure to find significant differences between the early- and late-lag guessing-time distributions is particularly reassuring, because this involved a comparison of a derived (i.e., late-lag) guessing-time distribution to an observed (i.e., early-lag) guessing-time distribution.<sup>6</sup>

**Guessing accuracies.** The guessing accuracies were obtained by separately applying Equation 4 from Meyer

**Table 3**  
**Derived Guessing Accuracies With Standard Errors (SE) for Experiment 1**

	Lag 1				Lag 2			
	True		False		True		False	
	Accuracy	SE	Accuracy	SE	Accuracy	SE	Accuracy	SE
Low imageability	.68	.07	.55	.07	.65	.11	.73	.07
High imageability	.66	.06	.53	.09	.63	.10	.81	.04
<i>M</i>	.67		.54		.64		.77	

**Table 4**  
**Derived Guessing Sensitivities With Standard Errors (SE)**  
**for Experiment 1**

	Lag 1		Lag 2		<i>M</i>
	<i>d'</i>	<i>SE</i>	<i>d'</i>	<i>SE</i>	
Low imageability	.54	.17	1.36	.42	.95
High imageability	.60	.37	1.29	.36	.95
<i>M</i>	.57	.20	1.32	.27	

et al. (1988; also given in the introduction) to the data from each subject's final test session. The resulting guessing accuracies (see Table 3) were subjected to a signal detection analysis (Green & Swets, 1966) in order to distinguish guessing sensitivity from guessing bias. This was done by using the probability of guessing "true" given a true sentence as the hit rate, and the probability of guessing "true" given a false sentence as the false alarm rate for each subject. The mean guessing sensitivities are shown in Table 4. They were subjected to a  $2 \times 2$  (lag  $\times$  imageability) repeated measures ANOVA. The only significant effect was for lag [ $F(1,7) = 7.47$ ,  $p = .03$ ]. A comparable ANOVA on the guessing biases yielded no significant effects or interaction.

As the standard errors in Table 4 show, significant levels of partial information are clearly present (i.e.,  $d' > 0$ ), even at the early lag where decomposition is not necessary due to the very early temporal placement of the signal. Furthermore, this partial information grows significantly from the early lag to the late lag, indicating a minimum of four states of response information (including the initial state of zero sensitivity and the final state of asymptotic sensitivity) during the course of a typical trial.

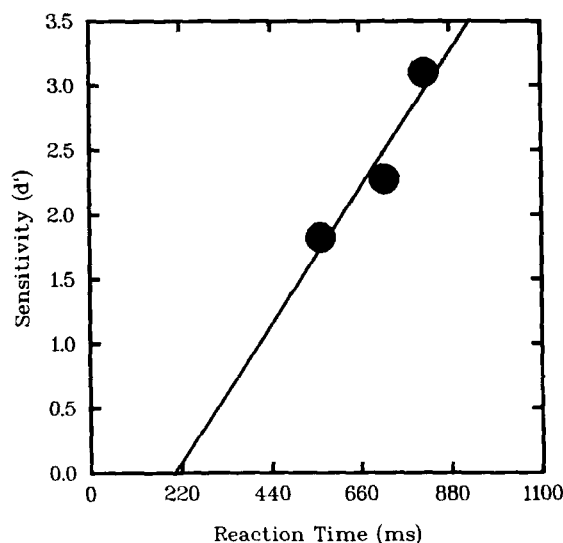
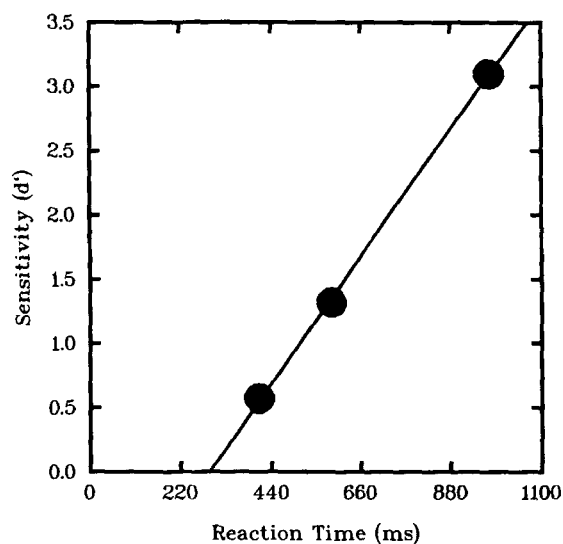
Interestingly, the imageability factor (and its correlated factors) had no discernible effect on the level of partial information (see Table 4). It can therefore be asserted (albeit tautologically) that the partial information detected in this experiment was not generated by processes sensitive to imagistic properties of the information being processed (or to any correlated confounds). This information was apparently abstract in nature.

