

SEMANTIC MEMORY RETRIEVAL: ANALYSIS BY SPEED ACCURACY TRADEOFF FUNCTIONS*

ALBERT T. CORBETT AND WAYNE A. WICKELGREN

*Psychology Department, University of Oregon, Eugene,
Oregon 97403, U.S.A.*

Semantic memory retrieval for verifying category-example associations was tested by a speed accuracy tradeoff method: present the category for 2 s, present a correct or incorrect example followed, after a variable lag (0, 0.1, 0.2, 0.3, 0.4, 0.6, 0.8, 1, 2, or 3 s), by a signal to make a "yes-no" response in about 0.2 s. Although the strength of the category-example association is higher for high dominance examples of a category, retrieval dynamics did not vary with dominance level. Recognition for category-example associations appears to be a direct-access (parallel) retrieval process. Priming a category by repeated testing of the same category over three consecutive trials had no effect on either asymptotic strength or retrieval dynamics. Partitioning into short, medium, and long latency responses at each lag produced microtradeoff functions which did not lie on the same macrotradeoff function. Retrieval dynamics were invariant with long-term practice.

Introduction

People decide faster that *A robin is a bird* than that *A chicken is a bird*. A robin is also rated to be a more typical example of a bird than is a chicken. In general, the instances of a category which are rated as most typical of the category or are most frequently produced by subjects as examples of that category are also most rapidly verified as examples of that category (Loftus, 1973a; Rips, Shoben and Smith, 1973; Rosch, 1973; Smith, Shoben and Rips, 1974). Why are some category-example associations verified faster than others? Rosch (1973) suggests a serial search model in which a category is entered at its core meaning and examples are searched in a relatively fixed order from the best (most typical) examples to the worst (least typical) examples. Anderson and Bower (1973, pp. 380-81) also assume a similar serial search.

Smith, Shoben and Rips (1974) have a two-stage model that distinguishes between the defining and the characteristic features of examples and categories. *Robins*, for instance, are said to have defining features (such as biped, wings, distinctive shape and colouring, etc.) which are relatively essential features for

* This research was supported by NSF grant BMS75-03145, by NIE grant NIE-G-77-008, and by contract F 44620-73-C-0056 from the Advanced Research Projects Agency, D.O.D., to Ray Hyman. Requests for reprints should be sent to Wayne A. Wickelgren, Psychology Department, University of Oregon, Eugene, Oregon 97403, U.S.A.

an entity to be classified as a *robin*, but *robins* are also said to have characteristic features (such as perch in trees, undomesticated, etc.) which are less essential for being classified as a *robin*. They propose a two-stage model of verification for category-example associations similar to that of Juola, Fischler, Wood and Atkinson (1971). In this model, all features (defining and characteristic) of both category and example are retrieved and matched with equal weighting in the first stage but only defining features (or all features, but weighted by degree of definingness) are retrieved and matched in the second stage. If the degree of match (similarity) in the first stage is above some high criterion, a fast "yes" is produced without having to go through the second stage. If the first stage similarity is below some low criterion, a fast "no" is produced, avoiding the need for a second stage. Only for intermediate degrees of similarity is it necessary to invoke the second stage which operates only on defining features. Thus, according to Smith *et al.* (1974), the average verification time for typical examples should be less than for atypical examples because of a difference in the frequency with which the second stage is required.

The similar two-stage model of Juola *et al.* (1971) would make the same prediction, since the examples which subjects rate as most typical of a category also tend overwhelmingly to be the examples with the highest strength of association to the category as measured by the frequency of subjects' producing the example when generating instances of the category (Battig and Montague, 1969; Hunt and Hodge, 1971). Juola *et al.* assume that the first-stage of retrieval is a direct-access process which produces a "yes-no" decision if the retrieved strength of association is above a high criterion or below a low criterion. If associative strength has a value between the two criteria, a serial search of the examples of a category (in this case) is conducted. Although typicality ratings and associative strength measured by production norms are highly correlated, Smith *et al.* (1974) report that it is actually associative strength that is the better predictor of category-example verification time. Hence, the present study uses category-example production norms as a measure of what Loftus (1973a) calls instance *dominance*, the strength of association from a category to an instance (as distinct from category dominance which is the reverse association measured by the frequency of generating superordinates of an example, e.g. Loftus and Scheff, 1971). Wilkins (1971) and Loftus (1973a) have shown that high instance dominance produces faster verification time for category-example associations when the category is presented a second or so before the example, while Loftus (1973a) has shown that high category dominance produces faster verification time for such associations when the example is presented before the category. In the present study the category name was presented before the example, so the term dominance will always refer to instance dominance.

