

## And and or, or the Comprehension of Pseudoimperatives<sup>1</sup>

FREDERICK J. SPRINGSTON AND HERBERT H. CLARK

Stanford University, Stanford, California 94305

It was proposed that the *or* in pseudoimperatives such as *Sit down or I'll scream* is represented in comprehension as equivalent to an *and* plus a negative element. Subjects were timed as they drew deductions from pseudoimperatives containing either *and* or *or* in such problems as: If the sign says "Don't flip the switch or the fan goes on" and the fan did not go on, then did you flip the switch? The latencies supported the proposal. They suggested that the pseudoimperative *p and q* is represented as the biconditional ( $p \leftrightarrow q$ ) and that *p or q* is represented as  $\sim(p \leftrightarrow q)$ , not as  $(\sim p \leftrightarrow q)$  or as  $(p \leftrightarrow \sim q)$ , which are its logical equivalents. The results also showed that so-called "invited inferences" take no longer to draw than inferences that constitute the literal meaning of these pseudoimperatives.

Several linguists (Fodor, 1970; Karttunen, 1971; Lakoff, 1970; Zwicky, 1971) have recently been concerned with the logical properties of certain complex sentences. Among the questions they have examined are: When is one sentence a paraphrase of another? When does one sentence logically presuppose or imply another? And how should the meaning of certain words and constructions be represented in logical form? In the present paper we will be concerned with the latter question for sentences containing *and* and *or*. These sentences, we will argue, can logically be represented in several alternative ways. The question is, therefore, which of these alternatives most closely fits how people normally represent these sentences once they have comprehended them. In order to choose among these alternatives, we will examine how long it takes people to carry out simple deductions based on these sentences.

The sentences we will examine are the so-called pseudoimperatives, as illustrated in (1):

<sup>1</sup> This research was supported in part by a Predoctoral Fellowship from the National Science Foundation and by United States Public Health Service Research Grant MH-20021 from the National Institute of Mental Health. We thank William P. Banks, Eve V. Clark, Russell Revlis, and Lance Rips for their helpful comments on earlier drafts of the paper.

- (1) a. Sit down and I'll scream.  
b. Sit down or I'll scream.

As Jespersen (1940), G. Lakoff (1966), and R. Lakoff (1971) have pointed out, these sentences, despite their superficially imperative form, are not imperatives at all: they are conditionals. Both sentences express the notion that if one thing happens (or fails to happen), then something else will inevitably happen. Hence sentences (1a) and (1b) can be paraphrased, respectively, by the explicit conditionals in (2a) and (2b):

- (2) a. I'll scream if you sit down.  
b. I'll scream if you don't sit down.

The most important point these paraphrases illustrate is that the *and* in (1a) entails a positive conditional, whereas the *or* in (1b) requires a negative in the conditional clause. Our main objective in this study is to determine the role this negative plays in the comprehension of pseudoimperatives.

The pseudoimperatives in (1), however, are often interpreted in everyday contexts as something more than the simple conditionals in (2). As many investigators have observed, *if-then* conditionals are often construed as if they were *if and only if* biconditionals (Noordman, 1972; Taplin, 1971). Geis and Zwicky

(1971) have argued that this comes about by a process they call "conditional perfection" or "invited inference." To take an example, *If you take out the garbage, I'll give you a dollar* is normally interpreted as entailing the sentence *If you don't take out the garbage, I won't give you a dollar*. But, as Geis and Zwicky note, this is a logically invalid inference, for the former conditional does not necessarily imply that I will not give you a dollar if you do not take out the garbage; nevertheless, in most situations, the former conditional "invites" the inference anyway. If (1a) and (1b) also invite such inferences, they would be paraphrased not simply by the conditionals in (2), but by the biconditionals in (3):

- (3) a. I'll scream if and only if you sit down.  
 b. I'll scream if and only if you don't sit down.

The biconditional in (3a), for example, expresses the primary or literal meaning of (1a)—namely, *I'll scream if you sit down*—plus the invited inference—*I won't scream if you don't sit down*.

But if people do make invited inferences when interpreting pseudoimperatives, this raises an interesting question about *or*. Note that the paraphrase of *Sit down or I'll scream* given in (3b) is logically equivalent to (3c) and (3d):

- (3) c. I won't scream if and only if you sit down.  
 d. It is not the case that: I'll scream if and only if you sit down.

These three paraphrases of (1b) differ in where the negative required by *or* is attached: In (3b) it is attached to the clause *you sit down*; in (3c) it is attached to the clause *I'll scream*; and in (3d) it denies the biconditional connection of the two positive propositions *you sit down* and *I'll scream*, but is directly attached to neither. The psychological question is, which of these three paraphrases—(3b), (3c), or (3d)—corresponds most accurately to how people represent the logical content of *Sit down or I'll scream* when it is treated as biconditional.

An additional question of interest whenever pseudoimperatives are interpreted as biconditionals is the following: Is there any difference in how quickly people can make use of the primary meaning of these pseudoimperatives as contrasted with their invited inferences?

To examine these questions we required subjects to solve simple deduction problems containing pseudoimperatives that were invariably to be construed as biconditionals, and we compared the latencies in solving the problems against predictions derived from a model of negation developed by Chase and Clark (1972), Clark (1970, 1973), and Clark and Chase (1972). In selecting sentence content, we attempted to avoid sentences that could only be interpreted either as promises (*Sit down and I'll give you a dollar*) or threats (*Sit down and I'll scream*), since these lead to some very odd pseudoimperatives (e.g., *Sit down or I'll give you a dollar*). For this reason we chose sentences about the causal relation between the position of a switch and the activity of a fan. Thus, some subjects were shown a "cover" sentence like *Flip the switch or the fan goes on* followed by the information *You flipped (or didn't flip) the switch*, and they were to answer the question "Did the fan go on?" The remaining subjects received the same cover sentences, but these were followed by information about whether the fan was on or not; these subjects were to answer the question "Did you flip the switch?" Schematically, there were six possible pseudoimperative cover sentences:<sup>2</sup>

- A. *p* and *q*  
 B. not-*p* and *q*  
 C. *p* and not-*q*  
 D. *p* or *q*  
 E. not-*p* or *q*  
 F. *p* or not-*q*,

<sup>2</sup> There are of course two other possible cover sentences *not-p* and *not-q* and *not-p* or *not-q*. Since cover sentences with double negatives were anticipated to be very difficult to comprehend, only a partial sample was included.

where  $p$  refers to the propositional content *Flip the switch*, and  $q$ , to the propositional content *the fan goes on*. These were followed by four possible "object" sentences:  $p$ ,  $not-p$ ,  $q$ , and  $not-q$ . The subjects receiving the object sentences  $p$  and  $not-p$  will be said to be in the  $p$  condition; those subjects receiving  $q$  and  $not-q$  will be said to be in the  $q$  condition. The Clark and Chase model of negation can be extended to make quite straightforward predictions about cover sentences A, B, and C. But it makes quite different predictions about D depending on how it is interpreted. If the negative is attached to  $p$ , then D is equivalent to B (i.e.,  $not-p$  and  $q$ ) both in content and latency predictions; if the negative is attached to  $q$ , then D is equivalent to C (i.e.,  $p$  and  $not-q$ ); but if the negative is attached to the biconditional relation itself, it is equivalent to the denial of A, which we will represent as  $not(p$  and  $q)$ . In our preliminary discussion, we will be concerned only with cover sentences A through D.