**Comparison of guessing- and regular-trial sensitivities.** It is also useful to directly compare the guessing sensitivities obtained for the early and late response-signal lags with the sensitivity obtained on regular trials. The regular-trial  $d'$  values were 2.98 for low imageability and 3.24 for high imageability, with a mean of 3.11. These values were obtained by separately averaging the hit and false alarm rates prior to computing  $d'$  (because some of the subjects yielded hit rates of 1.00—see Macmillan & Creelman, 1991, p. 275). These results are plotted in Figure 1 (top panel) as a function of mean guessing time and mean regular-trial RT (including error trials) along with the best-fitting linear regression function (viz.,  $d' = .0045T - 1.30$ ). This line collapses across levels of imageability, because imageability had no discernible effect on guessing sensitivity and no main effect on regular-trial sensitivity.

Because only three data points are involved, it may be premature to draw strong inferences from this figure alone. However, the extremely tight fit of the regression line suggests a theoretically important possibility, namely, that response information accrues at a constant rate over the course of the average trial.

## EXPERIMENT 2

Why is such linearity theoretically important? In a previous SAD verification study using class-inclusion statements (e.g., "All dogs are animals"), Kounios et al. (1987) found that partial information grew at an ap-



**Figure 1.** Mean guessing sensitivity ( $d'$ ) for two response-signal lags and mean regular-trial sensitivity plotted as a function of reaction time (in milliseconds) for Experiments 1 (top panel) and 2 (bottom panel). The lines were fitted by a least squares regression.

proximately linear rate. However, in contrast to the present results, they noted that extrapolation of regression lines fitted to the guessing sensitivities predicted regular-trial performance that was either much less accurate or much slower than was actually obtained. One explanation for these results was that verification might be accomplished by two independent, parallel mechanisms in their task. The first is a comparatively slow feature-comparison mechanism yielding partial information. The other is a relatively fast, discrete, all-or-none search of semantic memory yielding no partial information. Whichever mechanism yields a verification judgment first determines the response on that trial.

Under this scenario, mean RT (and accuracy) on regular trials is a mixture of the RTs (and accuracies) produced by the search and comparison mechanisms. Since SAD removes the contribution of all-or-none processes, the partial-information accumulation function would therefore reflect partial outputs only from the comparison mechanism. So, extrapolation of this partial-information function would predict the accuracy and latency of the slow comparison process that comprises only one component of regular-trial performance; it would not reflect the hypothetical fast, discrete search of semantic memory (see also Kounios, 1993).

In contrast, because a single regression line fits the average guessing sensitivities and the regular-trial sensitivity, the present results do not imply the existence of two verification processes utilizing information available in different grain sizes. At present, it is unknown whether this pattern characterizes many (or most) relational concepts, or whether it is a peculiarity of the ownership relation.

Because the apparent linearity of information accumulation in this task has significant theoretical consequences, this finding is worth further examination. For instance, it could be the case that a later response-signal lag might show that partial information asymptotes at an intermediate level and does not rise to the level of information exhibited on regular trials (cf. Meyer et al., 1988; Ratcliff, 1988). This could be determined by replicating the experiment with a larger number of signal lags, thereby obtaining a more precise picture of the overall shape of the information accumulation function.

Unfortunately, we were unable to compile enough acceptable stimuli (while maintaining a balanced experimental design) to support an increase in the number of signal-lag conditions. The alternative we opted for was to retain the use of two lags, but to shift them to a later temporal region. The strategy was to obtain estimates of partial information available late during a trial, thereby clarifying whether partial-information accumulation does, in fact, continue its apparently linear trend up to the level of sensitivity exhibited on regular trials.

This approach also has difficulties. The later the placement of a response signal, the more likely it is that the subject will respond before the signal. In this case, estimates of partial information are based on less guess-

**Table 5**  
Mean Reaction Times (RT, in Milliseconds)  
for Correct Responses, Standard Errors (SE),  
and Accuracies for Regular Trials From Experiment 2

	True			False		
	RT	SE	Accuracy	RT	SE	Accuracy
Low imageability	906	150	.91	741	96	.93
High imageability	720	106	.96	902	155	.95

ing data, since responses occurring before a signal are based on normal, complete processing. If a signal is placed too late, then the derived guessing accuracies will be based on much less guessing data and may therefore be less reliable. Consequently, a certain amount of experience (and luck) is necessary to anticipate where to put a very late signal (for each subject) so that it is late enough to yield evidence about the final phase of a trial, and yet not so late that the interpretation of the data becomes problematic.

### Method

**Procedure.** The signal lags varied somewhat across subjects. Our goal was to place the signals so that guessing responses to the late signal would cut the distribution of regular-trial RTs at about its median (meaning that only about half of the late-signal responses would be guesses). Early-signal onsets averaged (across subjects) 180 msec earlier than late-signal onsets. Otherwise, conditions for the additional and previous subjects were exactly the same.

