PSYCHOPHYSIOLOGY

Vol. 20, No. 4
Printed in U.S.A.

Brain Potentials Related to Stages of Sentence Verification

IRA FISCHLER, PAUL A. BLOOM,
Department of Psychology, University of Florida

DONALD G. CHILDERS, SALIM E. ROUCOS,
Department of Electrical Engineering, University of Florida

AND NATHAN W. PERRY, JR.
Department of Clinical Psychology, University of Florida

ABSTRACT

Subjects were shown the terms of simple sentences in sequence (e.g., "A sparrow / is not / a vehicle") and manually indicated whether the sentence was true or false. When the sentence form was affirmative (i.e., "X *is a* Y"), false sentences produced scalp potentials that were significantly more negative than those for true sentences, in the region of about 250 to 450 msec following presentation of the sentence object. In contrast, when the sentence form was negative (i.e., "X *is not a* Y"), it was the true statements that were associated with the ERP negativity. Since both the false-affirmative and the true-negative sentences consist of "mismatched" subject and object terms (e.g., sparrow / vehicle), it was concluded that the negativity in the potentials reflected a semantic mismatch between terms at a preliminary stage of sentence comprehension, rather than the falseness of the sentence taken as a whole. Similarities between the present effects of semantic mismatches and the N400 associated with incongruous sentences (Kutas & Hillyard, 1980) are discussed. The pattern of response latencies and of ERPs taken together supported a model of sentence comprehension in which negatives are dealt with only after the proposition to be negated is understood.

DESCRIPTORS: Event related potentials (ERPs), Language, Sentence comprehension, Semantic memory, N400.

The use of event-related potentials (ERPs) has made it possible to study ongoing cognitive processes that may not be directly observable with behavioral measures. Distinctive ERP waveforms have recently been associated with a variety of cognitive states and processes, among them selective attention (Hink, Hillyard, & Benson, 1978), memory search (Gomer, Spicuzza, & O'Donnell, 1976), and preparation of responses (Deecke, 1977).

It now appears that unique components of ERPs may also be associated with the "meaning" of a verbal event—that is, its linguistic and semantic content. The "verbal event" may be a single word, a pair of words, a phrase, or an entire sentence. With the exception of a series of studies by Chapman and his colleagues on the connotative meaning of isolated words (e.g., Chapman, McCrary, Chapman, & Bragdon, 1978), research on ERP indicants of linguistic meaning has made use of manipulations of context. In these studies, a critical set of words is preceded by different types of contexts, and the ERPs to the critical words are compared across the different context conditions. Contexts which have been found to produce discriminable ERP patterns for target words include associates of the target words (Thatcher, 1977; Megela, Teyler, & Hesse, 1977; Vaughan, Sherif, O'Sullivan, Herrmann, & Weldon, 1982), phrases which bias a verb or noun interpretation of a homophone (i.e., "a pretty /roz/" vs "the boatman /roz/"; Brown, Lehmann, & Marsh, 1980), and a series of words from the same semantic category (e.g., animals) as the target, versus words from a different category (Polich, Vanasse, & Donchin, 1981).

Several studies have used sentences as contexts, with the target word as the final word of the sentence (e.g., Friedman, Simson, Ritter, & Rappin, 1975). The most relevant of these to the present research was reported by Kutas and Hillyard (1980).

The collection and analysis of the data reported here was greatly facilitated by Tony Arroyo and David Glicksberg.
Paul Bloom is now with CocaCola USA, Atlanta, GA 30322. Salim Roucos is with Bolt, Beranek and Newman, Inc., Cambridge, MA 02138.

Address requests for reprints to: Ira Fischler, Psychology Department, University of Florida, Gainesville, Florida 32611.

This is just summarizing what we already know about N400: there is a big N400 when someone reads a word that does not fit the sentence context

They found a late negativity in the ERP associated with words that completed a sentence in a semantically anomalous fashion (e.g., "He spread the warm bread with SOCKS"). The negativity was found for all subjects, and had a peak amplitude near 400 msec. This "N400" component was of greater amplitude, as measured by the difference between ERPs for the congruent and incongruent conditions, for "strongly anomalous" completions (e.g., "You can't make a silk purse out of a cow's CHAIR") than for "weakly anomalous" completions (e.g., ". . . out of a cow's SKIN"). Also consistent with a semantic interpretation of the N400 was its absence when the final word was semantically acceptable but physically anomalous (i.e., in a larger type print). The negativity was observable as early as 200 msec after target word onset, indicating that the retrieval of the meaning of the final word and its integration with the meaning of the preceding context occurs very rapidly (cf. Fischler & Bloom, 1980). Polich et al. (1981) also reported a negativity in the 200–400 msec range for semantically incongruent final words of sentences.

The possibility of ERP indicants of sentence meaning is of particular importance to the study of semantic memory and language. Sentences are composed of propositions, the smallest unit of knowledge which may be true or false. This *verifiability* is what distinguishes sentences from words and phrases, which cannot be considered true or false. Also, the meaning of a word is often not fixed, but determined by the sentence context in which it appears (R. Anderson & Ortony, 1975). Most of the significant efforts to model semantic memory in recent years have considered the proposition to be a basic unit of semantic knowledge (e.g., J. Anderson, 1976; Kintsch, 1974).

The negativity associated with anomalous sentences reported by Kutas and Hillyard (1980) and by Polich et al. (1981) suggests that a basic process in sentence comprehension is the monitoring of the consistency or validity of the propositions asserted by the sentence, with a negativity associated with the disruption of that process. The effect appears to be distinct from other types of disruption of expectancies, which typically produce an enlarged N2–P3 complex when *rare* or unexpected stimuli are presented (e.g., Näätänen, Hukkanen, & Jarvilehto, 1981). In contrast, anomalous words produce a negativity when anomalous completions are as likely as acceptable ones (Kutas & Hillyard, 1982; Polich et al., 1981). In fact, Fischler (Note 1) showed that words that were anomalous completions of sentence contexts produced longer lexical decision latencies than did words in acceptable contexts, even when two-thirds of the sentences were anomalous.