By contrast to the foregoing models, the simplest direct-access model of the verification of category-example associations assumes that the category (e.g. *bird*) and example (e.g. *robin*) nodes are directly activated without search. The stored strength of association between these two nodes (perhaps mediated by a propositional node: e.g. the node representing *a robin is a bird*) is also retrieved directly, without serial searching or scanning of alternative associations (Anderson,

1976; Collins and Loftus, 1975; Wickelgren, 1975). Since there is no serial searching of examples associated with a category, and no two-stage processing differentially associated with low-dominance examples, the simplest direct-access model predicts no difference in the retrieval process for high vs. low dominance category-example associations. What this means will be explained in more detail after discussing the speed accuracy tradeoff (SAT) method for studying memory retrieval dynamics.

Although retrieval is assumed to be direct-access, it is clearly not instantaneous. As a subject is given more time in which to make a recognition decision (item, paired associate, etc.), accuracy improves up to some asymptotic level set by the strength of the relevant association(s) in memory (Corbett, 1977; Doshier, 1976; Reed, 1973, 1976; Wickelgren and Corbett, 1977). A speed accuracy tradeoff (SAT) function may provide a direct measure of the entire time course of memory retrieval dynamics, whereas the reaction time (RT) method yields the equivalent of a single point on such a retrieval dynamics function (RT plus its associated error rate). Also, asymptotic accuracy (with unlimited retrieval time) reflects the strength of an association in storage, not the dynamics of the retrieval process.* A number of SAT studies of memory retrieval (Corbett, 1977; Doshier, 1976; Reed, 1973, 1976; Wickelgren and Corbett, 1977) show that the time course of retrieval is well approximated by an initial period of chance accuracy ($d' = 0$) up to time $T = \delta$ (the intercept) followed by an exponential approach to asymptotic accuracy (λ), namely:

$$d_T = \lambda (1 - e^{-\beta(T-\delta)}) \quad (1)$$

where $\{T-\delta\} = T-\delta$ for $T > \delta$ and 0 elsewhere.

In Eq. (1), d_T is a d' accuracy measure after T ms of processing time, λ is the asymptotic accuracy level, δ is the time intercept of the SAT function. The exponential form of the retrieval function asserts that the absolute amount of retrieval per unit time is proportional to the amount left to be retrieved. The rate parameter, β , is this proportion of remaining, unretrieved, trace strength which is retrieved per ms of retrieval time after the intercept (δ). If the exponential form of retrieval is even approximately valid for category-example associations, it indicates that RT studies of dominance effects may simply be showing that dominant category-example associations have higher asymptotic strength (λ), with no difference from low dominance associations in the intercept and rate parameters (δ and β) which reflect the dynamics of retrieval process. Because of the empirical fact that retrieval accuracy approaches a limit, even small differences in asymptotic accuracy for two conditions can produce large differences in the time to reach a high criterion level of accuracy even with no

* The terms *retrieval time* and *retrieval process* will be used throughout this paper to refer in a theoretically neutral manner to: (a) the entire time from onset of a test category-example pair to the response and (b) all the processes that go on in this period, respectively. One of us (Wickelgren, 1976) believes that serial stage analyses of information processing tasks are appropriate structurally, but not dynamically (at least not in the usual sense). Accordingly, we use the term *retrieval time* in a strictly empirical sense to be equivalent to *processing time*: it is the time allowed for retrieval.

differences in the retrieval process. This is especially true since instructions to subjects in RT studies usually emphasize achieving high accuracy. The present study uses the SAT method to determine whether associative dominance in semantic memory affects asymptotic strength or retrieval dynamics or both.

Collins and Quillian (1970) found that subjects could more rapidly verify the truth of propositions such as "A canary is a bird" or "A canary is yellow" after judging the truth of another proposition about a "canary" than after judging a proposition about some other concept. That is, their results indicated that processing any proposition about canaries *primed* the retrieval from memory of other propositions about canaries, producing a faster verification time. Ashcraft (1976) in recognition and Loftus (1973*b*) in recall have found similar priming effects. Once again, the SAT method can indicate whether priming affects asymptotic accuracy (via a short-term boost in associative strength) or retrieval dynamics.

Also we wished to determine whether the process of retrieval of one association from a node (animal-horse) is responsible for the priming effect on the retrieval of another association from the same node (animal-wolf) or whether accessing the common category (animal) concept alone is sufficient to produce priming effects. Therefore, in the present category-example verification study, the category name was always presented well in advance (2 s) of the example so that priming effects resulting from accessing the category concept will be present in verifying the example even in the first presentation of a category, and no facilitation effects from the first to second to third presentation of a category is expected if priming is based entirely on activation of the category nodes alone. If facilitation effects are still obtained with category repetition they must stem from actually *processing* the category-example relation the first time the category is presented. In the present procedure retrieval time is measured from onset of the example. This procedure eliminates category node priming as the cause of any facilitation effect obtained in repeating a category across trials.

In the current study, we used the response signal method of generating speed-accuracy tradeoff curves. Our SAT functions for verification of a category-example association (e.g. fruit-pear), were generated by the following procedure: A category (fruit) was presented followed in 2 s by an example (pear). At a variable time (0-3 s) following onset of the example, the subject heard a brief tone which was the signal to make a yes-no response regarding whether the example belongs to the category. The delay between presentation of the example and presentation of the signal to respond is termed the lag. The *latency* is the delay between response signal and the subject's response. Total retrieval time is the lag + latency. In plotting our SAT curves, we typically plot accuracy as a function of mean retrieval time (lag + mean latency).