The basic notion behind the Clark and Chase model is that when the subject is required to compare one sentence against another (as he is in the present problems) this comparison will be easy when the two sentences match, or are congruent, in their underlying representations, but it will be difficult, taking more time, when the two sentences do not match, or are not congruent. Consider sentence

A,  $p$  and  $q$ . If this is followed by the object sentence  $p$ , the subject must compare  $p$  against the first conjunct of the cover sentence. Since the latter  $p$  matches the former  $p$ , this comparison is easy. But if sentence A is followed by the object sentence  $not-p$ , then there is a mismatch since  $not-p$  does not match  $p$ , and the comparison becomes more difficult. Thus, the Clark and Chase model would predict that, given cover sentence A, object sentence  $p$  should be faster than  $not-p$ . That is, A should produce what we will call a "positive  $p$ -difference." In contrast, for cover sentence B,  $not-p$  and  $q$ , the object sentence  $p$  does not match  $not-p$ , whereas the object sentence  $not-p$  does. Thus, given sentence B,  $p$  should be slower, not faster, than  $not-p$ , producing a "negative  $p$ -difference." Table I summarizes the "difference" predictions for sentences A, B, and C for both pairs of object sentences. Thus, there should be positive  $p$ -differences (denoted by pluses) for A and C, but a negative  $p$ -difference for B (denoted by a minus); similarly, there should be positive  $q$ -differences for A and B, but a negative one for C.

Precisely the same analysis can be applied to the three possible methods for dealing with cover sentence D,  $p$  or  $q$ . If D is equivalent to  $not-p$  and  $q$ , then it ought to have a pattern of pluses and minuses like B; if D is equivalent to  $p$  and  $not-q$ , then it ought to pattern like C. But if D is equivalent to  $not(p$  and  $q)$ , then

TABLE 1  
PREDICTIONS FROM THE CLARK AND CHASE MODEL FOR THE DIFFERENCES IN LATENCY BETWEEN THE  
OBJECT SENTENCES  $p$ ,  $not-p$ ,  $q$ , AND  $not-q$ , FOR COVER SENTENCES A, B, C, AND D

Cover sentence	Difference in object sentences	
	$not-p - p$	$not-q - q$
A. $p$ and $q$	+	+
B. $not-p$ and $q$	-	+
C. $p$ and $not-q$	+	-
D. $p$ or $q$		
(1) = $not-p$ and $q$	-	+
(2) = $p$ and $not-q$	+	-
(3) = $not(p$ and $q)$	+	+

there ought to be pluses in both columns, making it (in this sense) similar to A. These predictions are also shown in Table 1, and they constitute the main method for distinguishing among the three possible representations for *p* or *q*. There is one additional prediction that can be derived from the previous work on negation. Negatives have almost invariably been found to take longer to process than affirmatives (see Clark, 1970, 1973). If sentence D, *p* or *q*, contains an implicit negative, then it should take longer to deal with overall than A, *p* and *q*, which by hypothesis does not contain such a negative. The increased difficulty of D should be similar to the increased difficulty of B and C, which contain explicit negatives that also require additional processing.

The experiment to be reported, then, was designed to test these predictions. In addition, it included the more complex sentences E and F, which do not yield readily to unequivocal predictions from the Clark and Chase model. Thus, the experiment was more generally designed to help us construct a more complete model of the processes of comprehension and deduction in these problems. Such a model can be constructed on the basis of the relative overall latencies of sentences A through F as well as the relative sizes of *p*- and *q*-differences.

### METHOD

The subjects were required to solve a series of problems consisting of a cover sentence followed by an object sentence, and they were timed as they made their response. The 24 problems used consisted of one of the six cover sentences A through F followed by one of the four object sentences *p*, *not-p*, *q* or *not-q*. For the cover sentences, *p* was *Flip the switch*, *not-p* was *Don't flip the switch*, *q* was *the fan goes on*, and *not-q* was *the fan doesn't go on*; for the four object sentences, *p* was *You flipped the switch*, *not-p* was *You didn't flip the switch*, *q* was *The fan went on*, and *not-q* was *The fan didn't go on*. The problem consisting of *not-p* or *q* followed by *not-p*, for example, was seen in the following format, with the object sentence always centered below the cover sentence:

"Don't flip the switch or the fan goes on."  
You didn't flip the switch.

The subjects in the *p* condition solved only the 12 problems with object sentences *p* and *not-p*. They were told to treat the top sentence as a description of how a switch and fan worked and the bottom sentence as an action that had been carried out. On all problems these subjects were to answer the implicit question "Did the fan go on?" The subjects in the *q* condition solved the other 12 problems. Given a problem like "*Don't flip the switch and the fan goes on*" and *The fan didn't go on*, they were always to answer the implicit question "Did you flip the switch?" Each problem was typed on a card and was viewed by the subject at 14 in. in a tachistoscope.

On each trial the subject pressed the middle "ready" button of a hand-held three-button panel with either thumb, and half a second later, a problem appeared in the tachistoscope. On solving the problem the subject was to indicate his answer by pressing the left or right button with his left or right thumb to designate the answer "yes" or "no". The assignment of "yes" and "no" to the left and right buttons was counterbalanced across subjects. The latency of the subject's response on each trial was measured in hundredths of a second from the appearance of the problem to the first press of an answer button at which time the problem disappeared from view. Each subject was given 9 blocks of 12 trials, each block consisting of an individual shuffling of the 12 problems he was to receive. Since the problems were found to be quite difficult, the first four blocks were considered practice and were not considered in the analyses of the results. With short breaks between blocks, the total session lasted 45 min.