**Subjects.** Nine subjects were run in Experiment 2. Unfortunately, 4 of them speeded up so much in the test session (relative to the final practice session) that the signal lags chosen for them were too late to yield usable data. For example, one of these subjects made *all* of his regular-trial responses before the earlier of the two response signals.

### Results

**Regular-trial data.** The mean correct RTs and accuracies for the regular-trial data from Experiment 2 are presented in Table 5. A  $2 \times 2$  (truth  $\times$  imageability) repeated measures ANOVA was performed on the correct RT data. Only the interaction between truth and imageability was marginally significant [ $F(1,4) = 6.42, p = .06$ ]. A comparable ANOVA on the accuracy data yielded no significant main effects or interaction.

The 5 subjects from Experiment 2 were, overall, slightly faster and more accurate than the 8 subjects from Experiment 1. Experiment 2 also showed a cross-over interaction in the RTs (see Table 5), though these results are not complicated by the speed-accuracy trade-off noted in the regular-trial RT results for Experiment 1.

**Guessing times.** Because the guessing times for Experiment 2 were derived on the basis of thinner data (due to the lateness of the signals), they are necessarily somewhat noisier. This renders statistical tests for significant effects of various factors on the shapes of these guessing-time distributions less informative, since random noise in the data could, in principle, obscure systematic effects. For example, a repeated measures ANOVA on the

Table 6  
Derived Guessing Accuracies With Standard Errors (SE) for Experiment 2

	Lag 1				Lag 2			
	True		False		True		False	
	Accuracy	SE	Accuracy	SE	Accuracy	SE	Accuracy	SE
Low imageability	.70	.12	.92	.06	.67	.08	.89	.06
High imageability	.85	.11	.75	.10	1.02	.12	.71	.08
<i>M</i>	.78		.84		.85		.80	

Note—The speed-accuracy decomposition analysis can occasionally yield derived guessing accuracies slightly less than .00 or greater than 1.00 due to noise in the raw data.

guessing-distribution medians using the factors of imageability, truth, and lag yielded no significant main effects or interactions (all  $p$ s > .36). Consequently, the fit of the parallel sophisticated-guessing model cannot be conservatively evaluated for these data. Nevertheless, the highly systematic nature of the results indicates that the procedure for deriving these guessing accuracies is quite robust.

Early-signal guessing-time medians cut the distribution of regular-trial RTs at the 19th percentile. Late-signal guessing-time medians cut the regular-trial RT distribution at its 54th percentile.

**Guessing accuracies.** The derived guessing accuracies (see Table 6) were obtained with the same procedure used in Experiment 1. A repeated measures ANOVA (truth  $\times$  imageability  $\times$  lag) performed on these accuracies yielded only a significant interaction between truth and imageability [ $F(1,4) = 7.47, p = .05$ ]. Table 6 shows that this interaction followed the general pattern found for the regular-trial RTs. Apparently, the pattern of results observed on regular trials starts to take shape only during this late phase of a trial.

Unfortunately, the application of signal detection analysis to individual subject data would have been problematic, because noisiness in the raw data (due to the lateness of the signals) yielded some guessing accuracies in excess of 1.00. So, instead of rounding these individual values down (in some cases substantially) to .99, average hit and false alarm rates were calculated across subjects. As Table 6 shows, this procedure resulted in only one value (viz., true/high-imageability sentences for the late lag) slightly in excess of 1.00 (viz., 1.02), which was rounded down to .99 before calculating the group signal detection parameters.<sup>7</sup>

The response biases for guesses (in units of  $\log[\beta]$ ) were: high imageability,  $-.40$  for the early lag and  $-2.43$  for the late lag; low imageability,  $.84$  for the early lag and  $.65$  for the late lag. For both types of stimuli, the

biases became more negative from the early lag to the late, indicating an increasing tendency to guess "true" over the course of a trial. The regular-trial biases were  $-.28$  for high-imageability sentences and  $.13$  for low-imageability sentences.

The guessing sensitivities are shown in Table 7. Unfortunately, the method of calculating signal detection parameters on the basis of average hit and false alarm rates does not permit conventional significance testing. Nevertheless, the  $d'$  values indicate certain tentative conclusions that are consistent with other analyses from Experiment 1. First, there appear to be substantial levels of partial information in each condition (cf. standard errors in Table 6). The main effects of lag and imageability, if real, are apparently small. Furthermore, these effects are apparently due to one cell of Table 7, namely, the  $d'$  value for high-imageability sentences at the late lag. (This conclusion is supported by the significant interaction between imageability and truth on the guessing accuracies.) At this later lag, high-imageability sentences seem to yield substantially higher sensitivity than low-imageability sentences (cf. Table 6)—a difference not present at the earlier lag or in Experiment 1. This suggests an influence of imageability (or a correlated confound) at a comparatively late point during a trial.