If the negativity in the ERPs is indeed an index of a process that monitors the consistency or validity of prose, then it should be possible to use it to evaluate models of the real-time construction of sentence meaning from its constituent words. The purpose of the present experiment was to attempt such an evaluation using a sentence verification task. In this task, subjects are shown simple sentences of the form, "X is a Y" and "X is not a Y." These are called ***affirmative*** and ***negative*** sentences, respectively. The sentences are judged as true or false, based on either prior semantic knowledge (e.g., the true-negative sentence, "Seven is not an even number;" Wason, 1959; or the false-affirmative, "A canary is a fish;" Collins & Quillian, 1969), or on episodic information such as a picture presented along with the sentence (e.g., "The star isn't above the plus;" Clark & Chase, 1972).

Latency to verify such sentences commonly shows a striking interaction between sentence form (affirmative-negative) and veracity (true-false), such that for affirmative sentences, *false* decisions are slower than *true*, while for negative sentences, it is the *true* decisions that are slower. Moreover, negative sentences are in general responded to more slowly than affirmative sentences. This pattern of results can be explained by assuming that a negative sentence such as "A sparrow is not a vehicle" is understood as a supposition ("A sparrow is a vehicle") and its denial. This may be represented as follows:

$$\{\text{false} \mid \{\text{is a} \mid \text{sparrow, vehicle}\}\}$$

(see Clark & Clark, 1977, pp. 100–113; Kintsch, 1974). Negative sentences require additional time to deal with the denial. Moreover, it is assumed that during verification subjects *first* compare the "inner" supposition to the relevant semantic information. A mismatch at this preliminary stage occurs for two kinds of sentences: false-affirmative, and true-negative. This produces the observed interaction in response latency.

In our experiment, subjects were asked to verify "class inclusion" statements such as "A sparrow is not a vehicle." Our first question was whether a negativity in the ERPs could be obtained for false-affirmative sentences such as "a robin is a truck," relative to true-affirmative sentences, under conditions used in a typical sentence-verification study: individual sentences were repeatedly presented, false and true statements were equally probable, and subjects were required to decide if the statements were true.

Second, if the pattern of latencies to the different types of sentences corresponds to that described

You can ignore a lot of this theoretical background. The important question is this: we know that there will be a bigger N400 for "a sparrow is a VEHICLE" than for "a sparrow is a "BIRD", but we don't know if that's because vehicle is unrelated to sparrow, or if it's the word "vehicle" makes the sentence bad. The way to test this is by using negative sentences. In "a sparrow is not a VEHICLE", vehicle is unrelated to sparrow but it makes a true sentence; in "a sparrow is not a BIRD", bird is related to sparrow but it makes the sentence bad. So if N400 is responding to sentence comprehension, then there should be a bigger amplitude (more negative) N400 for "a sparrow is not a BIRD". if N400 is just about word relationships, there should be a bigger amplitude (more negative) N400 for "a sparrow is not a VEHICLE".

above, then a test of the supposition-matching hypothesis can be obtained by comparing the ERPs associated with true-negative and false-negative sentences. If the ERP negativity is associated with the falseness of the sentence as ultimately understood, then false-negative sentences should display this negativity. In contrast, if the negativity reflects a semantic mismatch at the point where the implicit supposition is compared to semantic memory, then the ERP difference for true and false sentences should reverse for the negative-form sentences, with *true*-negatives showing the ERP negativity.

Finally, we speculated that the ERPs associated with negative-form sentences should be discriminable from those of the affirmatives, since they are more difficult to process than affirmatives (see Kutas, McCarthy, & Donchin, 1977). More importantly, such a difference should be observed at a longer latency than that for the true-false differences, since according to the model the negative is "set aside" until the inner supposition has been matched.

Method

Materials

Sentence subject-words ("S words") were 18 concrete nouns (e.g., "robin"), and sentence object-words ("O words") were the immediate (e.g., "bird") and superordinate (e.g., "animal") category names for those nouns. The complete hierarchy of words used is shown in Figure 1A. Categories and instances were selected from Battig and Montague (1969). Across categories at a given level in the hierarchy, the mean length, frequency of occurrence in English (Kučera & Francis, 1967), and number of syllables of the words were equated.

Four types of sentences were generated from the words in Figure 1A, by inclusion of verbs ("V words") of both affirmative ("*is a*") and negative ("*is not a*") forms. Examples of each sentence type are shown in Figure 1B. Sentences (1) and (2) are true, while sentences (3) and (4) are false. Sentences (1) and (3) are affirmative, while sentences (2) and (4) are negative. Note also that sentences (1) and (4) contain suppositions that match known facts (is a | robin, bird) while sentences (2) and (3) contain suppositions that do not (is a | robin, tree).

Use of all combinations of S and O words produced 36 true-affirmative (TA) sentences and 36 false-negative (FN) sentences. Equal numbers of true-negative (TN) and false-affirmative (FA) sentences were generated by rematching the S and O words such that the two terms of a sentence came from different superordinate categories. In the resulting set of 144 sentences, then, each type of sentence was equally likely. Note that across the four different types of sentences, the identical set of words was used as the critical O words.

Procedure

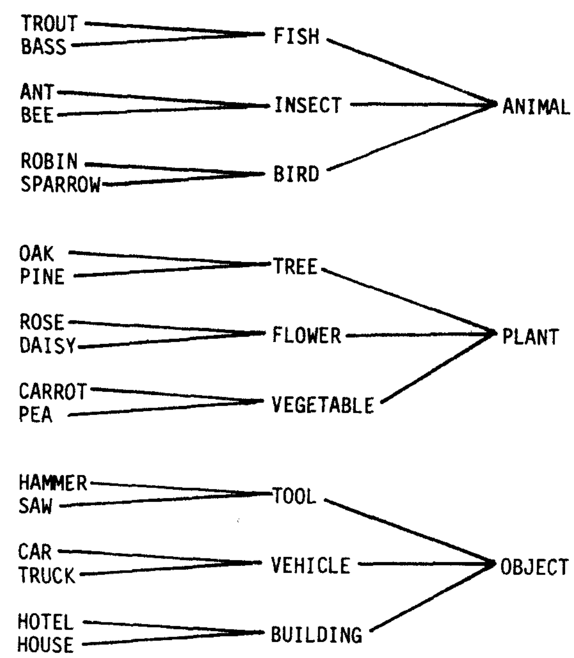
Eight male subjects from the Department of Electrical Engineering participated in the experiment. Three had previously served in ERP studies in our laboratory, but none had experience with the sentence verification task prior to the experiment.