The subjects were 48 Stanford University undergraduates fulfilling a course requirement. After the elimination of 12 subjects who made more than 40% errors on at least one problem over the last five blocks, there were 36 subjects, half of whom were in the *p* condition and half in the *q* condition. The subjects were urged to treat the sentences naturally, yet to work as quickly as possible while maintaining a high degree of accuracy. The experimentally correct answers were those of the biconditional interpretation; hence, half the problems were to be answered "yes" and half "no." The subjects were not told anything about how to interpret the sentences; they were simply corrected whenever they made an error, so they could adjust their interpretations of the problems if they found they were in conflict with the experimentally determined correct response. Below, however, we will offer evidence that subjects did construe the cover sentences as biconditionals without any prompting.

### RESULTS

The main results of interest are the latencies and errors for the 24 problems, shown in

Tables 2 and 3, respectively. Before these are examined, however, there are three preliminary issues to be resolved. First, since the model of interest is concerned with latencies, what is the appropriate measure of latency? Comparing the mean latencies of correct responses and the number of errors on each problem indicates that problems with longer latencies tend to elicit more errors too: the rank-order correlation of errors and latencies by problem was .89 ( $p < .01$ ) in the  $p$  condition and .80 ( $p < .01$ ) in the  $q$  condition. In contrast, the rank-order correlation of latencies and errors by subject was  $-.27$  for the subjects in the  $p$  condition and  $-.01$  for those in the  $q$  condition, both negligible correlations. These correlations suggest that hard problems caused subjects to take longer and lower their accuracy at the same time; so if they had kept their accuracy high on all problems, they should have taken even longer on the more difficult problems. It therefore seemed appropriate to take the median of the subject's five attempts at a problem while counting errors as infinite times; this would attenuate the problem of differential accuracy across problems. Table 2 shows the mean of the median latencies for each problem, and Table 3 shows the percentage of errors over the last five attempts at each problem by each subject. The results to be examined, however, are quite robust, for the mean latencies of the correct responses lead to the same qualitative results as the mean median latencies actually used.

The second issue is whether artifacts were introduced by eliminating those 12 subjects for which medians could not be calculated for each problem, or by eliminating the first four blocks as practice. The answer appears to be "no" on both counts. First, the error rates of the rejected subjects correlated highly with those of the analyzed subjects, with rank-order correlations by problems of .88 and .64 for the subjects in the  $p$  and  $q$  conditions, respectively. This suggests that the relative difficulty of each problem was very similar for the analyzed and rejected subjects. Second, the error rates for

the practice trials correlated very highly with those of the analyzed trials, with rhos of .95 and .88 for the subjects in the  $p$  and  $q$  conditions, respectively. Thus, relative difficulty as indicated by error rate changed very little, if any, from the practice to the analyzed trials.

The third issue is whether subjects naturally construed the pseudoimperatives as biconditionals, and there is evidence that they did. Each cover sentence has a literal meaning and an invited inference, and therefore half the problems for each such sentence require the subject to make a genuine inference and half, an invited inference. On cover sentence A, for example, the problems with object sentences  $p$  and  $not-q$  require genuine inferences; those with the object sentences  $not-p$  and  $q$  require invited inferences. If the subject was loath to draw the invited inference, he should therefore make many more errors on the invited inference problems—especially during practice. But this does not occur. On the practice problems of all subjects combined (both analyzed and rejected), there were 30 and 31% errors on problems demanding genuine and invited inferences, respectively; the percentages are 17 and 16% when one considers only those contrasts orthogonal to the processing parameters discussed below. On the easiest cover sentence, A, these percentages were 2 and 2%, respectively, showing virtually no errors in either instance. This evidence suggests that when subjects made errors they did so because the problems were difficult, not because they failed to construe the cover sentence as biconditionals.

The latencies for problems with cover sentences A, B, and C were in good qualitative agreement with the predictions made by the Clark and Chase model. As shown in Table 1, the  $p$ -differences for A, B, and C were predicted to be +, -, and +, respectively, and this was consistent with the +487, -1440, and +1936 msec differences actually found. (The reliabilities of these and all subsequent differences were evaluated using the Wilcoxon matched-pairs signed-ranks test; there are 18 signed-

TABLE 2  
MEAN MEDIAN LATENCIES (IN MSEC) FOR THE 24 COVER SENTENCES BY OBJECT SENTENCE CONDITIONS

Cover sentence	Object sentences					
	<i>p</i>	not- <i>p</i>	(not- <i>p</i> - <i>p</i> )	<i>q</i>	not- <i>q</i>	(not- <i>q</i> - <i>q</i> )
A. <i>p</i> and <i>q</i>	2116	2603	+487	2024	2693	+669
B. not- <i>p</i> and <i>q</i>	4492	3052	-1440	3704	5914	+2210
C. <i>p</i> and not- <i>q</i>	2955	4891	+1936	3790	3248	-542
D. <i>p</i> or <i>q</i>	3976	5685	+1709	3981	5306	+1325
E. not- <i>p</i> or <i>q</i>	5325	4993	-322	5197	5707	+510
F. <i>p</i> or not- <i>q</i>	4669	5817	+1148	4021	4729	+708

ranks in every case and the reported *T* always refers to the smaller sum of signed-ranks.) The *T*s for these three *p*-differences were all found to be less than 5,  $p < .001$ . Similarly, the *q*-differences for A, B, and C were predicted to be +, +, and -, respectively. This was confirmed with the actual *q*-differences of +669, +2210, and -542 msec, respectively ( $T=1$ ,  $p < .001$ ;  $T=1$ ,  $p < .001$ ; and  $T=31$ ,  $p < .01$ ). For the critical cover sentence D, the *p*-difference was positive (+1709 msec,  $T=5$ ,  $p < .001$ ), and the *q*-difference was also positive (+1325,  $T=37$ ,  $p < .02$ ). Thus, cover sentence D showed a pattern of *p*- and *q*-differences similar to cover sentence A and dissimilar to B and C. This suggests that the negative in D is not attached to either *p* or *q*, but rather is attached to the biconditional connection itself. The overall mean latencies for cover sentences A, B, C, and D (collapsed over object sentences) were

also consistent with the Clark and Chase model. This model predicted that A would take less time than B, C, and D, since A does not contain a negative, whereas B, C, and D do. In agreement with this prediction, the mean latencies were 3259, 4290, 3721, and 4737 msec, respectively, with the first significantly less than each of the others in both the *p* condition and the *q* condition taken separately ( $T=0$ ,  $p < .001$ , in all cases). In short, the results are qualitatively consistent with the thesis that *p* or *q* is represented with a negative that is attached to neither of its component propositions, but rather is represented as something like *not* (*p* and *q*).