Figure 1 (bottom panel) shows the  $d'$  values for early- and late-lag guesses (plotted as a function of median guessing time) and for the regular trials (plotted as a function of mean RT for correct and incorrect responses) with the best-fitting regression line. The regular-trial  $d'$  values were 2.81 for low imageability and 3.39 for high imageability, with a mean of 3.10; these were obtained by separately averaging the hit and false alarm rates prior to computing  $d'$  (because some of the subjects yielded hit rates of 1.00—see note 7). The fit of this line is clearly not as good as the fit of the regression line for Experiment 1, and no strong inferences can therefore be made from these results alone. Furthermore, the presence of an imageability effect at the late lag raises the possibility that irrelevant heuristics may have contributed some partial information. Nevertheless, it is very suggestive that the regression lines for Experiments 1 and 2 have similar slopes and intercepts (i.e., Experiment 1,  $d' = .0045T - 1.30$ ; Experiment 2,  $d' = .0048T - .98$ ). These results suggest that partial information does grow to high levels in this task and does not as-

Table 7  
Derived Guessing Sensitivities ( $d'$ ) for Experiment 2

	Lag 1	Lag 2	<i>M</i>
Low imageability	1.92	1.67	1.80
High imageability	1.72	2.88	2.30
<i>M</i>	1.82	2.23	



ymptote at an intermediate level (cf. Meyer et al., 1988; Ratcliff, 1988).

In spite of the practical difficulties involved in drawing detailed conclusions from the guessing results of Experiment 2, these data do support at least one theoretically important conclusion concerning the overall time course of information accumulation in this task. Even though rounding the late-signal hit rate down to precisely .99 is arguably somewhat arbitrary (but see note 7), the "true" late-signal sensitivity is undoubtedly very high, indicating that partial information does not asymptote at an intermediate level in this task (in contrast to the double-string lexical decision task used by Meyer et al., 1988). This supports the notion that the partial-information accumulation results from Experiment 1 can, in fact, be linearly extrapolated to the level of sensitivity obtained on regular trials.<sup>8</sup>

A further (speculative) observation suggested by these results is that although partial information (i.e., sensitivity) seems to accumulate at a constant rate, this accumulation may not consist of just one type of information. The absence of an imageability effect early during a trial (Experiment 1) indicates that the information being processed in this time window is abstract in nature. In contrast, in Experiment 2, the form of the interaction between lag and imageability (cf. Tables 6 and 7) suggests that partial-information accumulation late during a trial may be largely or completely due to imagistic information.

## DISCUSSION

### Summary

The primary purpose of the present study was to determine whether semantic relations other than the much studied class-inclusion relation should be best conceived of as atomistic, unitary associations (e.g., as links in a classical network), or as complex concepts consisting of a number of relational elements. The relation of ownership was selected as a test case (following the suggestion of Glass & Holyoak, 1975), because subjects cannot effectively compute the truth of this relation by using a similarity-based heuristic. The complexity of this relation was examined by applying SAD to the verification of ownership statements in order to determine whether subjects achieve their final state of response information in a single, discrete, all-or-none transition, or whether the relevant processes yield partial response information representing intermediate states of knowledge. The rationale was that the retrieval of a unitary relational link should be an all-or-none affair. In contrast, a set of relational elements need not be processed as a unitary bundle, thereby allowing intermediate response-information states. Most importantly, we found solid evidence for such partial information, lending support to the general approach embodied by relation element theory (Chaffin & Herrmann, 1988). In addition, the present study also yielded evidence that response information may grow in

a linear fashion to the final level of sensitivity exhibited on regular trials. This suggests a relatively simple process whereby relational elements are evaluated at a constant rate (cf. Kounios, 1993; Kounios et al., 1987).

### Alternate Interpretations

Do the partial-information results presented in this report necessarily demonstrate the componential structure of the ownership relation? There are three other interpretations that should be considered.<sup>9</sup>

**Exemplar model.** First, suppose that the subjects verified each sentence by retrieving exemplars corresponding to the relation. For instance, a subject might verify the sentence "Many people own dogs" by retrieving instances of specific people that own dogs. Every retrieved instance of a person who owns a dog would add a quantal unit of response information supporting the notion that the sentence is true; every retrieved instance of a person who does not own a dog similarly adds to the evidence that the sentence is false. The subject adopts information thresholds for responding "true" or "false" based, in part, on his or her interpretation of the quantifier *many* (Glass, Holyoak, & O'Dell, 1974). Whichever threshold is exceeded first determines the response. According to this approach, the ownership relation between each instance of a person and the owned dog can be unitary; it is the retrieval of many such instances that yields partial information. So there is no need to posit any intrarelational structure.