Subjects sat upright in a ventilated Faraday shielded room. Sentences were presented on a TV monitor located about 80 cm from the subjects' eyes. A fixation box was displayed white-on-black near the center of the screen. The box subtended 0.8 deg vertically and 3.5 deg horizontally, and remained visible throughout the experiment. An asterisk appeared black-on-white in the left side of the box between trials, and remained visible for approximately 4 sec. Its disappearance marked the beginning of a new trial, and its reappearance marked the end of the recording epoch.

On each trial, the asterisk disappeared, and was followed 800 msec later by three successive frames, containing the S, V and O words respectively. The interval between frames was 800 msec, and the duration of each frame was 175 msec. The S and O frames subtended visual angles of 0.6 deg vertically, and 1.5 to 3.0 deg horizontally, depending on the particular sentence. All words appeared black-on-white justified to the left within the fixation box.

EEG was recorded from F_3, F_4, C_3, C_4, and C_z in the International 10/20 system, using nonpolarizing cup electrodes (Beckman) attached to the subject's scalp. Each

(A)



(B)

(1)	True, Affirmative (TA):	A robin / is / a bird.
(2)	True, Negative (TN):	A robin / is not / a tree.
(3)	False, Affirmative(FA):	A robin / is / a tree.
(4)	False, Negative (FN)	A robin / is not / a bird.

Figure 1. Stimulus words used in the experiment and their hierarchical relations (A), and examples of each of four types of sentences (B).

understanding these example stimuli will help understand the design of the experiment electrode location was referenced to the linked mastoids. Vertical and horizontal eye movements were recorded by two additional electrodes around the right eye, one placed nasally above the supraorbital ridge and the second behind the external canthus. Trials with EOG deflections from baseline greater than 15% of the amplitude of a typical eyeblink for a given subject were rerun at the end of the trial block.

The amplifiers had been specifically designed for evoked-potential recording (see Childers, 1977). They were essentially flat over the bandwidth of 1–50 Hz with a 50% attenuation at 1 Hz and 50 Hz. A 10 Hz, 50 μV signal was used to calibrate all channels at the start of each session. A train of photic stimuli was then presented to assess electrode contact and signal quality. EEG recording on each trial began 400 msec prior to the onset of the S word, and ended 2.1 sec after the onset of the O word. Data were digitized on-line at 125 samples/sec/location and stored for later analysis.

Two blocks of trials were presented, with all 144 sentences shown once in each block. Within blocks, the first and second halves contained an equal number of trials of each sentence type. Sentences were randomized within block for each subject. A short rest was provided after every 72 trials. The entire data collection period lasted about 40 min.

Subjects were instructed to classify each sentence as true or false by pressing one of two response switches with the forefinger and middle finger of the right hand. The assignment of switch to response category was counterbalanced across subjects. The required movement was small and required little force. Response latency was recorded. Subjects were free to respond as soon after the onset of the O word as they wished, but no explicit speed or accuracy instructions were given. Trials with incorrect responses or with latencies longer than 1.7 sec were rerun later in the session. Feedback about errors, eye movements, and long latencies was displayed at the end of the trial in which the problem occurred.

Results

Across all sentence types, 3% of the trials were rerun due to incorrect or long responses, while 6% of the trials were rerun due to eye movements or blinks.

Response Latency

Mean latency in msec for each subject for each sentence type was obtained. Negative sentences (TN and FN) were responded to 170 msec more slowly than affirmative sentences (TA and FA). For affirmatives, true sentences were responded to more quickly (969 msec, SE = 40) than false (1104 msec, SE = 37); but for negative sentences, the direction of the difference reversed, with false sentences (FN) producing faster responses (1182 msec, SE = 35) than true sentences (1236 msec, SE = 36). There were substantial individual differences in mean latency, ranging from about 900 msec across sentence types to about 1200 msec.

A 2 × 2 analysis of variance on these differences showed significant main effects of sentence form (affirmative-negative), $F(1/7) = 250.99$, $p < .01$, and sentence veracity (true-false), $F(1/7) = 9.09$, $p < .05$. The interaction was also significant, $F(1/7) = 31.75$, $p < .01$. Tests of the simple effects showed TA faster than FA, and FN faster than TN, both $p < .01$.

ERP Analyses

Averaged ERPs aligned to the onset of the O word of the sentence were obtained for each sentence type, location, and subject. Grand averages comparing all affirmative versus all negative sentences across subjects for each location were obtained. The ERPs for C_z are presented in Figure 2. The horizontal line here and in subsequent figures indicates zero voltage. Presentation of each term of the sentence produced a distinct ERP, the most prominent feature being a positive peak at about 200 msec. For several hundred msec prior to each frame, the ERP tended to be negative-going, resembling a CNV in anticipation of the upcoming stimuli (see Tecce, 1972). The overall pattern is quite similar to that for the successive words of the sentence contexts shown by Kutas and Hillyard (1980, Figure 1A).