Method

Subjects

Four subjects were paid \$1.50/hour to participate in the experiment.

Materials

Fifty-two category names and 1560 category members (30 from each category), were selected from category production norms (Battig and Montague, 1969; Hunt and Hodge, 1971). Each category name consisted of one, two or three words. The category members were restricted to one or two words and three to 16 letters. Fifty of the categories served as experimental categories, the other two were employed in practice trials at the beginning of each session. The 30 examples in each category were subdivided into 10 high, 10 medium, and 10 low dominance examples.

Procedure

There were a total of 3000 experimental trials divided into 20 sessions. In each session the experimental trials were preceded by six practice trials for a total of 156 trials/session. Subjects participated in two sessions a day, one in the morning and one in the afternoon, each approximately 25 min long. Stimuli were presented on a cathode ray tube. The experiment was controlled by a PDP-15 computer.

At the beginning of each trial READY appeared on the screen for 1000 ms then was replaced by a category name. The category name remained on the screen for 2000 ms then was replaced by a test example. Subjects had to decide if the test example was a member of the category whose name had just been displayed. At a variable interval or lag after the onset of the test example a tone, which served as a response cue, sounded for 50 ms. Subjects were instructed to make their yes-no response by pushing one of two keys approximately 200 ms after the tone. After responding, the test example was removed from the screen and subjects rated their confidence in their yes-no decision on a scale from one to six. The subjects were instructed to think back and rate their confidence in their yes-no decision at the time that decision was made, not their confidence in their yes-no decision at the time of the confidence rating itself. After the confidence rating the subject was provided feedback on his or her latency from the onset of the tone. After 2-3 s the experimenter initiated the next trial. The lags employed in the task (i.e. the times between the stimulus and response cue onset) were 0, 100, 200, 300, 400, 600, 800, 1000, 2000, 3000 ms. Subjects became familiar with the procedure in an initial practice phase consisting of four sessions which involved stimuli other than those employed in the test sessions.

Design

Each of the 50 experimental category names appeared in three trials in succession in each session. The order in which the category names were presented was randomized for each session. The three successive presentations of a category were not blocked with respect to truth value (whether the test example was or was not a member of the test category) nor lag. Rather, the truth value, test example and lag were selected randomly for each trial with three constraints. Across the 20 sessions each category name was paired in the first, second and third presentation position with one correct example and one incorrect example at each of the 10 lags. Each test example was paired once with its correct category name and once with an incorrect category name. Finally, no example of a category appeared twice in the experiment until each example of that category had appeared once.

Results

Lag-latency functions

A very small number of trials with latencies greater than 1 s were excluded from all analyses. The mean latencies of the subjects' yes-no responses were

no different in the lag-latency functions for different dominance and priming conditions so Table I presents the latencies averaged over dominance level and over the first, second, and third presentations of a category (subsequently referred to as positions 1, 2 and 3). Latencies are longest at the shortest lag and decrease by 60 ms on the average from the shortest to the longer lags, with most of the decrease occurring before subjects are performing above chance (below the intercept of the SAT function). These are typical lag-latency functions for SAT studies using the response-signal method, with close to constant latency at each lag. The small differences in latency across lags may reflect psychologically interesting processes, but we attribute the differences largely to response strategy. It seems reasonable at this point to assume that the small differences in latency across lags may be satisfactorily incorporated in the speed accuracy tradeoff functions by plotting accuracy as a function of mean processing time (lag plus mean latency).

TABLE I
Lag-latency functions for category-example verification (latencies in ms)

Subject	Lag (s)									
	0	0.1	0.2	0.3	0.4	0.6	0.8	1.0	2.0	3.0
DC	226	186	178	180	170	171	180	179	187	185
NV	293	279	253	234	209	197	187	189	201	199
PS	223	197	184	185	184	174	168	166	159	159
DS	220	199	198	170	163	157	166	172	180	179
Mean	240	215	203	192	182	175	175	177	182	180

SAT functions

The accuracy at each lag was assessed by d_T , a d' -type measure based on the pair of probabilities: (a) for a yes response under a hit condition and (b) its comparable false alarm condition, but adjusted for the effects of non-unit slope as indicated by the confidence judgement data (Reed, 1973). Since confidence judgements were obtained after the yes-no response, they may be based on more retrieved information than was available to the subject at the time of the yes-no response. Accordingly, the confidence judgement data were used *only* to estimate the relation between d' and the slope of the operating characteristics, in particular the regression of \ln slope on d'_S (the d' value obtained from the intersection of the operating characteristic and the negative diagonal). Then instead of entering the tables of Elliott (1964), which assume unit slope, we enter hit and false alarm rates for the yes-no response into the equivalent of a table which assumes that the true slope has the value indicated by the regression derived from the confidence judgement data. In a few cases, there was no computable d_T score for the two shortest lags because of heavy response bias (e.g. subjects always pushing the yes button). These points were excluded from

all analyses. In the few cases (only for subject DS) where there was no computable d_T score for the five longest lags because performance was perfect (infinite d_T), d_T was arbitrarily set at 4 (a value equivalent to the highest obtained d_T score for DS).