Although exemplar representations of semantic relations have not previously been explored in the literature, exemplar representations of other types of concepts and categories are currently receiving much attention (Brooks, 1987). Indeed, such a model neatly accounts for the existence of partial information without recourse to a componential ownership relation. Unfortunately, although the exemplar approach cannot be conclusively disproved by the present data, it cannot be considered a particularly plausible explanation for the presence of partial information in this task. For example, how might exemplars be represented in such a model? Subjects might retrieve images of specific people with or without dogs. However, it would seem that these exemplars cannot be represented by imagery in this task, since imageability had no effect on partial information accumulation during a large (i.e., early) part of a trial.

If, however, the hypothesized exemplars are not represented imagistically, then how else might they be represented? They could be represented by unitary ownership links between nodes representing specific individuals and the objects they own. This is tantamount to saying that people store lists of all the objects owned by each individual they know, or lists of all the people that own each specific object they know. Suppose that people do actually store such lists of relational exemplars. In this case, the retrieval of these items from memory would have to proceed at a relatively constant rate over time in order to yield a linear accumulation of partial informa-

tion. This would mean that the retrieval of relational exemplars does not exhibit the sorts of chunking or clustering phenomena that characterize the retrieval of other types of information from long-term memory (see Crowder, 1976, chap. 10). So, although an exemplar model does not seem a particularly persuasive explanation for the present results, the possible role of relational exemplars in semantic memory is nevertheless an interesting and important topic for further investigation.

#### **Continuous probabilities assigned to unitary links.**

Another possibility is that the ownership relation consists of a unitary link, but that this link is associated with a level of activation or a probability weight that varies continuously over time, as might occur in a spreading activation model (cf. Collins & Loftus, 1975). These continuous activation values might manifest themselves in terms of accumulating partial information.

This sort of interpretation has at least two difficulties—one theoretical and one empirical. First, if continuous probabilities are assigned as weights to unitary links, then how are these probabilities determined? The computation of continuous probabilities most plausibly results from the processing of continuous (or at least quasi-continuous) information (Kounios, 1993; see also Miller, 1988). In other words, continuous, fine-grain information cannot result from computations involving discrete, large-chunk information, unless the latter has an internal structure allowing it to be decomposed into smaller-grain constituent elements. So it is difficult to imagine how continuous weights could be computed by a classical associative network incorporating discrete, unitary links. An alternative would be to have the probabilities computed outside of the network, but then one would have to hypothesize small-grain (i.e., continuous) semantic information stored outside the network as well. Such a model is, at best, strained.<sup>10</sup>

Second, it is not clear how continuous weights (however they are computed) assigned to unitary links could yield a linear growth of sensitivity (and no reliable change in response bias; cf. Kounios et al., 1987). In contrast, such a pattern of results is an intuitive outcome for a sequential-sampling model in which features are compared at a constant rate.

**Timer model.** A third alternative interpretation argues that partial information may be generated from a simple search combined with a "timer" mechanism (Kounios, 1985). For instance, a subject may start a timer while searching memory for a (unitary) stored relation. As time passes without finding the relation, the subjective probability that the relation will not be found may increase (i.e., because false propositions are not stored). When some deadline is exceeded without retrieving the proposition, the sentence is judged to be false.

Such a timer model might, in fact, predict a continuous accrual of response accuracy. Nevertheless, a timing mechanism is not entirely consistent with other aspects of the data. First, in the present study as well as in the previous SAD study by Kounios et al. (1987), evidence

for a linear growth of sensitivity over time was found. In order to produce such a pattern of results, a timer model would have to output  $d'$  units of sensitivity as a linear function of (real) time elapsed. However plausible or implausible such a linear semantic memory clock might be, this notion loses credibility because the rate of such a clock would have to be independent of the rate of the ongoing search. For instance, if the timer generates sensitivity at a constant rate and if this rate is in fact influenced by the rate of the search process (as one might expect), then this implies that relations are being retrieved (in a serial search) at a constant rate. But even at the limits of neural performance, a serial search of a person's vast semantic memory store would take an unrealistic amount of time. A much more likely possibility is that the search is parallel. However, for a parallel mechanism to search at a constant rate would mean that there would have to be a precise tradeoff between the number of relations evaluated per unit of time and the number of relations being evaluated (in parallel) at any given moment (cf. the fan effect; see Anderson, 1990). For example, in a model in which a search fans out from a starting point in a network (cf. Collins & Loftus, 1975), the fanning would increase (over time) the number of relations being simultaneously evaluated. This increase in the number of parallel operations would have to precisely complement changes in the duration of the evaluation of each relation in order to yield a linear growth of sensitivity.

Second, and more importantly, a timer model should predict an increasing bias to respond "false" over the course of a trial. Kounios et al. (1987) did not find consistent, reliable changes in response bias. We found an increase in the bias to respond "false" from the early lag to the late lag in Experiment 1; however, this change did not approach significance. The trend was in the opposite direction in Experiment 2.