Areas of difference between the waveforms for affirmative and negative sentences appear to be the slightly larger P2 peak following the V frame, the slightly more negative area for negative sentences just prior to the O frame, and the extended area of greater positivity for negative sentences from about 700–1200 msec following O frame onset. The appearance of the waveforms in the latter region suggests a slightly delayed positive peak for the negative sentences, The first two of these differences did not appear for a majority of subjects at any location. However, the very late difference was

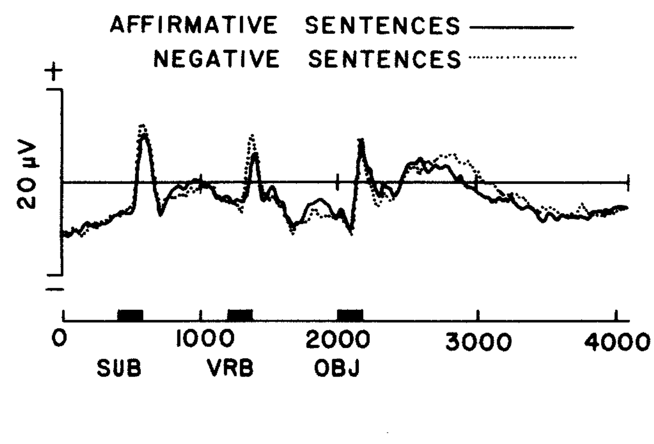
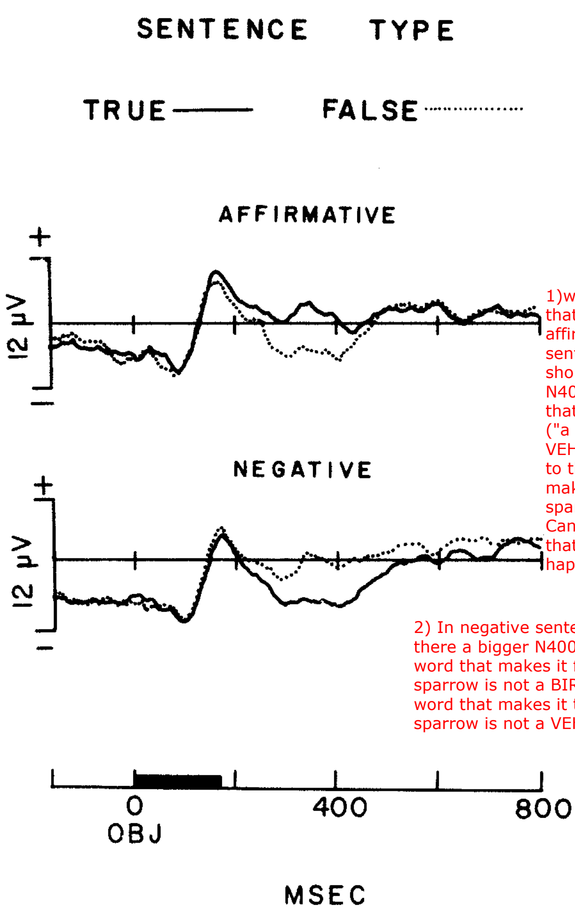


Figure 2. Averaged ERPs for all affirmative- versus all negative-sentence trials, 8 subjects, location C_z. Onset and duration of the sentence subject (SUB), verb (VRB), and object (OBJ) are indicated. The horizontal lines indicate zero voltage.

Don't worry too much about all these technical details and numbers. The only important thing is to look at the graph of the results (figure 3) and see if there was a bigger N400 for "a sparrow is not a VEHICLE" or for "a sparrow is not a BIRD" apparent for 6 of 8 subjects consistently across locations. The direction of difference was strongly reversed only for one subject, who also had the fastest response latency. The mean difference in μV over the interval from 800–1120 msec after O frame onset between affirmative and negative trials (Neg minus Aff) was calculated for each location for each subject. The mean differences (SEs in parentheses) were: C_z, 1.80 (1.01); C_3, 1.03 (0.81); C_4, 1.29 (0.98); F_3, 1.95 (0.83), $p < .05$; F_4, 1.15 (0.91). Individual *t*-tests showed that the differences fell short of significance, except for F_3. In all cases, however, the trend was for the ERPs to negative sentences to be slightly more positive than those for affirmative trials throughout this interval.

True versus False Sentences. Averaged ERPs for true versus false sentences were obtained for each subject for each location for affirmative and for negative sentences separately. The grand averaged ERPs across subjects for affirmative sentences for location C_z are shown in the top portion of Figure 3. The interval 160 msec prior to O frame onset was used to align the ERP for *false* trials to that for *true* trials for each subject. The adjustment averaged less than one μV. Following an N1–P2 complex in which little difference is seen, the ERP for the false-affirmative sentences is distinctly more negative than that for the true-affirmative sentences in the region from about 250 msec to 450 msec after O frame onset. The averages then reconverge.

The averaged ERPs for affirmative sentences are presented for the individual subjects in Figure 4. For all subjects except Sb. 4, the greater negativity for *false* trials around this region is apparent. The mean differences in μV for the ERPs to true and false sentences in the region between 320 and 480 msec after O frame onset, adjusted for any differences during the 160-msec prestimulus baseline, were obtained for each subject-location combination. Mean differences across subjects for each location are presented in Table 1. In each case, false sentences are associated with significantly more negative ERPs in the region around 400 msec; *t*-tests showed the difference to be significant for all locations. A one-way ANOVA on these difference scores with electrode location as the single factor



This result is the most important part. There are two things to check:

1)we already know that in the affirmative sentences, there should be a bigger N400 for the word that makes it false ("a sparrow is a VEHICLE") compared to the word that makes it true ("a sparrow is a BIRD"). Can we confirm that that really happened?

2) In negative sentences, is there a bigger N400 for the word that makes it false ("a sparrow is not a BIRD") or the word that makes it true ("a sparrow is not a VEHICLE")?

Figure 3. Averaged ERPs for true- versus false-sentence trials, all subjects, location C_z, from 400 msec before to 800 msec following the onset of the sentence object (O frame). The 160-msec period prior to O frame onset was used as a baseline to superimpose the ERPs for the False trials on those of the Trues (see text). The horizontal lines indicate zero voltage for True ERPs.

showed that the effect did not differ significantly in magnitude across locations, $F(4/28) = 1.31$.

Grand averaged ERPs across subjects for negative sentences are shown in the bottom portion of Figure 3. The averages are similar in most respects to those for the affirmative sentences, with one striking exception: whereas false-affirmative sentences produce the greater ERP negativity around 400 msec, it is the *true* sentences that show this greater negativity for sentences of negative form (TN vs FN).