The pattern of latencies in Table 2, however, contains much more information than is accounted for in the simple predictions made in Table 1, for there are wide variations in the *p*- and *q*-differences as well as in the overall latencies for each cover sentence. A more

TABLE 3  
PERCENTAGE OF ERRORS FOR THE 24 COVER SENTENCE BY OBJECT SENTENCE CONDITIONS

Cover sentence	Object sentences			
	<i>p</i>	not- <i>p</i>	<i>q</i>	not- <i>q</i>
A. <i>p</i> and <i>q</i>	0	0	2	2
B. not- <i>p</i> and <i>q</i>	3	2	1	24
C. <i>p</i> and not- <i>q</i>	4	10	2	3
D. <i>p</i> or <i>q</i>	7	17	8	23
E. not- <i>q</i> or <i>q</i>	13	14	19	17
F. <i>p</i> or not- <i>q</i>	8	19	11	13

explicit process model of what subjects are doing in this task might account for many of these additional factors. Thus, drawing on the previous work of Clark (1970, 1973), Clark and Chase (1972), Just and Carpenter (1971), and Trabasso, Rollins, and Shaughnessy (1971), we will develop such a model and compare it to the pattern of latencies in Table 2.

#### A MODEL FOR COMPREHENSION AND DEDUCTION

The model to be developed consists of four stages: (1) a representation stage in which the subject codes the cover sentence; (2) a representation stage in which the subject codes the object sentence; (3) a comparison stage in which the subject compares the object sentence against the appropriate component of the cover sentence and computes the answer; and (4) a response stage in which the answer is converted into an overt response. For the present we will develop the model only for cover sentences A, B, C, and D; afterwards, we will examine the more complex sentences E and F.

At Stage 1 the subject must encode the cover sentence. The model represents each cover sentence as being encoded as a biconditional. The first column of Table 4 shows the biconditional representations for A, B, C, and D;

these four representations differ in where the negative is attached, in agreement with the qualitative results just examined. In addition, we assume, consistent with previous findings (see Clark and Chase, 1972; Just and Clark, 1973), that the cover sentences containing negatives (B, C, D) will take longer to encode than those without (A). The time increment consumed by encoding the negative will be denoted as  $n$ . *A priori* we did not predict differences among the three negative cover sentences, B, C, and D. Although there is some evidence for such differences, for the present, we will assume uniform encoding difficulty for these negatives.

At Stage 2 the subject must encode the object sentence. This, we assume, is represented as a simple proposition or its negative, as schematized for the  $p$  condition in the second column of Table 4. Half of the object sentences contain a negative, and we might have assumed these would take longer to encode. The latencies, however, show no evidence of this extra time. If they did take longer, the absolute  $p$ -difference should be larger for A than for B, and the absolute  $q$ -difference larger for A than for C. The former prediction is disconfirmed by a result significantly in the wrong direction ( $T = 25$ ,  $p < .01$ ), and the latter prediction, by a result that was small and unreliable. One possibility here is that the difficulty in encoding negatives is in deter-

TABLE 4  
INITIAL REPRESENTATIONS OF COVER AND OBJECT SENTENCES AND VALUES OF THE ANSWER INDEX AT THREE POINTS IN THE COMPARISON PROCESS

Cover sentence	Representations		Initial value of Answer Index	Change effected by Rule 1	Change effected by Rule 2	Final value of Answer Index
	Cover sentence	Object sentence				
A. $p \leftrightarrow q$	$p$	$q$	$q$			$q$
B. $\sim p \leftrightarrow q$	$\sim p$	$q$	$q$	$\sim q$		$\sim q$
C. $p \leftrightarrow \sim q$	$p$	$\sim q$	$q$	$\sim q$		$q$
D. $\sim(p \leftrightarrow q)$	$p$	$\sim q$	$\sim q$	$\sim(\sim q)$	$q$	$q$
	$\sim p$	$q$	$\sim q$	$\sim(\sim q)$	$q$	$q$

mining where the negative element is attached. In the cover sentences there is a choice of where to attach the negative and so this should take time; but the object sentences consist only of a single member, so there is no choice and the time expended should be minimal. In any case the present model makes no provisions for encoding differences in the object sentence.

In representing the problem to himself the subject must also set up an Answer Index, a register that will eventually contain the correct answer. The Answer Index must therefore contain the conjunct of the cover sentence that is queried by the problem. The subjects in the  $p$  condition, for example, are asked "Did the fan go on?" and their Answer Index, then, would initially contain the right-hand conjunct of the cover sentence. Any negative attached to the queried conjunct is also entered in the Answer Index. Moreover, whenever the cover sentence has an *or* in it, as in D, a negative is directly assigned to the Answer Index in conformity with the assumed representation of D: *not (p and q)*. The initial values of the Answer Index for A, B, C and D in the  $p$  condition are shown in Table 4.

At Stage 3 the subject must compute the correct answer. The subject does this by comparing the object sentence (say, *not-p*) against the conjunct with the same propositional content in the cover sentence (e.g.,  $p$  in A, C, or D) and noting whether the two conjuncts match in polarity (i.e., in the presence or absence of a negation indicator). The result of this comparison operation is implemented as Rule 1.

Rule 1: If the Stage 2 code does not match the polarity of the corresponding conjunct of the Stage 1 code, then add a negative to the Answer Index.

Computations for Stage 3 are completed with the subsequent application of Rule 2.

Rule 2: If the Answer Index contains two or more negatives, delete two of them.

As shown in Table 4, Rule 1 has the effect of changing  $q$  to  $\sim q$  for two of the eight problems and of changing  $\sim q$  to  $\sim(\sim q)$  for two other problems. Rule 2 then notes that the latter two Answer Indices contain double negatives, that is  $\sim(\sim q)$ , and it reduces them simply to  $q$ . The result of these two operations is that the Answer Index now contains the correct answer for each problem, as shown in the final column of Table 4. This algorithm, then, is logically sufficient for the task. The important assumption is that Rules 1 and 2 consume time whenever they are required to carry out changes: Rule 1 consumes  $m$  increment of time for every mismatch, and Rule 2 consumes  $r$  increment of time for every double negative reduction.

What evidence is there for Rules 1 and 2? We have already examined the evidence for Rule 1. It states that the comparison stage should take longer whenever the object sentence does not match the corresponding conjunct of the cover sentence. But this was just the basis for the predictions shown in Table 1 and confirmed in Table 2. Rule 2 is proposed to account for some of the variation in the absolute values of the  $p$ - and  $q$ -differences. As Table 4 shows, this rule predicts that the  $p$ -differences should be larger (in absolute value) for C and D than for A and B. The actual  $p$ -differences are in good accord with this prediction: The  $p$ -differences for C (1936 msec) and D (1709 msec) are significantly greater than the  $p$ -differences for A (487 msec;  $T = 14, p < .001$ ) and for B (1440 msec;  $T = 47, p < .05$ ). Unexpectedly, however, the  $p$ -difference was larger in absolute value for B than for A ( $T = 25, p < .005$ ), and this remains unexplained. Rule 2 also predicts that the  $q$ -differences for B (2210 msec) and D (1325 msec) should be larger in absolute value than for A (669 msec) and C (542 msec), and here the data are in full agreement. The  $q$ -differences for B and D are significantly greater than for either A ( $T = 18, p < .001$ ) or C ( $T = 14, p < .001$ ). In sum, the  $p$ - and  $q$ -differences, when evaluated ordinally, corroborate Rules 1 and 2



as components of the comprehension model.