#### **Future Directions**

As for future research, the most obvious strategy would be to determine whether the verification of other semantic relations yields partial information as well. The fact that ownership-statement verification yields partial information does not guarantee that other relations behave in a similar fashion. If complex relations (e.g., ownership) are composed of sets of unitary elemental featural relations, then, theoretically, the verification of individual isolated elemental relations would be accomplished in a discrete, all-or-none fashion. So, the strategy utilized in the present research might be expanded to examine a broad range of relations in an empirically based search for fundamental relational elements. Since there are a large number of candidate relations, such a program of research would, of practical necessity, need to be guided by theoretical linguistic analyses (e.g., Chaffin & Herrmann, 1988).

Furthermore, the case we have given against alternate theoretical interpretations of the partial information ob-

tained in this task depends heavily on the notion that guessing sensitivity accrues at a constant rate. Consequently, it is important for future research to map out more precisely the exact shape of this partial-information accumulation function. This will require the use of other relations yielding a larger number of usable stimuli.

## REFERENCES

- ANDERSON, J. R. (1990). *Cognitive psychology and its implications*. New York: W. H. Freeman.
- BROOKS, L. T. (1987). Decentralized control of categorization: The role of prior processing episodes. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 141-174). Cambridge: Cambridge University Press.
- CHAFFIN, R., & HERRMANN, D. J. (1988). The nature of semantic relations: A comparison of two approaches. In M. Evens (Ed.), *Relational models of the lexicon* (pp. 289-334). Cambridge: Cambridge University Press.
- CHANG, T. M. (1986). Semantic memory: Facts and models. *Psychological Bulletin*, **99**, 199-220.
- COLLINS, A. M., & LOFTUS, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review*, **82**, 407-428.
- COLLINS, A. M., & QUILLIAN, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning & Verbal Behavior*, **8**, 240-247.
- CROWDER, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.
- DE JONG, R. (1991). Partial information or facilitation? Different interpretations of results from speed-accuracy decomposition. *Perception & Psychophysics*, **50**, 333-350.
- FRANCIS, W. N., & KUČERA, H. (1982). *Frequency analysis of English usage*. Boston: Houghton Mifflin.
- GLASS, A. L., & HOLYOAK, K. J. (1975). Alternative conceptions of semantic memory. *Cognition*, **3**, 313-319.
- GLASS, A. L., HOLYOAK, K. J., & O'DELL, C. (1974). Production frequency and the verification of quantified statements. *Journal of Verbal Learning & Verbal Behavior*, **13**, 237-254.
- GREEN, D. M., & SWETS, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- HINTON, G. E., MCCLELLAND, J. L., & RUMELHART, D. E. (1986). Distributed representations. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 1. Foundations* (pp. 77-109). Cambridge, MA: MIT Press.
- HOLLAN, J. D. (1975). Features and semantic memory: Set-theoretic or network model? *Psychological Review*, **82**, 154-155.
- JORDAN, M. I. (1986). An introduction to linear algebra in parallel distributed processing systems. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 1. Foundations* (pp. 365-422). Cambridge, MA: MIT Press.
- KOUNIOS, J. (1985). *A new approach to the study of human semantic memory: Analyzing the intermediate products of rapid semantic processes*. Unpublished doctoral dissertation, University of Michigan, Ann Arbor.
- KOUNIOS, J. (1993). Process complexity in semantic memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 338-351.
- KOUNIOS, J. (1994). *On the continuity of thought and the representation of knowledge: RT, SAT, SAD, and ERPs reveal levels of structure in semantic memory*. Manuscript submitted for publication.
- KOUNIOS, J., & HOLCOMB, P. J. (1992). Structure and process in semantic memory: Evidence from event-related brain potentials and reaction times. *Journal of Experimental Psychology: General*, **121**, 459-479.
- KOUNIOS, J., & HOLCOMB, P. J. (1994). Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **121**, 459-479.
- KOUNIOS, J., OSMAN, A. M., & MEYER, D. E. (1987). Structure and process in semantic memory: New evidence based on speed-accuracy decomposition. *Journal of Experimental Psychology: General*, **116**, 3-25.
- MACMILLAN, N. A., & CREELMAN, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.
- McCLOSKEY, M. (1980). The stimulus familiarity problem in semantic memory research. *Journal of Verbal Learning & Verbal Behavior*, **19**, 485-502.
- McCLOSKEY, M., & GLUCKSBERG, S. (1979). Decision processes in verifying category membership statements: Implications for models of semantic memory research. *Cognitive Psychology*, **3**, 313-339.
- MEYER, D. E., & IRWIN, D. E. (1994). *Speed-accuracy decomposition based on intersensory facilitation, illusory temporal correlation, and extrasensory premonition: A reply to De Jong*. Manuscript in preparation.
- MEYER, D. E., IRWIN, D. E., OSMAN, A. M., & KOUNIOS, J. (1988). The dynamics of cognition and action: Mental processes inferred from speed-accuracy decomposition. *Psychological Review*, **95**, 183-237.
- MILLER, J. (1988). Discrete and continuous models of human information processing: Theoretical distinctions and empirical results. *Acta Psychologica*, **67**, 191-257.
- RATCLIFF, R. (1988). Continuous versus discrete information processing: Modeling accumulation of partial information. *Psychological Review*, **95**, 238-255.
- RATCLIFF, R., & MCKOON, G. (1982). Speed and accuracy in the processing of false statements about semantic information. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **8**, 16-36.
- RIPS, L. J., SMITH, E. E., & SHOEN, E. J. (1975). Set-theoretic and network models reconsidered: A comment on Hollan's "Features and semantic memory." *Psychological Review*, **82**, 156-157.
- SMITH, E. E., SHOEN, E. J., & RIPS, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, **81**, 214-241.