Averaged ERPs for individual subjects are shown in Figure 5. As before, although the details of the waveforms differ, the greater negativity for TN sen-

Table 1

Mean differences in ERP magnitude (μV) between true and false sentences (true minus false) for affirmative and negative sentences in the region 320–480 msec following onset of sentence object (O frame)

	Mean Differences (SEs in Parentheses)				
Sentence Forms	C_z	C_3	C_4	F_3	F_4
Affirmative	3.82 (1.05)	2.87 (0.77)	2.95 (0.95)	2.30 (0.71)	2.83 (1.00)
Negative	−3.84 (0.53)	−2.18 (0.53)	−2.86 (0.40)	−1.95 (0.44)	−2.59 (0.66)

Note.—All differences are significant at $p < .01$ by *t*-test, except F_3, negatives, where $p < .05$.

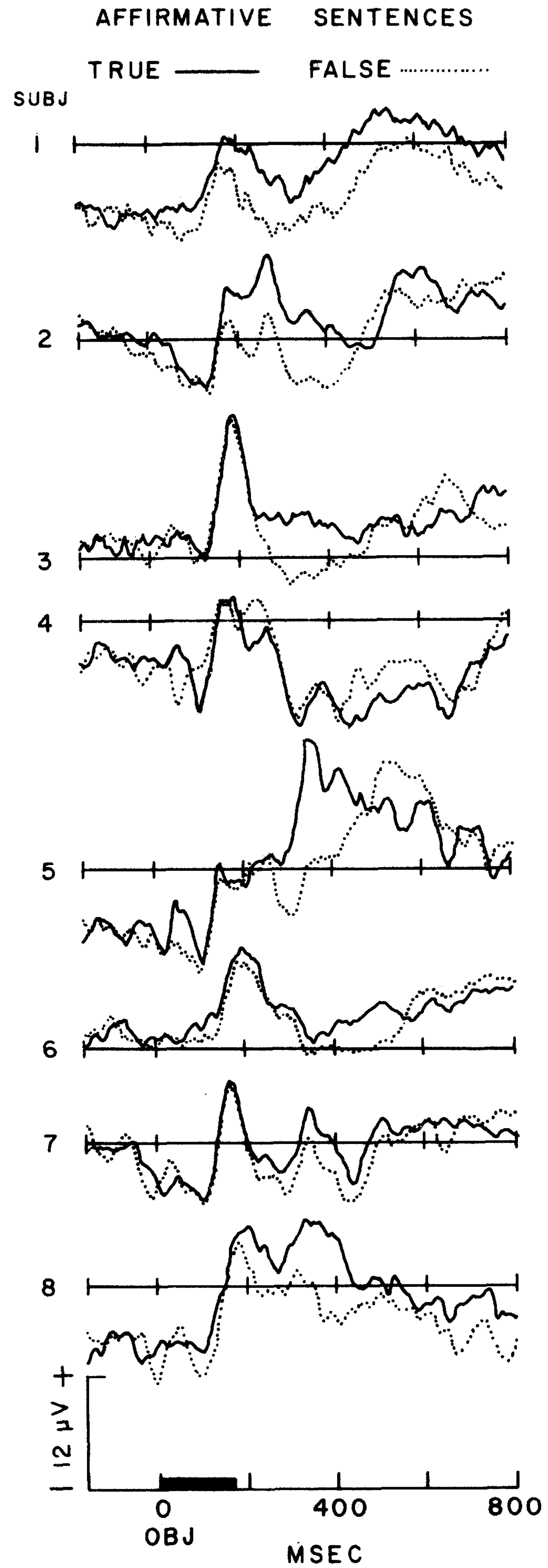


Figure 4. Averaged ERPs for true- versus false-sentence trials for individual subjects, affirmative sentence trials, 400 msec before to 800 msec after onset of sentence object (O frame). The 160-msec period prior to O frame onset was used as a baseline to superimpose the ERPs for False trials onto those for True trials. The horizontal lines indicate zero voltage for True ERPs.

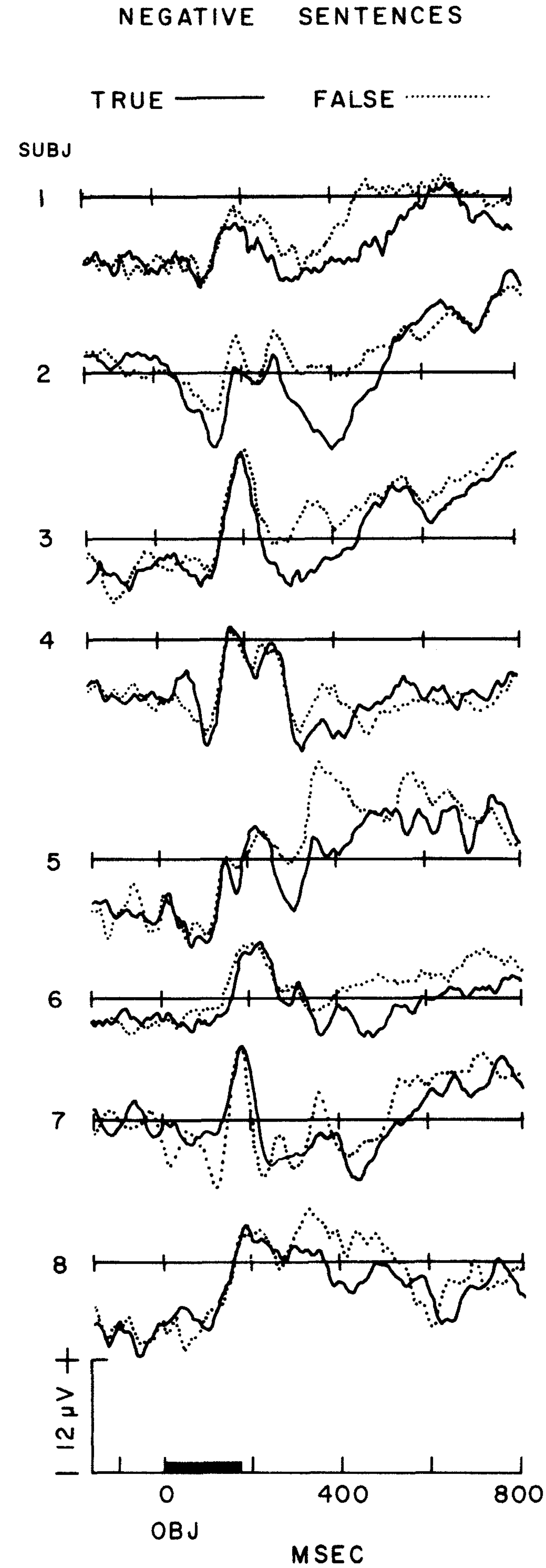


Figure 5. Averaged ERPs for true- versus false-sentence trials for individual subjects, negative sentence trials, 400 msec before to 800 msec after onset of sentence object (O frame). The 160-msec period prior to O frame onset was used as a baseline to superimpose the ERPs for False trials onto those for True trials. The horizontal lines indicate zero voltage for True ERPs.