At Stage 4 of the model, the subject takes the answer as finally computed—namely, the final value of the Answer Index, as shown in Table 4—and transduces it into the physical response of pressing the “yes” or “no” button. Since the number of “yes” and “no” responses was equated, no response-bias effect was postulated.

To summarize the latency predictions of the model, we have postulated three time increments:  $n$ , the time required to encode a negative;  $m$ , the time required to effect Rule 1; and  $r$ , the time required to effect Rule 2. Least squares estimates of these three parameters (utilizing only cover sentences A, B, C, and D) were calculated to allow us to present the

predictions of the model in graphic form, and they are:  $n$ , 1416 msec;  $m$ , 636 msec; and  $r$ , 1308 msec. In addition, this model requires a base time  $t_0$ , estimated at 2070 msec, which is the time necessary to execute all operations common to every condition. In accordance with the rationale limned out in the description of Stages 1 and 3 of the model, these parameter estimates are associated with particular stimulus conditions, and the results are displayed in Fig. 1a. These predicted values should be compared with the obtained values in Fig. 1b. The model is intended only to describe the central tendencies of the latency distributions; specific claims about the underlying distributions do not appear justifiable at this point. But the degree of success obtained

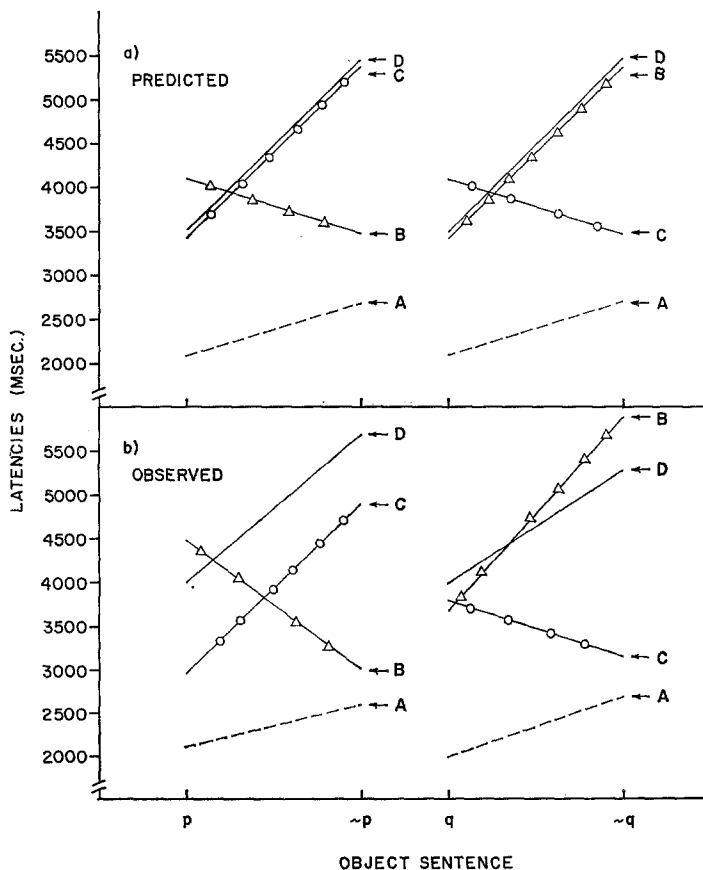


FIG. 1. Predicted and observed latencies for problems containing the cover sentence A, B, C, or D and the object sentence,  $p$ ,  $\sim p$ ,  $q$ , or  $\sim q$ .

TABLE 5  
INITIAL REPRESENTATIONS OF COVER AND OBJECT SENTENCES AND VALUES OF THE ANSWER INDEX AT  
THREE POINTS IN THE COMPARISON PROCESS

	Representations		Initial	Change	Change	Final
	Cover sentence	Object sentence	value of Answer Index	effected by Rule 1	effected by Rule 2	value of Answer Index
<i>Method 1</i>						
E	$\sim(\sim p \leftrightarrow q)$	$p$	$\sim q$	$\sim(\sim q)$	$q$	$q$
		$\sim p$	$\sim q$			$\sim q$
F	$\sim(p \leftrightarrow \sim q)$	$p$	$\sim(\sim q)$		$q$	$q$
		$\sim p$	$\sim(\sim q)$	$\sim(\sim(\sim q))$	$\sim q$	$\sim q$
<i>Method 2</i>						
E	$(p \leftrightarrow q)$	$p$	$q$			$q$
		$\sim p$	$q$	$\sim q$		$\sim q$
F	$(p \leftrightarrow q)$	$p$	$q$			$q$
		$\sim p$	$q$	$\sim q$		$\sim q$

in verifying the predictions of the model with nonparametric statistics demonstrates considerable concurrence among subjects in using the hypothesized algorithm to solve these problems.

Now we turn to cover sentences E and F, which contain two negatives each. Table 5 outlines two methods subjects could have used to deal with E and F. The first is simply the same model as we elaborated above, and it leads to the predicted values shown in Fig. 2a. The second method is a kind of "conversion" model for handling these sentences (Clark, 1970, 1973; Clark and Chase, 1972). In it the subject is assumed to "convert" the double negatives into a simpler form at the initial encoding stage, and this makes them easier to deal with at the later comparison stage. Thus, the subject would code *not-p or q* and *p or not-q* both as equivalent to the simpler pseudoimperative *p and q*. That is, when he reads the sentence *Don't flip the switch or the fan goes on*, he transforms it into *Flip the switch and the fan goes on*—and then proceeds as before. Thus, after this initial encoding difference, the model for E and F are identical to the one outlined previously. There is considerable precedent for this type of initial "conversion" or "pre-processing." Wason and Jones (1963) note, for

example, that in verifying sentences like *Eight is not even* some of their subjects reported converting *not even* to *odd* before attempting to decide on truth or falsity. Similar examples of conversions are given in Clark and Chase (1972), Trabasso, Rollins, and Shaughnessy (1971), and Young and Chase (1971). If we assume in this second method that the initial "conversion" consumes an increment of time  $r$ —that is, it is equal to the increment  $r$  for the reduction of double negations *after* comparison—then this model predicts the latencies shown in Fig. 2b.