## NOTES

1. The equation used to decompose the distribution of reaction times obtained on response-signal trials (thereby extracting the distribution of guessing times) is:

$$1 - P(t_g > C) = 1 - [P(t_s > C)/P(t_n > C)],$$

where  $[1 - P(t_g > C)]$  is the cumulative distribution function that gives the probability that the time to make a guessing response on a response-signal trial with a particular response-signal lag is less than constant  $C$ ,  $P(t_s > C)$  is the probability that the time taken to respond on a signal trial is greater than constant  $C$ , and  $P(t_n > C)$  is the probability that the time taken to complete normal processing of the stimulus (i.e., reaction time on regular trials) is greater than constant  $C$ . The explanation and derivation of this equation are given by Meyer et al. (1988).

2. De Jong (1991) recently argued that, under certain circumstances (e.g., Meyer et al., 1988), but not others (e.g., Kounios et al., 1987; Ratcliff, 1988), SAD might yield spurious evidence of partial information. According to this hypothesis, the presentation of an auditory response signal can sometimes cause intersensory facilitation of the normal processes, invalidating the assumption of the parallel sophisticated-guessing model that the normal and guessing processes are independent. A rebuttal of De Jong's argument is beyond the scope of this article. A complete analysis and critique of De Jong (1991) will be provided by Meyer and Irwin (1994). However, two points should be made here. First, specific procedural details of the present study were designed to preclude the possibility of such intersensory facilitation (e.g., using a visual pattern mask as a response signal instead of using a tone). And second, we have recently discovered a task that yields no measurable partial information, namely, an anagram-decision task in which the subject must decide whether or not any combination of the letters in a stimulus string represents a legal word. This finding is virtually impossible to obtain with any applicable variant of De Jong's model and further supports the viability of SAD.

3. In contrast to the typical RT experiment, SAD experiments usually involve collecting large amounts of data from a small number of highly trained subjects.

4. Of the 8 subjects whose data are reported here, 1 subject yielded much slower mean regular-trial RTs during training. In order to obtain guessing data from her that were comparable to those from the other subjects, she was given later response-signal lags (early signal, 160 msec; late signal, 340 msec). These lags maintained the 180-msec difference between the early and late response-signal onsets.

5. Our original intention was to delete all signal trials with RTs longer than 350 msec after the onset of the response signal in order to preclude the possibility that the tone indicating that the response was too slow might interfere with processes of interest (De Jong, 1991). However, such trials accounted for only about 1% of the data and did not have a significant influence on the results. We therefore retained these few "slow" trials.

6. While none of the effects on the shapes of these distributions were significant, it should be noted that it is not presently known precisely how statistically robust the SAD analysis is. It is possible that major violations of the parallel sophisticated-guessing model would be of little consequence to the extraction of guessing accuracy, or that minor violations might have major consequences. Fortunately, the critical finding of partial information does not depend on the assumption that SAD is very robust. This is because the early signal lag was placed early enough that virtually all responses in this condition occurred before any regular-trial responses. In this case, any sensitivity found for responses to the early signal must reflect partial information. The detection of such partial information does not rely on the SAD analysis, because no decomposition is necessary for signal-trial responses occurring this early.