tences can be seen in the ERPs for each subject. The adjusted difference scores were obtained as with the affirmative sentences, and the mean differences across subjects for each location are presented in Table 1. In each case, the ERPs for true sentences (TN) were significantly more negative than those for the false sentences (FN) in the region from 320 to 480 msec. The one-way ANOVA revealed a significant effect of location, $F(4/28) = 3.73$, $p < .05$. A Newman-Keuls test indicated that the effect was greater for C_z than the other locations, $p < .01$, which did not differ significantly from each other. As with the affirmative sentences, however, the trend was for the effect to be somewhat greater for the central than frontal loci, and greater for the right than left loci.

Discussion

The major results of the experiment can be summarized briefly. The response latencies replicated the basic pattern found in numerous sentence-verification studies: Overall, sentences of negative form produced slower responses; for affirmative sentences, "false" responses were slower, while for negative sentences, "true" responses were slower.

The pattern of ERPs associated with the various sentence types was also clear. For affirmatives, the averaged ERPs for false sentences were substantially more negative than those for the true sentences in the region from about 300 to 500 msec following onset of the final word of the sentence. The difference within the region 320–480 msec was significant at all locations, with a magnitude of 2–4 μV. In contrast, for negative sentences, it was *true* statements that were associated with the greater negativity in the ERPs. The magnitude and latency of the difference for negative sentences were similar to that for affirmatives. In both cases, the effect was greatest at C_z, and although other differences were not significant, the central and right locations tended to show a greater true-false difference than the frontal and left locations. The details of this topographic pattern were also obtained in an earlier version of the experiment, and we feel the pattern is genuine.

A direct comparison of the ERPs for affirmative and negative sentences showed no consistent differences between the two classes at least up to about 700 msec after onset of the final word of the sentence; there followed a nonsignificant tendency for the affirmative sentences to show a positive peak somewhat earlier than did negative sentences.

The pattern of behavioral and psychophysiological data is consistent with the model of sentence verification outlined earlier. Deciding whether or not a sentence is true can require a series of steps or stages. A first step is representing the sentence in a form compatible with the "target" information, in our case, propositions in semantic memory. For negative sentences such as "A robin is not a truck," this apparently means understanding the "inner" supposition without the negative—"A robin is a truck"—discovering that this is false, then subsequently reversing this decision when the negative is added. The mismatch between the subject and object of the false-affirmative sentences and the true-negative sentences during the initial stage of comparison apparently results in the observed difference in ERPs, compared to the "matching" suppositions for true-affirmatives and false-negatives. The difference appears quite early after the final word of the sentence is presented, compared to the overall time needed for the false decision.

It is worth emphasizing that the reversal in the true-false difference in ERPs between affirmative and negative sentences occurred independently of the decision and response required. This is strong evidence that the difference observed reflects a semantic process. When the difference in meaning is confounded with differential responses (e.g., "count the synonyms"), it is always possible that observed ERP effects are generated by events subsequent to the analysis of meaning. (See Posner, Klein, Summers, & Buggie, 1973, Exp. 4 for a clear demonstration of this; see also Sutton, 1979 for a discussion of the problem.)

There was little evidence suggesting that ERPs to "ultimately" false statements—FA and FN—were distinguishable from true statements. It may be that the neural process associated with the final decision does not produce differential ERPs. Alternatively, the variability in overall decision time for each sentence type within as well as across subjects may have obscured any such differences in the averaged data. We have done an extensive study of the ERPs associated with sentence verification with the epochs aligned to the initiation of the response, since the final decision of truth should just precede response execution. While the ERPs associated with true and false decisions were discriminable at better than chance accuracy, the basis of discrimination appeared to be tied to execution of the response itself (see Childers, Bloom, Arroyo, Roucos, Fischler, Achariyapaopan, & Perry, 1982).

We may refer to the greater negativity for FA and TN sentences as an "N400" to reflect the apparent direction and latency of the difference observed from TA and FN sentences. The appearance of the ERPs in Figure 3 certainly suggests a distinct component observed for FA and TN but not TA and FN sentences. Still, it could be argued that the

You can ignore most of this discussion; the important question is whether the results suggest that N400 is about word relationships, or N400 is about the overall sentence meaning.

difference reflects not a "genuine" negative component but the delayed appearance of late positive components associated with sentences that are more difficult to process (cf. Kutas et al., 1977). Several of the individual subject ERPs in Figures 4 and 5 at least suggest such a positive peak in the region of 400 msec for the faster conditions (TA and FN). While it is true that false affirmatives are slower than true affirmatives, and the reverse is true for negatives, there are good reasons for rejecting this interpretation of the ERP differences. First, the latency of the N400 appeared to be relatively stable across subjects (see Figures 4 and 5), in the face of great variation in response latency. More compelling is the comparison between affirmative and negative sentences in Figure 2. Overall, response latency for negatives is some 170 msec slower than for affirmatives, yet there is little trace of a difference between the ERPs to these two classes of sentences in the region around 400 msec. Finally, there is no evidence for a subsequent region in which the ERPs for FA and TN sentences "rebound" to become relatively more positive than those for TA and FN sentences (see Figure 3).

The occurrence of an N400-like feature for the FA and TN sentences under the present task conditions adds support to the reality and generality of late negativity in tasks involving discrepant semantic material. The present N400 is somewhat smaller and less sharply peaked than that reported by Kutas and Hillyard (1980, in press). This could be due, among other things, to variation in task strategy used by subjects when an explicit decision is required, or to the amount of repetition—both of individual words and of complete sentences—in our study. It is worth emphasizing that in spite of such repetition, significant ERP differences based on semantic content were obtained. The need for repetition of events for derivation of averaged ERPs can be at odds with the substantial changes in how a linguistic event may be processed with repetition. Requiring an overt semantic decision on each trial may maintain semantic processing despite such repetition (cf. Megela & Teyler, 1977).