To determine the goodness of fit to the data of an exponential approach to a limit [Eq. (1)], a measure R^2 (percentage of variance accounted for) was employed, which adjusts for the number of free parameters with the equation

$$R^2 = 1 - \frac{\sum_{T=1}^N (d_T - \hat{a}_T)^2 / (N - k)}{\sum_{T=1}^N (d_T - \bar{d}_T)^2 / (N - 1)} \quad (2)$$

where N is the number of empirical points d_T , k is the number of parameters in the theoretical function, \hat{a}_T is the theoretical d_T value for condition T , and \bar{d}_T is the grand mean of the d_T (Reed, 1976). Reducing the number of degrees of freedom by the number of free (estimated) parameters is the conventional method of adjusting for differences in this factor across different models. However, our criteria for evaluating models place as much weight on consistency of relations among parameter estimates across conditions and subjects as on goodness of fit as measured by R^2 . Furthermore, we follow the criterion that if two models have approximately equal fit, the simpler model is preferred.

Priming

There is no evidence of any difference in either asymptotic strength or retrieval dynamics among the speed accuracy tradeoff curves for presentation positions 1, 2 and 3. Fitting an exponential approach to a limit in which asymptote (λ), retrieval rate (β) and intercept (δ) were held constant across the three curves for each subject yielded as good an R^2 as that provided by independent fits to the three curves (that is, 3 λ , 3 β s, and 3 δ s). Therefore, Figure 1 shows the speed accuracy tradeoff functions for category-example verification dynamics pooled over all three positions, yielding one curve per subject. The best fitting parameters obtained in fitting an exponential approach to a limit to these pooled curves are displayed in Table II.

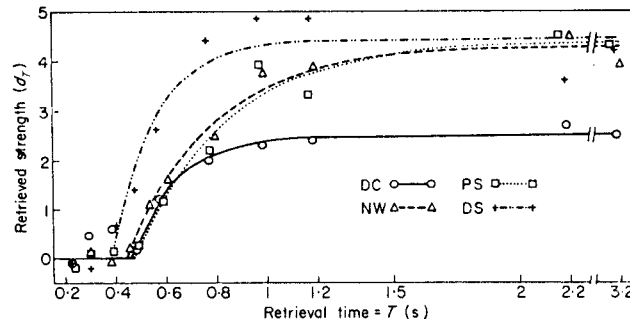


FIGURE 1. Speed accuracy tradeoff functions for category-example verification for each subject pooled over all positions. Lines are the best-fitting exponential function ($d_T = \lambda (1 - e^{-\beta(T-\delta)})$).

TABLE II

Best fit parameter estimates for category-example SAT functions
 $d_T = \lambda (1 - e^{-\beta(t-\delta)})$

Subject	λ	β	δ (ms)	R^2
DC	2.5	0.0062	465	0.91
NV	4.3	0.0032	445	0.90
PS	4.3	0.0029	469	0.95
DS	3.9	0.0060	375	0.97
Mean	3.8	0.0045	438	0.93

Advance presentation of the category alone 2 s before presentation of the example appears to completely eliminate the priming effect of prior category-example verification. Thus, engaging in a formal retrieval-decision process for one category-example pair is unnecessary to achieve maximum priming of the retrieval-decision process for another example from the same category.

The result is most compatible with a nodal activation explanation of previous priming effects. That is, repetition of a concept (in this case the category) results in faster accessing of the same concept node. A concept-example associative-activation model cannot be entirely ruled out, however, since accessing a node (i.e. reading the category name) may result in a spreading associative activation of links from that node, even without engaging in the process of retrieval of (example) associations from that (category) node (Collins and Loftus, 1975).

Dominance

The 30 correct category-example pairs for each category were subdivided into thirds representing the high, medium, and low (instance) dominance associations according to the Battig and Montague (1969) and Hunt and Hodge (1971) norms. The entire set of incorrect pairs was used to determine d_T values for the three dominance groups at each lag. The SAT functions obtained for high, medium, and low dominance pairs were fit with the exponential function assuming four different models for the parameters: (a) 3 λ s, 3 β s, and 3 δ s, (b) 3 λ s, 3 β s, and 1 δ , (c) 3 λ s, 1 β , and 3 δ s, and (d) 3 λ s, 1 β , and 1 δ . The first three models assume that dominance affects both asymptotic strength (λ) and retrieval dynamics (β and/or δ). The fourth model assumes that dominance affects only asymptotic strength of association and has no effect on the retrieval process.