7. Macmillan and Creelman (1991, p. 275) present a rationale for averaging hit rates and false alarm rates (separately) across subjects before computing  $d'$  when some subjects yield hit rates of 1.00 or false alarm rates of 0.00. They also discuss several methods for correcting or adjusting perfect hit and false alarm rates (p. 10). Unfortunately, none of these latter methods are applicable in the present circumstances, because these proportions are estimated by the SAD technique on the basis of pooled subject data rather than computed from trial data (as in a psychophysical experiment). However, as Macmillan and Creelman state: "The largest possible finite value of  $d'$  depends on the number of decimal places to which  $H$  and  $F$  are carried. When  $H = .99$  and  $F = .01$ ,  $d' = 4.65$ ; many experimenters consider this an effective ceiling" (p. 10). So, our rationale in rounding the  $d'$  of 1.02 to .99 was simply to round the hit rate down to the largest meaningful and practical value, since above .99,  $d'$  explodes. Fortunately, though the results of the present study are strengthened by the fact that this adjusted

data point lies right on the line fitted to the other points (see Figure 1), they do not depend on this point, because the line fitted to the other three points alone is virtually identical. Furthermore, even if the correction employed yielded an underestimate of the actual guessing sensitivity, there would still be little doubt that the obtained guessing sensitivity for the late lag of Experiment 2 ruled out the possibility that the real guessing sensitivity did not rise after the early lag of the same experiment.

8. As a speculative exercise, consider what happens when a linear regression is performed on the combined results of Experiments 1 and 2. The four points in this regression correspond to the average guessing sensitivities for Experiment 1, Lag 1; the average of Experiment 1, Lag 2 and Experiment 2, Lag 1; Experiment 2, Lag 2; and the regular-trial sensitivities averaged across Experiments 1 and 2 (all plotted as a function of average guessing time or regular-trial RT). It was deemed justifiable to average the sensitivities and guessing times for Experiment 1, Lag 2 and Experiment 2, Lag 1, because the signal placements, guessing sensitivities, and guessing times were quite close. (The similarity of these two data points further justifies the assumption that the data from Experiments 1 and 2 are comparable and combinable.) The best-fitting linear regression function (i.e.,  $d' = .0057T - 1.76$ ) provides an extremely tight fit and further substantiates the hypothesized linearity of information accumulation suggested by Experiment 1 (see Figure 1, top panel) and the previous SAD results of Kounios et al. (1987). Suggestively, the problematic late-signal point from Experiment 2 falls almost precisely on this regression line.

9. The exemplar interpretation was suggested by Robert Lorch.

10. It should be noted that although classical associative networks may have a difficult time producing linear partial-information accumulation, neither relation element theory nor the data presented in this report are necessarily incompatible with the notion that relations may be stored as distributed representations in a parallel distributed processing (PDP) network (Hinton, McClelland, & Rumelhart, 1986). For example, if relations are in fact composed of different combinations of elements drawn from a common pool, then this pool of elements may form a subset of the microfeatures used to represent concepts in a PDP network. Furthermore, some PDP models can yield linear (i.e., "semilinear") activation functions analogous to the linear accumulation of partial information observed in the present study (Jordan, 1986, pp. 420–421). However, the actual implementation of relational representation in PDP systems is not a straightforward matter (e.g., the "binding problem"; see Hinton et al., 1986).

## APPENDIX

High Imageability		Low Imageability	
True	False	True	False
backpacks	squirrels	insurance	breezes
bathrobes	jets	substances	constellations
beds	lizards	bonds	days
bracelets	palaces	companies	environments
brooms	armor	music	God
cameras	bridges	investments	asteroids
cards	buffaloes	wills	capitals
cars	castles	stocks	cities
chairs	deer	businesses	continents
clocks	eagles	annuities	counties
coats	mummies	pensions	countries
combs	spiders	accounts	galaxies
couches	eels	policies	ghosts
crayons	panthers	property	harbors
dice	robots	equipment	hemispheres
dishes	spears	supplies	illness
dogs	doves	merchandise	immortality
earrings	zebras	things	ozone
envelopes	monkeys	equity	seasons
flashlights	tanks	collections	thunder
food	cobras	franchises	universes
fruit	harpsichords	items	Venus
glasses	hawks	proxies	wind
gloves	camels	ideas	bays
goldfish	snails	patents	canyons
hats	frogs	copyrights	comets
houses	ships	belongings	craters
jewelry	cranes	provisions	dormitories
keys	crowns	accessories	Earth
luggage	dolphins	products	highways
mittens	leopards	devices	Jupiter
necklaces	owls	implements	lava
oranges	pyramids	wares	Mars
pencils	rhinoceroses	goods	mermaids
pillows	rockets	spices	moons
purses	salamanders	commodities	nature
radios	spaceships	objects	Neptune
records	witches	materials	planets
rings	apes	securities	raindrops
scissors	piers	stores	rivers
shampoo	thrones	reserves	Saturn
skis	tigers	savings	seas
socks	towers	funds	slate
sofas	trains	containers	sleet
stereos	armies	receptacles	snowflakes
sweaters	dynamite	hardware	stars
telephones	elephants	assets	streets
televisions	fields	possessions	suns
towels	pastures	gadgets	trolls
umbrellas	railroads	resources	universities
wallets	tracks	inventions	valleys
watches	shields	contraptions	volcanos