A comparison of the topographic distribution of the N400 observed here and in Kutas and Hillyard (in press) also suggests a common origin of the two effects. In both cases, the trend was for right and central locations to show a larger effect than left and frontal. In both cases, the differences between locations were small, indicating a widespread spatial distribution of the effect; Kutas and Hillyard found that for the lateral comparisions, only a temporal-parietal location over Wernicke's area and its right hemisphere homologue produced an N400 significantly greater on the right than the left side. Their N400 was actually largest over P_z, but our study did not include that location, so a complete comparison is not possible.

It remains possible that the observed difference between matched and mismatched sentences is better characterized as an enhanced positivity for matching sentences (TA and FN) than an enhanced negativity for mismatched sentences (TN and FA). We have chosen the latter because of the form of the difference seen in Figure 3, and because of the analogy we see to other instances of negativity produced by semantic mismatching. We are presently testing the two alternatives directly using sentences whose truth value is unknown to the subjects.

The incongruity effect in these studies may be semantic, but it may not necessarily be propositional. In Kutas and Hillyard's studies, for example, the final anomalous word may be incongruent with the meaning of individual words or groups of words in the preceding context. Polich et al. (1981) reported that when the preceding context was a list of nouns from a single category, the negativity to a final word from a different category was "very much like" that elicited by the final anomalous word in their sentence context condition. There are other indications in the literature that a mismatch between the meaning of two sequentially presented words can produce relatively more negativity in the associated ERPs than is found in the ERPs for "matched" pairs, in the 200–500 msec region. Several studies using antonym pairs (e.g., hot-cold), which are incongruent on at least one dimension of meaning, have produced a similar pattern (see Thatcher, 1977, p. 440, Figure 6A; Vaughan et al., 1982, Figure 1). Sanquist, Rohrbaugh, Syndulko, and Lindsley (1980) reported an enhanced late positivity for the second word of a synonym pair, compared to unrelated pairs of words, but only when the task oriented subjects to word meaning, as opposed to orthographic or phonemic aspects of the stimuli (see their Figure 1). Finally, Boddy and Weinberg (1981) used a semantic categorization task similar in content to the affirmatives of our study, where a category question ("Is it a fish?") was followed by a positive or negative instance. The negative instances were associated with a slightly but reliably enlarged N1–P2 complex. The negative instances appeared also to produce greater late ERP negativity (see their Figure 2), but no analyses of this latter difference were presented.

The variation in tasks and in the conventions of ERP analysis across these studies makes it difficult to assess whether the late negativity reported for word pairs and for sentences is a single phenomenon reflecting a common process. It is also made difficult by the fact that even when words are presented without sentence structure, subjects who are oriented toward a semantic task may form implicit propositions attempting to relate the words. A possible test of the issue within our paradigm may be to present sentences like, "A table is a chair." Such false-affirmative sentences are associated with *faster* response times than FAs with distantly related or unassociated terms (Holyoak & Glass, 1975). If the N400 marks a mismatch between two isolated words, it should not be seen in such sentences; it should on the other hand occur if it marks the failure of a proposition placed in working memory to be consistent with corresponding information stored in semantic memory.

The sentence verification task has been criticized as being somewhat artificial (e.g., Tanenhaus, Carroll, & Bever, 1976; Evans, 1982). Wason, for example (e.g., 1965, 1980), has pointed out that in everyday speech, negatives are used to deny a supposition that is reasonable (e.g., "a whale is not a fish") or to point out exceptions within a context (e.g., "Senator Smith isn't a man"). Such "plausible" negatives are more easily verified than those like "A robin is not a vehicle." Also, the pattern of latencies in such tasks can be affected by factors such as whether the negative statement can be converted into a positive one (i.e., "Seven isn't even" into "Seven is odd;" see Carpenter & Just, 1975). But it seems less important to determine what kind of process or materials is most "natural" than to be able to specify how such factors influence the way people construct a representation of the meaning of a sentence, and then try to understand it in terms of what is already known. Our results suggest that evidence from ERPs can play an important role in testing alternative models of sentence comprehension in this larger sense.

REFERENCES

Anderson, J.R. *Language, memory and thought*. New York: Lawrence Erlbaum, 1976.

Anderson, R.C., & Ortony, A. On putting apples into bottles—a problem of polysemy. *Cognitive Psychology*, 1975, *7*, 167–180.

Battig, W.F., & Montague, W.E. Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monographs*, 1969, *80*, No. 3, 1–46.

Boddy, J., & Weinberg, H. Brain potentials, perceptual mechanisms and semantic categorization. *Biological Psychology*, 1981, *12*, 43–61.

Brown, W.S., Lehmann, D., & Marsh, J.T. Linguistic meaning-related differences in evoked potential topography: English, Swiss-German, and imagined. *Brain & Language*, 1980, *11*, 340–353.

Carpenter, P.A., & Just, M.A. Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review*, 1975, *82*, 45–73.

Chapman, R.M., McCrary, J.W., Chapman, J.A., & Bragdon, H. Brain responses related to semantic meaning. *Brain & Language*, 1978, *5*, 195–205.

Childers, D.G. Evoked responses: Electrogenesis, models, methodology, and wavefront reconstruction and tracking analysis. *Proceedings of the IEEE*, 1977, *65*, 611–626.

Childers, D.G., Bloom, P.A., Arroyo, A.A., Roucos, S.E., Fischler, I.S., Achariyapaopan, T., & Perry, Jr., N.W. Classification of cortical responses using features from single EEG records. *IEEE Transactions on Biomedical Engineering*, 1982, *29*, 423–438.

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

Clark, H.H., & Clark, E.V. *Psychology and language*. New York: Harcourt Brace Jovanovich, 1977.

Collins, A.M., & Quillian, M.R. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1969, *8*, 240–247.