The fit of the exponential function was very good and nearly equivalent for the four models, though model (d) fit best for three of the four subjects and model (a) fit best for one subject [average R^2 gave a modest advantage to model (a) with $\bar{R}^2 = 0.946$ vs. $\bar{R}^2 = 0.942$ for model (d)]. Thus, there is certainly no reason based on comparative goodness of fit to prefer any of the more complex models [(a), (b) or (c)] to the simpler model (d) which assumes that the dominant category-example associations differ from the less dominant associations only in asymptotic associative strength, with no difference in retrieval dynamics. The same conclusion emerges from examining the consistency of the parameter

TABLE III

Best fit parameter estimates for high (H), medium (M) and Low (L) dominance SAT functions $d_T = \lambda (1 - e^{-\beta(t-\delta)})$

Subject	λ_H	λ_M	λ_L	β	δ (ms)	R^2
DC	2.6	2.5	2.5	0.0048	462	0.96
NV	3.9	3.6	3.7	0.0028	436	0.94
PS	5.0	4.6	3.7	0.0033	473	0.92
DS	3.6	3.3	3.3	0.0058	377	0.95
Mean	3.8	3.5	3.3	0.0042	437	0.94

estimates across individuals for the four models. Models (a) and (b) show no consistent relation between dominance and the retrieval dynamics parameters. Model (c) showed some tendency for intercept (δ) to decrease with increasing dominance, but the overall fit of model (c) was actually slightly *poorer* than the fit of the simpler model (d) both on the average ($\bar{R}^2 = 0.939$ vs. $\bar{R}^2 = 0.942$) and for three of the four subjects. Table III shows the best fitting parameter values and goodness of fit for the 3 λ , 1 β , and 1 δ model in which asymptotic strength decreases from the high to medium to low dominance associations (with one small reversal for NV) but retrieval dynamics is invariant with dominance. Figure II shows the high and low dominance data (averaged over the four subjects) fit by the exponential function with the parameters averaged over the four subjects.

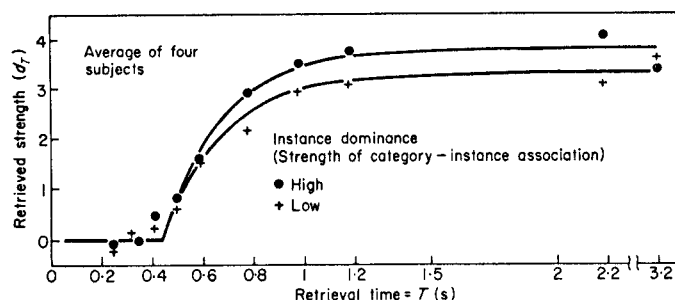


FIGURE 2. Speed accuracy tradeoff functions for category-example verification for high vs. low dominance examples. Data are d_T and T values for each lag averaged over the four subjects. Lines are best-fitting exponentials with parameters derived from averaging the best-fitting values for each subject.

Partitioning and microtradeoff

The distribution of a subject's latencies for a given lag typically have a rather small standard deviation with the response signal method, which minimizes the potential evils of averaging. Nevertheless, it would be desirable to assess whether such pooling of trials with various latencies (in a given lag condition) is distorting the form of the SAT function or systematically biasing its parameters. To investigate this, the data were partitioned into trials with short, medium and