The actual latencies, shown in Fig. 2c, are consistent with a mixture of the regular and conversion methods. First, the two methods converge in their predictions on cover sentence F in the  $p$  condition and on cover sentence E in the  $q$  condition, and these agree fairly well with the data. The  $p$ -difference for F should be positive, and it is. Furthermore, its absolute value (1148 msec) should be approximately the same as for A and B (mean, 964 msec) and smaller than for C and D (mean, 1822 msec); it is not significantly different from the former mean, and in the latter case the predicted difference approaches conventional significance levels ( $T = 53, p < .08$ ). As for cover sentence E, its  $q$ -difference should be positive, and it is.

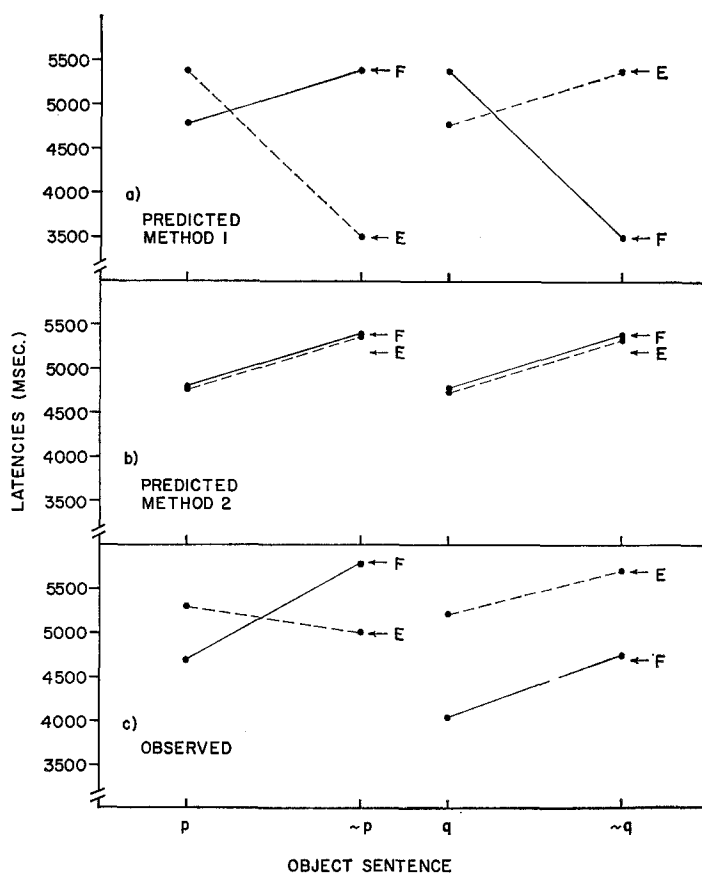


FIG. 2. Latencies predicted by Methods 1 and 2 and observed latencies for problems containing the cover sentence E or F and the object sentence  $p$ ,  $not-p$ ,  $q$ , or  $not-q$ .

Furthermore, its absolute value (510 msec) should be the same as that for A and C (mean, 606 msec) and smaller than that for B and D (mean, 1768 msec), and both of these predictions are confirmed, with the latter difference significant with  $T = 38, p < .02$ .

The two methods, however, diverge on the remaining predictions. The  $p$ -difference for cover sentence E is compatible with a mixture of the two strategies, for its value is negative, but too small (322 msec) to be comparable to the value predicted by the "conversion" model (1944 msec). The  $q$ -differences for cover sentence F is consistent with the conversion method. Its value is positive, as it should be, and its absolute value (708 msec) is close to the 636 msec predicted for it. There is, however,

an unpredicted difference in the overall latency levels between sentences E and F in the  $q$  condition ( $T = 16, p < .001$ ), again suggesting some differences in the coding time of the two types of negatives. In any case, this analysis suggests that the subjects in the  $p$  condition were using some mixture of the regular and conversion methods on sentences E and F, while the subjects in the  $q$  condition were using mainly the conversion method.

#### *Genuine versus Invited Inferences*

There is very little evidence to promote the belief that invited inferences are more difficult than genuine inferences. This is evident whether one (a) simply contrasts all problems requiring invited inferences against those re-

quiring genuine inferences, or (b) considers only those contrasts orthogonal to the four parameters of the present model (i.e., contrasts of the  $q$ -differences of A, C and B with the  $p$ -differences of A, B, and C, respectively). By the first method, invited inferences are 323 msec *faster* than genuine inferences; by the second, invited inferences are a small unreliable 74 msec slower than genuine inference. Moreover, this finding of no difference is not confined to the subjects with low error-rates on which our latency analyses were based. Considering the error-rate for all subjects over all trials, we find invited inferences favored over genuine inferences 15–19% by the first method, and 7–9% by the second method. Clearly, these data offer no support for the hypothesis that invited inferences are more difficult to draw than genuine inferences.

#### DISCUSSION

The present findings give considerable support for the analysis of *or* as *and* plus a negative. More specifically, they favor the proposal that  $p$  or  $q$  is represented mentally as *not* ( $p$  and  $q$ ). The first requirement was to demonstrate that our subjects actually did construe the present pseudoimperatives as biconditionals, and this was shown by the fact that these subjects made no more errors on invited than genuine inferences, even during practice. Given this prerequisite there are three pieces of evidence that support various aspects of this analysis. First,  $p$  or  $q$  took an average of about 1500 msec longer than  $p$  and  $q$ . This result is consistent with the typical finding that sentences with negatives take longer to comprehend on the average than the corresponding sentences without. Second, the pattern of  $p$ - and  $q$ -differences for  $p$  or  $q$  was different from that for *not*- $p$  and  $q$  and  $p$  and *not*- $q$ , which have negatives directly attached to the  $p$  and  $q$  conjuncts, respectively, but its pattern was similar to that for  $p$  and  $q$ , which has no negative attached to either conjunct. By exclusion, the assumed negative in  $p$  or  $q$  must be attached to the biconditional con-

nective, making it the denial of  $p$  and  $q$ , i.e., *not* ( $p$  and  $q$ ). Finally, the absolute magnitudes of the  $p$ - and  $q$ -differences showed that  $p$  or  $q$  requires Rule 2 in its Stage 3 comparison process in exactly the manner that *not*- $p$  and  $q$  and  $p$  and *not*- $q$  do. Rule 2, it will be recalled, reduces double negatives to positives, taking an extra increment of time  $r$ . This evidence argues that the presumed negative implicit in  $p$  or  $q$  behaves in the same way as do the explicit negatives in cover sentences B and C.