Deecke, L. Functional significance of cerebral potentials preceding voluntary movement. In D. Otto (Ed.), *Multidisciplinary perspectives in event-related potential research*. Washington, DC: Environmental Protection Agency (publication EPA-600/9-77-043), 1977. Pp. 87–92.

Evans, J. St. B.T. *The psychology of deductive reasoning*. London: Routledge & Kegan Paul, 1982.

Fischler, I., & Bloom, P.A. Rapid processing of the meaning of sentences. *Memory & Cognition*, 1980, *8*, 216–225.

Friedman, D., Simson, R., Ritter, W., & Rappin, I. The late positive component (P300) and information processing in sentences. *Electroencephalography & Clinical Neurophysiology*, 1975, *38*, 255–262.

Gomer, F.E., Spicuzza, R.J., & O'Donnell, R.P. Evoked potential correlates of visual item recognition during memory scanning tasks. *Physiological Psychology*, 1976, *4*, 61–65.

Hink, R.F., Hillyard, S.A., & Benson, P.J. Event-related potentials and selective attention to acoustic and phonetic cues. *Biological Psychology*, 1978, *6*, 1–16.

Holyoak, K.J., & Glass, A.L. The role of contradictions and counterexamples in the rejection of false sentences. *Journal of Verbal Learning and Verbal Behavior*, 1975, *14*, 215–239.

Kintsch, W. *The representation of meaning in memory*. Hillsdale, NJ: Lawrence Erlbaum, 1974.

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

Kutas, M., & Hillyard, S.A. Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 1980, *207*, 203–204.

Kutas, M., & Hillyard, S.A. The lateral distribution of event-related potentials during sentence processing. *Neuropsychologia*, 1982, *20*, 579–590.

Kutas, M., McCarthy, G., & Donchin, E. Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. *Science*, 1977, *197*, 792–795.

Megela, A.L., & Teyler, T.J. Event-related potentials associated with linguistic stimuli: Semantic versus lower-order effects. In D. Otto (Ed.), *Multidisciplinary perspectives in event-related potential research.* Washington, DC: Environmental Protection Agency (publication EPA-600/9-77-043), 1977. Pp. 267–270.

Megela, A.I., Teyler, T.J., & Hesse, G.W. ERP response decrement and recovery as a function of stimulus type and scalp location. *Physiology and Behavior*, 1977, *19*, 15–22.

Näätänen, R., Hukkanen, S., & Jarvilehto, T. Magnitude of stimulus deviance and brain potentials. In H. Kornhuber & L. Deecke (Eds.), *Progress in brain research* (Vol. 54). *Motivation, motor and sensory processes of the brain: Electrical potentials, behavior and clinical use.* Amsterdam: Elsevier/North-Holland, 1981. Pp. 337–342.

Polich, J., Vanasse, L., & Donchin, E. Category expectancy and the N200. *Psychophysiology*, 1981, *18*, 142. (Abstract)

Posner, M.I., Klein, R., Summers, J., & Buggie, S. On the selection of signals. *Memory & Cognition*, 1973, *1*, 1–12.

Sanquist, T.F., Rohrbaugh, J.W., Syndulko, K., & Lindsley, D.B. Electrocortical signs of levels of processing: Perceptual analysis and recognition memory. *Psychophysiology*, 1980, *17*, 568–576.

Sutton, S. P300—thirteen years later. In H. Begleiter (Ed.), *Evoked brain potentials and behavior.* New York: Plenum, 1979. Pp. 107–126.

Tanenhaus, M.K., Carroll, J.M., & Bever, T.G. Sentence-picture verification models as theories of sentence comprehension: A critique of Carpenter and Just. *Psychological Review*, 1976, *83*, 219–225.

Tecce, J.J. Contingent negative variation (CNV) and psychological processes in man. *Psychological Bulletin*, 1972, *77*, 73–108.

Thatcher, R.W. Evoked potential correlates of hemispheric lateralization during semantic information processing. In S. Harnad, R. Doty, L. Goldstein, J. Jaynes, & G. Krautheimer (Eds.), *Lateralization in the nervous system.* New York: Academic Press, 1977. Pp. 429–448.

Vaughan, J., Sherif, K., O'Sullivan, R.L., Herrmann, D.J., & Weldon, D.A. Cortical evoked responses to synonyms and antonyms. *Memory & Cognition*, 1982, *10*, 225–231.

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

Wason, P.C. The contexts of plausible denial. *Journal of Verbal Learning and Verbal Behavior*, 1965, *4*, 7–11.

Wason, P.C. The verification task and beyond. In D.R. Olson (Ed.), *The social foundations of language and thought.* New York: Norton, 1980. Pp. 28–46.

REFERENCE NOTE

1. Fischler, I. *Studies of the effects of sentence contexts on lexical decision speed: Implications for models of word recognition and reading.* Paper presented at the meeting of the Midwestern Psychological Association, Detroit, April 1981.

Announcements

Eleventh Annual Scientific Meeting Psychophysiology Society (London)

From December 14th through 16th, 1983, the Eleventh Annual Meeting of the Psychophysiology Society will be held at Charing Cross Medical School, London. Details for the submission of papers may be obtained from: Dr. D. Papakostopoulos, Burden Neurological Institute, Stoke Lane, Stapleton, Bristol BS16 1QT, England. Details of conference registration may be obtained from: Dr. J. Gruzelier, Department of Psychiatry, Charing Cross Hospital Medical School, St. Dunstans Road, London W6 8RP, England.

Fifteenth Annual Meeting Biofeedback Society of America

From March 23rd through 28th, 1984, the Fifteenth Annual Meeting of the Biofeedback Society of America will be held at the Regent Hotel, Albuquerque, New Mexico. Workshops will be offered through the entire meeting and exclusively on the last day. The theme for the 1984 program is *Biofeedback in Perspective: Fifteen Years of Development.*

The deadline for submission of papers is October 1, 1983 (Program Chairman: Martin R. Ford, Institute of Living, Box 1929, Hartford, CT 06101). Further information may be obtained from: Biofeedback Society of America, 4301 Owens Street, Wheat Ridge, CO 80033, or phone 303/422-8436.