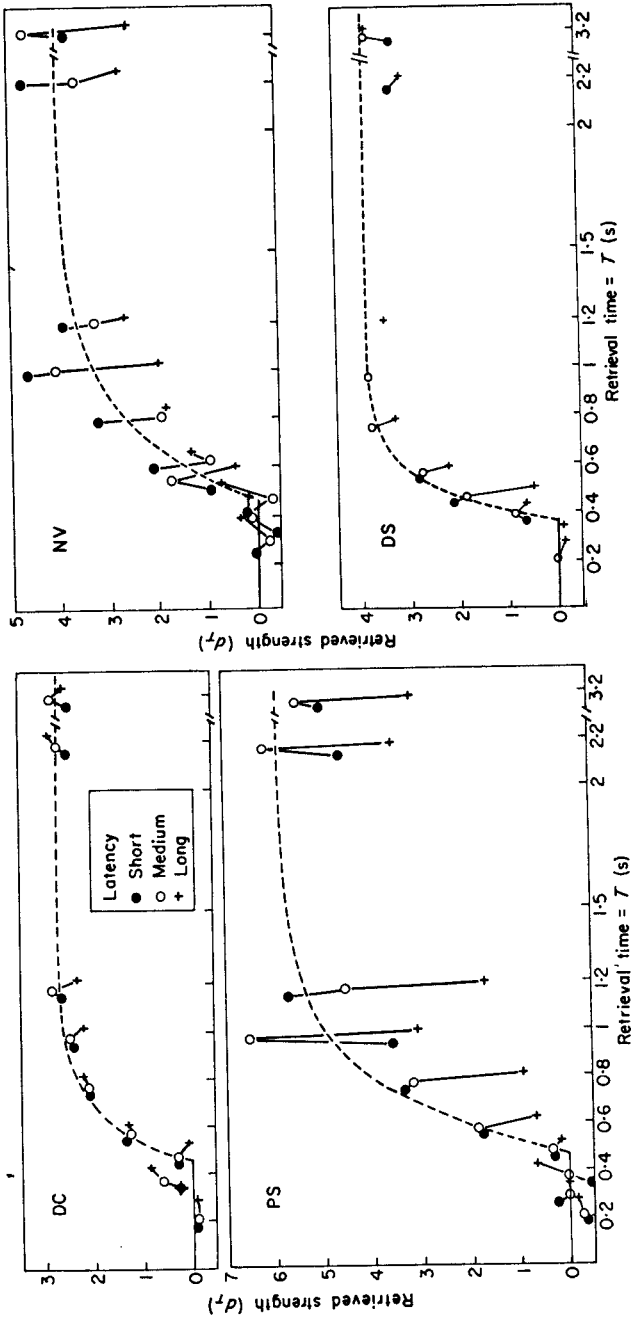


FIGURE 3. Three-point microtradeoff functions for short, medium, and long latency responses at each lag (connected by solid lines) with the best-fitting macrotradeoff function for the medium latency responses shown by dashed lines.

long RTs with separate SAT functions being plotted for each. In addition to these partitioned *macrotradeoff functions*, it is also of interest to examine the three-point *microtradeoff functions* for each lag condition. The simplest possible result would be for each microtradeoff function to lie on a common macrotradeoff function. This result was obtained in the only prior study of latency partitioning using the response signal method, a perceptual choice study (Schouten and Bekker, 1967).

The SAT functions for each subject for the shortest, medium, and longest latencies at each lag are shown in Figure 3. The microtradeoff functions for each lag are shown by connecting the three points at each lag by lines. Obviously, the microtradeoff functions for category-example verification do not lie on the same macrotradeoff function. In general, the longest latencies at each lag are less accurate than the medium and shortest latencies, rather than being more accurate as they should be if the microtradeoff functions were to lie on the macrotradeoff function.

TABLE IV

Best fit parameter estimates for short (S), medium (M) and long (L) latency SAT functions $d_T = \lambda (1 - e^{-\beta(t-\delta)})$

Subject	Asymptotes			Rate	Intercepts (ms)			R^2
	λ_S	λ_M	λ_L	β	δ_S	δ_M	δ_L	
DC	2.6	2.7	2.6	0.0056	415	454	514	0.94
NV	4.5	4.0	2.6	0.0032	388	448	458	0.91
PS	5.1	5.9	3.0	0.0035	432	465	585	0.92
DS	3.7	3.9	3.6	0.0077	336	368	486	0.96
Mean	4.0	4.1	3.0	0.0050	393	433	511	0.93

To get a more precise idea of the differences between the partitions, three separate macrotradeoff functions were fit to the three latency-partition curves for each subject.

Four models were fit to the curves (a) 3 λ , 3 β , 3 δ , (b) 3 λ , 1 β , 3 δ , (c) 3 λ , 3 β , 1 δ , and (d) 3 λ , 1 β , 1 δ . The best fit was obtained with the second model which assumes a common rate parameter (β) for all three curves for a given subject, but three different asymptotes (λ), and three different intercept (δ) parameters. This fit is shown in Table IV. In this fit the asymptote of the long-latency function is consistently lowest, and there is little difference between the short and medium functions. However, the intercept increases monotonically from short to medium to long latency SAT functions.

The greater variation over trials in the nature of the material being processed probably accounts for why the microtradeoff functions are so discrepant from the macrotradeoff function in this study, whereas they were not in the Schouten and Bekker (1967) study. In Schouten and Bekker (1967), there were only two different stimuli (lights one above the other), each of which was to be mapped on to one of two response buttons. In our study there were 1500 different

correct category-example pairs and the same number of different incorrect pairs. Whatever the reason for the intercept and asymptote differences among the latency-partitioned SAT functions, it was the case that the same exponential form of the function with virtually the same retrieval rate parameter provided as good a fit to the partitioned data as to the pooled data (β was slightly higher in the partitioned data). Thus, there is no reason to believe that the form of the SAT function or the retrieval rate parameter is being substantially distorted by pooling trials with different response latencies. Furthermore, the intercept parameter for the pooled data is close to and not systematically different from the average of the intercept parameters for the partitioned data.

Practice

All the trials for the first 10 sessions of the experiment were pooled for each subject. The same was done for the last 10 sessions. The empirical SAT functions for the first and last halves of the experiment were fit with four models: (a) 2 λ s, 2 β s, 2 δ s, (b) 2 λ s, 2 β s, 1 δ , (c) 2 λ s, 1 β , 2 δ s, and (d) 2 λ s, 1 β , 1 δ . No improvement in fit was afforded by the five and six parameter models over the four parameter model, which assumed constant retrieval dynamics (constant β and δ). Nor was there any consistent tendency for faster dynamics (greater rate and/or lower intercept) as a consequence of long-term practice in category-example verification within the SAT paradigm. For three of the four subjects there was not even much of a difference in the asymptotic strength (λ) value for the first vs. second half. The parameter estimates and goodness of fit for the best fitting (2 λ s, 1 β , 1 δ) model are shown in Table V.