This third piece of evidence, in effect, rules out an alternative model in which *and* and *or* are treated by separate rules, thereby obviating the need for the negative in the representation of *or* sentences. In this alternative, Rule 1 would be broken up into two rules, one for *and* and one for *or*:

Rule 1': If the sentence contains *and* and if the Stage 2 code does not match the polarity of the corresponding conjunct of the Stage 1 code, then add a negative to the Answer Index.

Rule 1'': If the sentence contains *or* and if the Stage 2 code *does* match the polarity of the corresponding conjunct of the Stage 1 code, then add a negative to the Answer Index.

This pair of rules would then be followed by Rule 2, as before. But the addition of Rule 1'' to take care of *or* would have the effect of eliminating the use of Rule 2—the reduction of double negatives, requiring an increment  $r$ —for sentence D, and this would predict the  $p$ - and  $q$ -differences to be the same for D as for A. This prediction is clearly confuted by the results. This prediction was generated on the assumption that increment  $m$  is required whenever Rule 1' or 1'' detects a mismatch, an assumption consistent with previous work by Clark and Chase (1972), Trabasso *et al.* (1971), and Seymour (1969). But if the increment  $m$  is assumed to be required whenever Rule 1' or 1'' adds a negative to the Answer Index, or if positive increments are assumed to be required both for detecting mismatches and

for adding a negative to the Answer Index, then this model generates even worse predictions; for example, the  $p$ - and  $q$ -differences should be even less for D than for A, a prediction directly counter to the results.

There is a critical difference between the two rules of the present model and two rules of Clark and Chase's (1972) model for the verification of negative sentences, even though both pairs of rules carry out nearly the same function. If the Clark and Chase rules were adjusted to fit the present model, they would result in only a single rule: If the Stage 2 code does not match the polarity of the corresponding conjunct of the Stage 1 code, then change the polarity of the Answer Index. This would have the effect of adding a negative when the Answer Index was positive and subtracting a negative when it was negative. But if this rule takes a constant increment of time  $m$ , then it predicts that the  $p$ - and  $q$ -differences for sentences A, B, C, and D should all have the same absolute value. This prediction, however, is clearly disconfirmed by the present results. The addition of Rule 2 in the present model effectively claims that changing the polarity of the Answer Index has two time increments associated with it, not just one. It takes  $m$  amount of time to change the Answer Index from positive to negative, but  $m+r$  to change it from negative to positive— $m$  for the mismatch and adding the negative (Rule 1) and  $r$  for reducing the double negative (Rule 2). Empirically this extra parameter  $r$  appears to be required in the present results, but not for the Clark and Chase results. At this time it is unclear why these two studies differ: the difference could lie in any number of factors that varied between the two studies.

Pseudoimperatives, we noted above, can be thought of as consisting of two conditionals, one genuinely asserted and the other merely invited. For  $p$  and  $q$  these are  $p \rightarrow q$  and  $q \rightarrow p$ , respectively. Since our subjects took no longer to make invited than genuine inferences, this suggests that the two conditionals are repre-

sented in codes that are equally accessible. The model proffered above satisfies this requirement by treating pseudoimperatives as if the two conditionals were amalgamated into a single biconditional in which there is no distinction between genuine and invited inferences. The adequacy of this solution is threatened, however, when one considers that people are able to distinguish genuine from invited inferences. Invited inferences can be revoked (or uninvited), as in *Sit down and I'll scream, but even if you don't sit down, I'll scream*; on the other hand, genuine inferences cannot be revoked, as can be seen in the unacceptable *Sit down and I'll scream, but if you do sit down, I won't scream*. If only *and* sentences were at issue, there would be no problem, for one could imagine  $p$  and  $q$  being represented by two equally accessible but differently labeled conditionals—"the literal meaning":  $p \rightarrow q$ , and "the invited meaning":  $q \rightarrow p$ . This solution, however, will not work for *or* sentences. That is,  $p$  or  $q$  would have to be given the representation "literal meaning":  $\sim p \rightarrow q$  and "invited meaning":  $q \rightarrow \sim p$ , whereas both of these representations are at variance with the evidence showing that the negative in  $p$  or  $q$  is attached to neither the  $p$  nor the  $q$ . Obviously, if the present evidence is correct, there will have to be a quite different resolution to the problem.

One plausible resolution is to distinguish, in the manner of Gordon and Lakoff (1971), the literal meaning of a sentence from its conversationally conveyed meaning. As they have noted, many sentences convey a meaning that is quite distinct from their literal meaning. *Can you open the window?*, for example, has the literal meaning "Are you able to open the window?" (a question), though in many contexts it *conveys* the meaning "Please open the window" (a request). According to Gordon and Lakoff, if the addressee is in a context where it is obvious to both him and the speaker that he is able to open the window, the addressee considers the literal meaning ("Does the addressee have the ability?") and the context

("The addressee obviously has the ability") together, applies a conversational postulate (roughly, "Whenever the speaker questions the addressee's obvious ability to do something, the addressee is being requested to carry out that action"), and thereby comes to construe the question as a polite request to open the window. Interestingly, such conveyed meanings can be revoked or averted as in "Can you open the window—if I tie both of your hands?" Here, the second clause overrides the addressee's assumptions about the context, and therefore the addressee can give the question a literal interpretation only. Second, while the conveyed meaning is related to the literal meaning, the literal meaning is not necessarily preserved "intact" in the conveyed meaning.

A similar analysis can be given for pseudo-imperatives. For *Sit down and I'll scream* ( $p$  and  $q$ ), for example, the addressee would consider the literal meaning ( $p \rightarrow q$ ) together with the context, apply a conversational postulate, and come to construe the sentence as a biconditional ( $p \leftrightarrow q$ ). Speculatively, we might characterize the intervening postulate as something like this: "If a speaker advises that doing  $p$  is sufficient for  $q$  to occur, then assume doing  $p$  is both necessary and sufficient for  $q$  to occur." Such a postulate, we propose, is invoked in a wide class of contexts: that is, whenever (1) the addressee considers the speaker to be "cooperative" (see Gordon and Lakoff for some discussion), and (2) the addressee has no reason to believe the contrary of the postulate. (If, for example, the speaker revokes the invited inference, the conversational postulate will be blocked as condition (2) is not met.) Our results imply, moreover, that once the addressee has judged the context to be appropriate, he routinely constructs the conveyed meaning, dropping the literal meaning, and then uses this representation for any subsequent deductions. This characterization, then, can reconcile the apparently paradoxical result that (1) our subjects represented  $p$  or  $q$  as  $\sim(p \leftrightarrow q)$ , which is not simply a conjunction of

the literal and invited meaning, yet (2) the so-called invited meaning can be blocked or revoked leaving only the literal meaning. The one residual problem is how *not-p and q* and *p or q* come to have distinct conveyed meanings ( $\sim p \leftrightarrow q$ ) and  $\sim(p \leftrightarrow q)$ , respectively, although they appear to share the same literal meaning ( $\sim p \rightarrow q$ ). To account for this, conversational postulates must, evidently, be made sensitive to the surface structure position of negatives, or else the literal meanings of *not-p and q* and *p or q* must be given distinguishing features not formalizable within a predicate calculus notation (e.g., *or*-sentences might be additionally marked as threats). Regardless of the resolution of this aspect of the problem, conversational postulates would seem to be necessary for describing how people interpret a wide range of prosaic sentences—pseudo-imperatives included.

Although there have been previous studies on biconditionals (equivalent to  $p$  and  $q$ ) and exclusive disjunctions (equivalent to  $p$  or  $q$ ), they are not directly comparable to the present study. Neisser and Weene (1962), Haygood and Bourne (1965), and Bourne (1970), for example, required subjects to discover concepts from instances; Trabasso *et al.* (1971) required subjects to verify instances against concepts; and Wason and Johnson-Laird (1972) used simple conditionals and required subjects to select instances that would test a conditional rule. Although the tasks were quite different from the present deduction task, they at least had in common with the present study that they attempted to decompose the connectives into simpler processes. In particular, the elementary operations underlying the Trabasso *et al.* study are almost identical to those underlying the Clark and Chase (1972) work on which the present study is based. But there are fundamental differences in the connectives studied. The *and* and *or* in pseudo-imperatives, for example, are asymmetrical, whereas those in Trabasso *et al.* study were symmetrical. The sentence *Sit down or I'll scream*, for instance, does not mean the same

thing as *I'll scream or (you'll) sit down*, if the latter sentence makes sense at all, and therefore the conjunctions are asymmetrical in their function (see Fillenbaum, 1971; Johnson-Laird, 1967, 1969; Staal, 1968, for further discussions of symmetrical and asymmetrical *and*). Also, Trabasso *et al.* found evidence that their subjects were not able to refrain from interpreting the experimentally defined exclusive *or* as an inclusive *or*, and this makes their results even more difficult to compare to ours. One can only conclude that there is much left to learn about logical connectives like *and* and *or* and their use in natural language.

## REFERENCES

- BOURNE, L. E. Knowing and using concepts. *Psychological Review*, 1970, 77, 546-556.
- CHASE, W. G., & CLARK, H. H. Mental operations in the comparison of sentences and pictures. In Gregg, L. (Ed.) *Cognition in learning and memory*. New York: Wiley, 1972.
- CLARK, H. H. How we understand negation. Paper presented at COBRE Workshop on Cognitive Organization and Psychological Processes. Huntington Beach, California, August 19, 1970.
- CLARK, H. H. Semantics and comprehension. In Sebeok, T. A. (Ed.), *Current trends in linguistics, Vol. 12: Linguistics and adjacent arts and sciences*. The Hague: Mouton, 1973.
- CLARK, H. H., & CHASE, W. G. On the process of comparing sentences against pictures. *Cognitive Psychology*, 1972, 3, 472-517.
- FILLENBAUM, S. On coping with ordered and unordered conjunctive sentences. *Journal of Experimental Psychology*, 1971, 87, (1), 93-98.
- FODOR, J. D. Formal linguistics and formal logic. In J. Lyons, (Ed.), *New horizons in linguistics*. Harmondsworth, Middlesex, England: Penguin Books, 1970.
- GORDON, D., & LAKOFF, G. Conversational postulates. In: *Papers from the Fifth Regional Meeting of the Chicago Linguistics Society*. Chicago: University of Chicago Press, 1971.
- GEIS, M., & ZWICKY, A. M. On invited inferences. *Linguistic Inquiry*, 1971, 2, 561-566.
- HAYGOOD, R. C., & BOURNE, L. E. Attribute- and rule-learning aspects of conceptual behaviour. *Psychological Review*, 1965, 72, 175-195.
- JESPERSEN, O. *A modern English grammar*, Part V, Section 24.3. London: Allen and Unwin, 1940.
- JOHNSON-LAIRD, P. N. Katz on analyticity. *Journal of Linguistics*, 1967, 3, 82.
- JOHNSON-LAIRD, P. N. "&". *Journal of Linguistics*, 1969, 6, 111-114.
- JUST, M., & CARPENTER, P. Comprehension of negation and quantification. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 244-253.
- JUST, M. A., & CLARK, H. H. Drawing inferences from the presuppositions and implications of affirmative and negative sentences. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 21-31.
- KARTTUNEN, L. Implicative verbs. *Language*, 1971, 47, 340-358.
- LAKOFF, G. Stative adjectives and verbs in English. *Mathematical linguistics and automatic translation*, Harvard University, NSF-17, 1966.
- LAKOFF, G. Linguistics and natural logic. *Synthese*, 1970, 22, 151-271.
- LAKOFF, R. If's and's and but's about conjunction. In C. J. Fillmore, and D. T. Langendoen, (Eds.), *Studies in Linguistic Semantics*. New York: Holt, Rinehart, and Winston, 1971.
- NEISSER, U., & WEENE, P. Hierarchies in concept attainment. *Journal of Experimental Psychology*, 1962, 64, 640-645.
- NOORDMAN, L. G. M. On the interpretation of conditional sentences. *Heymans Bulletins*, 1972, NR: HB-72-118 EX, Psychologische Instituut, Groningen University, Groningen, The Netherlands.
- SEYMOUR, P. H. K. Response latencies in classification of word-shape pairs. *British Journal of Psychology*, 1969, 60, 312-321.
- STAAL, J. F. "And." *Journal of Linguistics*, 1968, 4, 79-81.
- TAPLIN, J. E. Reasoning with conditional sentences. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 219-225.
- TRABASSO, T., ROLLINS, H., & SHAUGHNESSY, E. Storage and verification stages in processing concepts. *Cognitive Psychology*, 1971, 2, 239-289.
- WASON, P. C., & JOHNSON-LAIRD, P. N. *Psychology and reasoning: Structure and content*. London: Batsford, 1972.
- WASON, P. C., & JONES, S. Negatives: denotation and connotation. *British Journal of Psychology*, 1963, 54, 299-307.
- YOUNG, R., & CHASE, W. G. Additive stages in the comparison of sentences and pictures. Paper presented at Midwestern Psychological Association meetings, Chicago, April 1971.
- ZWICKY, A. M. Linguistics as chemistry: the substance theory of semantic primes. *Working papers in linguistics*, 1971, 8, 111-135.

(Received October 20, 1972)