TABLE V

*Best fit parameter estimates for SAT functions as a function of practice:
first vs. second halves of the experiment $d_T = \lambda (1 - e^{-\beta(t-\delta)})$*

Subject	λ_1	λ_2	β	δ (ms)	R^2
DC	2.6	2.5	0.0057	463	0.94
NV	3.5	4.7	0.0029	430	0.97
PS	4.3	4.5	0.0025	465	0.94
DS	3.5	3.7	0.0062	375	0.94
Mean	3.5	3.8	0.0043	433	0.95

Why there was any increase in the asymptotic strength of the category-example associations from the first to second half of the experiment is an open question. No association was ever tested more than once, but subjects might review their knowledge of the examples of the tested categories. In any case, a change in λ as a function of practice is of much less theoretical significance than the invariance of the memory retrieval dynamics. After all, we know category-example associations are subject to changes in stored strength as a function of learning and forgetting. Apparently, a few practice sessions are sufficient to produce SAT

functions which show no further practice effects in retrieval dynamics. Nor does repeated testing of categories produce any long-term facilitation of retrieval dynamics for examples in those categories.

Discussion

The conclusions of this study are primarily about method, though there are invariance results which, if replicated, have substantial theoretical importance. First, the present study demonstrates that even a relatively small difference in asymptotic strength can produce the 50-150 ms differences in categorization time for high- vs. low-dominance associations found in previous studies (see Smith *et al.*, 1974 for a review) without any necessary difference in retrieval dynamics. In fact, if error rate were constant and low, say 6% ($d_T = 3.10$, assuming equal hit and false alarm error rates) the RT difference between high and low dominance predicted by the theoretical SAT functions in Figure II is over 200 ms (840 vs. 1060 ms). If, as is typical, error rate is lower for the faster condition, for instance 5% error in the high dominance condition ($d_T = 3.28$) and 7% in the low dominance condition ($d_T = 2.94$), the predicted RT difference based on the theoretical SAT functions is only 40 ms (920 vs. 960). On the asymptotic sections of the SAT functions (approximately 3% errors in the high-dominant condition and 5% errors in the low dominant condition) very small changes in accuracy are associated with large changes in response time and a large range of RT differences may be obtained. Thus in a categorization task, or in other recognition memory tasks RT differences are consistent both with differences in retrieval dynamics or with asymptotic accuracy differences and do not actually constitute conclusive evidence for either (unless asymptotic accuracy is constant across conditions).

Specifically, previous RT studies of dominance effects on categorization time are consistent with the hypothesis that dominant associations merely have higher stored strength as well as with the hypothesis that there are genuine differences in retrieval dynamics across dominance levels. The present study suggests that there is no large difference in retrieval dynamics as a function of dominance and provides some support for the hypothesis that all example associations to a common category concept are retrieved in parallel at the same rate regardless of strength (dominance). Since this amounts to accepting the null hypothesis, this conclusion must remain tentative pending corroboration by other investigations. However, if retrieval dynamics is invariant with dominance, this finding contradicts all two-stage models, such as Juola *et al.* (1971) or Smith *et al.* (1974), which assume that dominant associations are retrieved faster because their retrieval more frequently terminates in the first stage. This finding also contradicts all serial search models of category-example verification such as Rosch (1973) or Anderson and Bower (1974) which assume that examples are retrieved in order of dominance.

Previous RT studies of priming are sufficient to show the existence of some priming of retrieval of an A-B association due to prior retrieval of an A-C association (though they cannot distinguish whether the priming effect is on

asymptotic strength or on dynamics). Thus, we accept the existence of such a priming effect. Our findings indicate that this effect is entirely due to prior activation of A, which occurs even if no retrieval of associations to A is required. Such a finding suggests that retrieval and priming of A-B associations are an automatic consequence of presentation of A (or B) and is not contingent upon the need to make a formal decision or response concerning any A-B association. This supports the spreading activation principle of retrieval and priming (Collins and Loftus, 1975).

Finally, the conclusions regarding the response-signal method of generating SAT functions are generally favourable, though it would simplify matters if microtradeoff functions lay on top of the macrotradeoff function. It is of practical importance that measured retrieval dynamics is invariant with practice in the SAT task, at least after a few practice sessions—a result which has also been found previously in a word-word paired associate task by Corbett (1977). It is also reassuring that the variation in intercept revealed by the partitioning analysis does not greatly distort the form of the pooled macrotradeoff SAT function. However, pooling should, and does, produce a small degree of initial positive acceleration in the vicinity of the intercept, in the form of a small bump. This is easiest to see in the d_T values at $T = 0.4$ s in Figure II, where this tendency is amplified still further by averaging over intercept variations across the four subjects. Examination of the individual SAT functions in Figure I reveals only a very small effect of this type. Furthermore, it must be emphasized that any such variation is probably in the subject, not in the response-signal method or any other method of obtaining RTs or SAT functions. There is a more complex model-fitting procedure by which any such intercept variation can be factored out of the parameter estimation and goodness of fit process by deliberately incorporating a submodel of this variation into all of the other models (a similar procedure is discussed in Wickelgren, 1977, p. 78). However, the magnitude of the effect in the present study appeared to be so small as not to justify this added complexity. In any case, obtaining a measure of the entire time course of retrieval dynamics, by some SAT method, provides a relatively direct way of answering such more detailed questions.

References

- ANDERSON, J. R. (1976). *Language, Memory, and Thought*. Hillsdale, N.J.: Erlbaum.
- ANDERSON, J. R. and BOWER, G. H. (1973). *Human Associative Memory*. New York: John Wiley & Sons.
- ASHCRAFT, M. H. (1976). Priming and property dominance effects in semantic memory. *Memory and Cognition*, **4**, 490-500.
- BATTIG, W. F. and MONTAGUE, W. E. (1969). Category norms for verbal items in 56 categories: a replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph*, **80** (3, Pt. 2).
- COLLINS, A. M. and LOFTUS, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, **82**, 407-28.
- COLLINS, A. M. and QUILLIAN, M. R. (1970). Facilitating retrieval from semantic memory: The effect of repeating part of an inference. *Acta Psychologica*, **33**, 304-14.
- CORBETT, A. T. (1977). Retrieval dynamics for rote and visual image mnemonics. *Journal of Verbal Learning and Verbal Behaviour*, **16**, 233-46.

- DOSHER, B. A. (1976). The retrieval of sentences from memory: a speed-accuracy study. *Cognitive Psychology*, **8**, 291-310.
- ELLIOTT, P. B. (1964). Tables of d' . In SWETS, J. A. (Ed.), *Signal Detection and Recognition by Human Observers*, pp. 651-84. London: Wiley.
- HUNT, K. P. and HODGE, M. H. (1971). Category-item frequency and category-name meaningfulness (m'): Taxonomic norms for 84 categories. *Psychonomic Monograph Supplements*, **4** (6, Whole No. 54).
- JUOLA, J. F., FISCHLER, I., WOOD, C. T. and ATKINSON, R. C. (1971). Recognition time for information stored in long-term memory. *Perception and Psychophysics*, **10**, 8-14.
- LOFTUS, E. F. (1973a). Category dominance, instance dominance, and categorization time. *Journal of Experimental Psychology*, **97**, 70-74.
- LOFTUS, E. F. (1973b). Activation of semantic memory. *American Journal of Psychology*, **86**, 331-7.
- LOFTUS, E. F. and SCHEFF, R. W. (1971). Categorization norms for fifty representative instances. *Journal of Experimental Psychology Monograph*, **91**, 355-64.
- PACHELLA, R. G. (1974). The interpretation of reaction time in information processing research. In KANTOWITZ, B. H. (Ed.), *Human Information Processing: Tutorials in Performance and Cognition*, pp. 41-82. New York: Wiley & Sons.
- REED, A. V. (1973). Speed-accuracy trade-off in recognition memory. *Science*, **181**, 574-6.
- REED, A. V. (1976). List length and the time course of recognition in immediate memory. *Memory and Cognition*, **4**, 16-30.
- RIPS, L. J., SHOBE, E. J. and SMITH, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behaviour*, **12**, 1-20.
- ROSCH, E. H. (1973). On the internal structure of perceptual and semantic categories. In MOORE, T. E. (Ed.), *Cognitive Development and the Acquisition of Language*. New York: Academic Press.
- SCHOUTEN, J. F. and BEKKER, J. A. M. (1967). Reaction time and accuracy. *Acta Psychologica*, **27**, 143-53.
- SMITH, E. E., SHOBE, E. J. and RIPS, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, **81**, 214-41.
- WICKELGREN, W. A. (1975). Dynamics of retrieval. In DEUTSCH, D. and DEUTSCH, J. A. (Eds), *Short-term Memory*. New York: Academic Press.
- WICKELGREN, W. A. (1976). Network strength theory of storage and retrieval dynamics. *Psychological Review*, **83**, 466-78.
- WICKELGREN, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, **41**, 67-85.
- WICKELGREN, W. A. and CORBETT, A. T. (1977). Associative interference and retrieval dynamics in yes-no recall and recognition. *Journal of Experimental Psychology: Human Learning and Memory*, **3**, 189-202.
- WILKINS, A. J. (1971). Conjoint frequency, category size and categorization time. *Journal of Verbal Learning and Verbal Behaviour*, **10**, 382-5.

Received 4 April 1977